Martingale

Chia-Min Wei

July 20, 2024

1 Martingale

Definition 1 (Filtration): We say that a sequence of sub-σ-algebras $\{\mathcal{F}_n\}_{n=1}^m$ defined over a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ is a *filtration* if $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for any $1 \leq n \leq m$. If the sequence is infinite, then we simply set $m = +\infty$. In this note, the notation $\{\mathcal{F}_n\}_{n\geq 1}$ and $\{\mathcal{F}_n\}_{n=1}^\infty$ is equivalent.

A filtration represents a gambler's information after the nth round of play. The requirement that $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ means that the gambler's information increases over time. Next we model the fortune of the gambler over time. After the nth round of play, the gambler knows his fortune, and therefore his information at time n, $\mathcal{F}_n$, should contain such information. This leads us to the notion of *adapted sequence*.

Definition 2 (Adapted sequence): Let $\{X_n\}_{n=0}^m$ be a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$. $\{X_n\}_{n=0}^m$ is said to be *adapted* to the filtration $\{\mathcal{F}_n\}_{n=0}^m$ if X_n is $\mathcal{F}_n-$measurable. We will call $\{X_n, \mathcal{F}_n\}_{n=0}^m$ an adapted sequence.

Definition 3 (Martingale): An adapted sequence $\{X_n\}_{n=0}^m$ with $\mathrm{E}\,|X_n| < \infty$ for all $0 \leq n \leq m$ is said to be a

(i) *sub-martingale* if $\mathrm{E}[X_{n+1}\,|\,\mathcal{F}_n] \geq X_n$ for all $1 \leq n \leq m$.

(ii) *super-martingale* if $\mathrm{E}[X_{n+1}\,|\,\mathcal{F}_n] \leq X_n$ for all $1 \leq n \leq m$.

(iii) *martingale* if $\mathrm{E}[X_{n+1}\,|\,\mathcal{F}_n] = X_n$ for all $1 \leq n \leq m$.

$\{X_n, \mathcal{F}_n\}$ is a sub-martingale, if given the gambler's information at time n, his fortune is, on average, going to increase after the next round of play. This means that the gambler is playing a *favorable* game. Similarly, a super-martingale represents an unfavorable game.

There is nothing super about the super-martingale. The name is related to superharmonic functions. Think about random walks on $\mathbb{R}^d$. Fix a radius $r > 0$ and let $\{X_n\}_{n\geq 1}$ be a random walk defined by

$$f_{X_{n+1}}(y\,|\,x_n, x_{n-1}, ..., x_0) = f_{X_{n+1}}(y\,|\,x_n) = \begin{cases} \frac{1}{r^{d-1}\omega_d} & \text{if } |x - x_n| = r, \\ 0 & \text{otherwise.} \end{cases}$$

where ω_d is the surface area of an R^d ball with radius 1. Let ϕ be a superharmonic function defined on $\mathbb{R}^k$.

Then the sequence $\{\phi(X_n), \mathcal{F}_n\}$ is a super-martingale.

Given a sequence of random variables $\{X_n\}_{n=1}^m$, there is a natural way of choosing the filtration to which $\{X_n\}$ is adapted: setting $\mathcal{F}_n = \sigma(X_0, ..., X_n)$. Indeed, if $\{X_n\}$ is a martingale with respect to some filtration, then it must be a martingale with respect to the natural filtration.

Example 1.1 (Branching Process): Let $\{\xi_{nk} : n \geq 1, k \geq 1\}$ be a double array of i.i.d random variables with $\mathrm{E}[\xi_{nk}] = \mu$ and Z_n be the size of the population of generation n. Consider the evolution of a population starting from a single person.

$$Z_0 = 1, \;\; Z_n = \sum_{k=1}^{Z_{n-1}} \xi_{nk}.$$

$\{Z_n\}_{n\geq 0}$ is called a *branching process*. Let $\mathcal{F}_n = \sigma(X_1, ..., X_n)$.

$$\mathrm{E}[Z_{n+1} \,|\, \mathcal{F}_n] = \mu Z_n.$$

Therefore, $\{Z_n, \mathcal{F}_n\}_{0=1}^{\infty}$ is a martingale, sub-martingale, super-martingale if and only if $\mu = 1$, ≥ 1 or ≤ 1.

Example 1.2 (Likelihood Ratio): Let $Y_1, Y_2, ...$ be a sequence of random variables on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$. Let Q be another probability measure defined on $\mathcal{F}$. Suppose for any n, the joint distribution of $(Y_1, ..., Y_n)$ under the two measures are both dominated by the Lebesgue measure with derivatives p_n and q_n. Consider the sequence $\{Z_n\}_{n\geq 1}$

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

and let $\mathcal{F}_n = \sigma(Y_1, ..., Y_n)$. Then $\{Z_n, \mathcal{F}_n\}_{n\geq 1}$ is a martingale under P. First,

$$\mathrm{E}\,|Z_n| = \int_{\mathbb{R}^n} \frac{q_n(y_1, ..., y_n)}{p_n(y_1, ..., y_n)} p_n(y_1, ..., y_n) d(y_1, ..., y_n) = 1 < \infty.$$

Now let $A \in \mathcal{F}_n$. Then $A = \{\omega : (Y_1(\omega), ..., Y_n(\omega)) \in B\}$ for some Borel set B in $\mathbb{R}^n$.

$$\begin{aligned}
\mathrm{E}[\mathrm{E}[Z_{n+1} \,|\, \mathcal{F}_n]\mathbf{1}_A] &= \mathrm{E}[Z_{n+1}\mathbf{1}_A] \\
&= \int_{B\times\mathbb{R}} \frac{q_{n+1}(y_1, ..., y_{n+1})}{p_{n+1}(y_1, ..., y_{n+1})} p_{n+1}(y_1, ..., y_{n+1})\, d(y_1, ..., y_n, y_{n+1}) \\
&= \int_B q_n(y_1, ..., y_n)\, d(y_1, ..., y_n) = \int_B \frac{q_n(y_1, ..., y_n)}{p_n(y_1, ..., y_n)} p_n(y_1, ..., y_n)\, d(y_1, ..., y_n) \\
&= \mathrm{E}[Z_n \mathbf{1}_A].
\end{aligned}$$

Therefore, $\mathrm{E}[Z_{n+1} \,|\, \mathcal{F}_n] = Z_n$.

2 Betting and Optional Stopping

Does there exist a betting strategy that allows one to "beat" a fair or unfavorable game? Suppose an adapted sequence $\{X_n, \mathcal{F}_n\}_{n=0}^m$ represents the fortune of a gambler playing the original game. At each nth round of play, he gains $X_{n+1} - X_n$. We now allow the gambler, after nth round of play, to choose how much to bet at the next round. By choosing $H_{n+1} \geq 0$, instead of gaining $X_{n+1} - X_n$, he gains $H_{n+1}(X_{n+1} - X_n)$. His betting strategy can only be based on his available information. Namely, for any two situations that the gambler cannot tell apart, the bet should be the same. Therefore, H_{n+1} should be measurable with respect to $\mathcal{F}_n$. This leads us to the notion of betting sequence.

Definition 4 (Betting Sequence): A *betting sequence* with respect to a filtration $\{\mathcal{F}_n\}_{n=0}^m$ is a sequence of nonnegative random variables $\{H_n\}_{n=1}^m$ such that H_n is measurable with respect to $\mathcal{F}_{n-1}$.

Example 2.1 (Martingale Strategy): In 19th century, there is a popular betting strategy named *martingale*. The betting strategy had a gambler double his bet whenever he loses a round of game and stop when he wins the first round. Consider the most simple form of game: the gambler wins 1 dollar if a fair coin comes up head and loses 1 dollar otherwise. Let $\{G_n\}_{n\geq 1}$ be the gambler's gain at each round of play. Then his fortune X_n after nth round forms a martingale (let $X_0 = 0$), defined by $X_n = X_0 + \sum_{j=1}^n G_j$ with respect to the natural filtration $\mathcal{F}_n = \sigma(X_0, ..., X_n) = \sigma(G_1, ..., G_n)$. The martingale strategy suggests a betting sequence $\{H_n\}_{n\geq 1}$ and an induced sequence of fortune $\{Y_n\}_{n\geq 0}$ defined as follows: $H_1 = 1$, and for $n \geq 2$,

$$Y_n = Y_0 + \sum_{j=1}^n H_j G_j$$

$$H_n = \begin{cases} 2|Y_{n-1}| & \text{if } G_j = -1 \text{ for all } j \leq n-1 \\ 0 & \text{if } G_j = 1 \text{ for some } j \leq n-1. \end{cases}$$

The following theorem says that the gambler cannot beat a fair game(?)

Theorem 1 (Betting Theorem): Let $\{X_n, \mathcal{F}_n\}_{n=0}^m$ be an adapted sequence and $\{H_n\}_{n=1}^m$ be a betting sequence. Define $Y_0 = X_0$ and $Y_n = X_0 + \sum_{j=1}^n H_j(X_j - X_{j-1}) = Y_{n-1} + H_n(X_n - X_{n-1})$. Then $\{X_n, \mathcal{F}_n\}_{n=0}^m$ is a

(i) martingale, then $\{Y_n, \mathcal{F}_n\}_{n=0}^m$ is also a martingale.

(ii) sub-martingale, then $\{Y_n, \mathcal{F}_n\}_{n=0}^m$ is also a sub-martingale.

(iii) super-martingale, then $\{Y_n, \mathcal{F}_n\}_{n=0}^m$ is also a super-martingale.

Proof. It is easy to see that $\{Y_n\}_{n=0}^m$ is adapted to $\{\mathcal{F}_n\}_{n=0}^m$ and

$$\mathrm{E}\,|Y_n| \leq \mathrm{E}\,|X_0| + \sum_{j=1}^n H_j(\mathrm{E}\,|X_j| + \mathrm{E}\,|X_{j-1}|) < \infty.$$

Now suppose $\{X_n, \mathcal{F}_n\}_{n=1}^m$ be a sub-martingale.

$$\begin{aligned} \mathrm{E}[Y_{n+1} \,|\, \mathcal{F}_n] &= \mathrm{E}[Y_n + H_{n+1}(X_{n+1} - X_n) \,|\, \mathcal{F}_n] \\ &= Y_n + H_{n+1}(\mathrm{E}[X_{n+1} \,|\, \mathcal{F}_n] - X_n) \\ &\geq Y_n. \end{aligned}$$

□

But wait! Is this theorem enough to claim that the gambler cannot beat a fair game? Suppose there is a large number I of gamblers (indexed by i) who play independently the coin flipping game and they all adopt the martingale strategy $\{H_n\}_{n\geq 1}$. Let $\{Y_{in}\}_{n\geq 0}$ be gambler i's fortune after each round n, and suppose $Y_{i0} = 0$ for all i. Then Theorem 1 suggests that $(1/I)\sum_{i=1}^{I} Y_{in} \approx E[Y_{1n}] = 0$. But why should we care about the average fortune of bidders at any particular time n instead of the average of their final fortune when they stop playing?

Certainly, with probability 1, the gambler is going to win. The martingale strategy involves *stopping* at some particular point when some condition is satisfied. This leads to the notion of *stopping time*.

Definition 5 (Stopping Time): Let T be a random variable taking values in $\mathbb{N} \cup \{0\} \cup \{+\infty\}$, defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ with a filtration $\{\mathcal{F}_n\}_{n\geq 0}$. We say that T is a *stopping time* if $\{T = n\} \in \mathcal{F}_n$ for all n. $T(\omega) = +\infty$ if $T(\omega) \notin \mathbb{N} \cup \{0\}$. Note that T is a stopping time if and only if $\{T \leq n\} \in \mathcal{F}_n$ for all n if and only if $\{T > n\} \in \mathcal{F}_n$. We say that T is a *proper stopping time* if $\mathrm{P}(T \neq \infty) = 1$.

In Example 2.1, the martingale strategy involves a stopping time defined by $T = \inf\{n > 0 : Y_n > 0\}$. What we really care, instead of $\mathrm{E}[Y_n]$ for some particular n, is $\mathrm{E}[Y_T]$. Certainly, $\mathrm{E}[Y_T] > 0$. However, this cannot happen in the real world as the following theorems suggest.

Theorem 2 (Doob's Optional Stopping Theorem I): Let $\{X_n, \mathcal{F}_n\}_{n\geq 0}$ be a sub-martingale and let T be a stopping time w.r.t $\{\mathcal{F}_n\}_{n\geq 0}$. Define $\tilde{X}_n = X_{\min\{T,n\}}$. Then $\{\tilde{X}_n, \mathcal{F}_n\}$ is also a sub-martingale and hence $\mathrm{E}[\tilde{X}_n] \geq \mathrm{E}[X_0]$.

Proof. Define for all $n \geq 1$, $H_n = 1$ if $T > n-1$ and $H_n = 0$ if $T \leq n-1$, Notice that $\{H_n = 0\} = \{T \leq n-1\} \in \mathcal{F}_{n-1}$. Therefore, $\{H_n\}_{n\geq 1}$ is a betting sequence. Observe that $\tilde{X}_n = X_0 + \sum_{j=1}^{n} H_j(X_j - X_{j-1})$. By Theorem 1, the result follows. □

Notice that if T is a proper stopping time, then $\tilde{X}_n \xrightarrow{a.s.} X_T$ as $n \to \infty$. But it does not guarantee that $\mathrm{E}[\tilde{X}_n] \longrightarrow \mathrm{E}[X_T]$. In Example 2.1, stopping time $T = \inf\{n \geq 1 : Y_n > 0\}$ and $\tilde{Y}_n = Y_n$. We see that $Y_n \xrightarrow{a.s.} Y_T$, but $\mathrm{E}[Y_n] = 0$ while $\mathrm{E}[Y_T] > 0$.

Theorem 3 (Doob's Optional Stopping Theorem II): Let $\{X_n, \mathcal{F}_n\}_{n\geq 0}$ be an adapted sequence and let T be a proper stopping time with respect to $\{\mathcal{F}_n\}_{n\geq 0}$. If for all n, $|X_{\min\{T,n\}}| < K$ for some K, then

(i) $\mathrm{E}[X_T] \geq \mathrm{E}[X_0]$ if $\{X_n, \mathcal{F}_n\}_{n\geq 0}$ is a sub-martingale.

(ii) $\mathrm{E}[X_T] \leq \mathrm{E}[X_0]$ if $\{X_n, \mathcal{F}_n\}_{n\geq 0}$ is a super-martingale.

(iii) $\mathrm{E}[X_T] = \mathrm{E}[X_0]$ if $\{X_n, \mathcal{F}_n\}_{n\geq 0}$ is a martingale.

Proof. Suppose $\{X_n, \mathcal{F}_n\}_{n\geq 0}$ is a sub-martingale. Since $X_{\min\{T,n\}} \xrightarrow{a.s.} X_T$, by Dominated Convergence Theorem, $\mathrm{E}[X_T] = \lim_{n\to\infty} \mathrm{E}[X_{\min\{T,n\}}] \geq \mathrm{E}[X_0]$. □

Example 2.2 (Martingale Strategy continued): In real life, the martingale strategy is not applicable because the gambler has limited fund. The betting sequence $\{H_n\}_{n\geq 1}$ and stopping time T should be revised:

$H_1 = 0$, and for some $A > 0$,

$$H_n = \begin{cases} 2|Y_{n-1}| & \text{if } G_j = -1 \text{ and } Y_j \geq A \text{ for all } j \leq n-1 \\ 0 & \text{if } G_j = 1 \text{ or } Y_j < -A \text{ for some } j \leq n-1. \end{cases}$$

$$T = \inf\{n \geq 1 : G_n = 1 \text{ or } Y_n < -A\}.$$

The induced sequence of fortune $\{Y_n\}_{n\geq 1}$ is still a martingale. Since it is uniformly bounded by $|Y_n| \leq 3A$, $\mathrm{E}[Y_T] = \mathrm{E}[Y_0] = 0$.

Suppose we have two stopping times S and T with $S \leq T$, how could we guess X_T at time S (which is random)? We shall first know the information available at time S. Although at any time n, you can always distinguish whether $\{S = n\}$ has occurred (by the definition of stopping time), knowing $S = n$ still brings additional information!

Definition 6 ($\mathcal{F}_T$): Let $\mathcal{F}_\infty = \sigma(\bigcup_{n\geq 0})$ and T be a stopping time. Define the gambler's information when he stops playing as $\mathcal{F}_T$,

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : A \cap \{T = n\} \in \mathcal{F}_n \text{ for all } n\}.$$

One can check that $\mathcal{F}_T$ is indeed a σ-algebra. Also, X_T is $\mathcal{F}_T$-measurable. And if $S \leq T$, then $\mathcal{F}_S \subset \mathcal{F}_T$.

Example 2.3 (Trivial Stopping Time): Let stopping time T be defined $T \equiv m$, namely, always stop playing after the mth round of play. Notice that $\{T = n\} = \varnothing$ if $n \neq m$ and $\{T = m\} = \Omega$ if $n = m$. Hence $\mathcal{F}_T = \mathcal{F}_m$.

Theorem 4 (Doob's Optional Stopping Theorem III): Let $\{X_n, \mathcal{F}_n\}_{n\geq 0}$ be a sub-martingale (super-martingale), and S and T be two stopping times such that $S \leq T$. If X_S and X_T are integrable, and if

$$\liminf_{n\to\infty} \mathrm{E}[|X_n|\mathbf{1}_{\{T>n\}}] \longrightarrow 0, \tag{1}$$

then

$$\mathrm{E}[X_T \,|\, \mathcal{F}_S] \geq (\leq) X_S.$$

If, in addition, $\{X_n\}_{n\geq 0}$ is a martingale, then the equality holds.

Proof. It suffices to prove that for any $A \in \mathcal{F}_T$,

$$\int_A X_T - X_S \, d\mathrm{P} \geq 0.$$

Let $\{n_k\}_{k\geq 1}$ be a subsequence along which the limit in Equation 1 is reached and define $T_k = \min\{T, n_k\}$

and $S_k = \min\{S, n_k\}$. Note that for all $k \geq 1$, $S_k \leq T_k \leq n_k$. It suffices to prove

$$\int_A X_{T_k} - X_{S_k}\, d\,\mathrm{P} \geq 0 \text{ for all } k \geq 1,$$
$$\int_A |X_{T_k} - X_T|\, d\,\mathrm{P} \longrightarrow 0, \quad \int_A |X_{S_k} - X_S|\, d\,\mathrm{P} \longrightarrow 0.$$

We may write

$$X_{T_k} - X_{S_k} = \sum_{n=1}^{n_k} (X_n - X_{n-1}) \mathbf{1}\{S_k + 1 \leq n \leq T_k\}.$$

For any $n \leq n_k$, $\{S_k + 1 \leq n\} = \{S \leq n-1\}$, while $\{n \leq T_k\} = \{n \leq T\} = \{T > n-1\}$. Since $A \in \mathcal{F}_S$, $A \cap \{S \leq n-1\} \in \mathcal{F}_{n-1}$. $A \cap \{S_k + 1 \leq n\} \cap \{n \leq T_k\} \in \mathcal{F}_{n-1}$. Define $B_n = A \cap \{S_k + 1 \leq n\} \cap \{n \leq T_k\}$,

$$\int_A X_{T_k} - X_{S_k}\, d\,\mathrm{P} = \sum_{n=1}^{n_k} \int_{B_n} X_n - X_{n-1}\, d\,\mathrm{P} \geq 0,$$

because $\int_{B_n} X_n\, d\,\mathrm{P} = \int_{B_n} \mathrm{E}[X_n \,|\, \mathcal{F}_{n-1}]\, d\,\mathrm{P}$.

$$\begin{aligned}
\int_A |X_{T_k} - X_T|\, d\,\mathrm{P} &= \int_A |X_{n_k} - X_T| \mathbf{1}_{\{T > n_k\}}\, d\,\mathrm{P} \\
&\leq \int_A |X_{n_k}| \mathbf{1}_{\{T > n_k\}}\, d\,\mathrm{P} + \int_A |X_T| \mathbf{1}_{\{T > n_k\}}\, d\,\mathrm{P} \\
&\longrightarrow 0.
\end{aligned}$$

□

Remark: If there exists $n_0 < \infty$ such that $\mathrm{P}\{T < t_0\} = 1$, then Equation 1 holds.

Using Theorem 4, we can derive a Markov-type inequality for martingales.

Theorem 5 (Doob's Maximal Inequality): Let $\{X_n, \mathcal{F}_n\}_{n \geq 0}$ be a sub-martingale and for each $0 \leq m$, let $M_m = \max\{X_1, ..., X_m\}$. Then for any $m \geq 0$ and $x \in (0, \infty)$,

$$\mathrm{P}(M_m > x) \leq \frac{\mathrm{E}[X_m \mathbf{1}_{\{M_m > x\}}]}{x} \leq \frac{\mathrm{E}[X_m^+]}{x}.$$

Proof. Define the stopping time S by:

$$S = \begin{cases} \inf\{0 \leq n \leq m : X_n > x\} & \text{on } A = \{M_m > x\}, \\ m & \text{on } A^c, \end{cases}$$

and T by $T \equiv m$. Since $S \leq T$ and $\mathrm{P}\{T \leq m\} = 1$, by Theorem 4, $\mathrm{E}[X_m \,|\, \mathcal{F}_S] = \mathrm{E}[X_T \,|\, \mathcal{F}_S] \leq X_S$. Note that for any $0 \leq n \leq m-1$, $A \cap \{S = n\} = \{S = n\} \in \mathcal{F}_n$ and $A \cap \{S = m\} = \{X_0, ..., X_{m-1} \leq x, X_m > x\} \in \mathcal{F}_m$. Therefore, $A \in \mathcal{F}_S$.

$$\mathrm{P}(A) \leq \frac{\int_A X_S\, d\,\mathrm{P}}{x} \leq \frac{\int_A X_T\, d\,\mathrm{P}}{x} = \frac{\mathrm{E}[X_m \mathbf{1}_A]}{x} \leq \frac{\mathrm{E}[X_m^+]}{x}.$$

□

Theorem 6 (Doob's L^p-Maximal Inequality for Sub-Martingales): Let $\{X_n, \mathcal{F}_n\}_{n\geq 0}$ be a sub-martingale and let $M_n = \{X_j : 1 \leq j \leq n\}$, Then for any $p \in (1, \infty)$,

$$\mathrm{E}\left[(M_n^+)^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathrm{E}\left[(X_n^+)^p\right] \leq \infty.$$

Proof. If $\mathrm{E}[(X_n^+)^p] = \infty$, then the inequality clearly holds. We therefore focus on the case when $\mathrm{E}\left[(X_n^+)^p\right] < \infty$. Note that for any nonnegative random variable Y, $x > 0$ and $p > 1$,

$$\begin{aligned}
\mathrm{E}[Y^p] = \mathrm{E}\left[\int_0^Y px^{p-1}\,dx\right] &= \mathrm{E}\left[\int_0^\infty px^{p-1}\mathbf{1}_{\{Y>x\}}\,dx\right] \\
&= \int_0^\infty px^{p-1}\,\mathrm{P}\{Y > x\}\,dx. && \text{(by Tonelli's Theorem)}
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathrm{E}[(M_n^+)^p] &= \int_0^\infty px^{p-1}\,\mathrm{P}(M_n > x)\,dx \\
&\leq \int_0^\infty px^{p-2}\,\mathrm{E}[X_n^+\mathbf{1}_{\{M_n>x\}}]\,dx && \text{(by Theorem 5)} \\
&= \frac{p}{p-1}\,\mathrm{E}\left[X_n^+(M_n^+)^{p-1}\right] && \text{(by Tonelli's Theorem)} \\
&\leq \frac{p}{p-1}\,\mathrm{E}\left[(X_n^+)^p\right]^{\frac{1}{p}}\,\mathrm{E}\left[(M_n^+)^p\right]^{\frac{p-1}{p}}. && \text{(by Hölder's Inequality)}
\end{aligned}$$

This implies

$$\mathrm{E}\left[(M_n^+)^p\right]^{\frac{1}{p}} \leq \frac{p}{p-1}\,\mathrm{E}\left[(X_n^+)^p\right]^{\frac{1}{p}}.$$

□

Corollary 1: Let $\{X_n, \mathcal{F}_n\}_{n\geq 0}$ be a martingale and let $\tilde{M}_n = \sup\{|X_j| : 1 \leq j \leq n\}$. Then for $p > 1$,

$$\mathrm{E}\left|\tilde{M}_n\right|^p \leq \left(\frac{p}{1-p}\right)^p \mathrm{E}\,|X_n|^p.$$

3 Martingale Convergence Theorem

Martingales converge under very mild conditions, making them highly tractable. The following are some results which will be proved in this section:

(i) Any nonnegative super-martingale converges almost surely.

(ii) Any sub-martingale $\{X_n\}_{n\geq 0}$ for which $\{\mathrm{E}\,|X_n|\}_{n\geq 0}$ is bounded converges almost surely.

(iii) Further, if the sub-martingale is nonnegative and $\{\mathrm{E}\,|X_n|^p\}_{n\geq 0}$ is bounded for some $p \in (1, \infty)$, then X_n converges almost surely as well as in L^p.

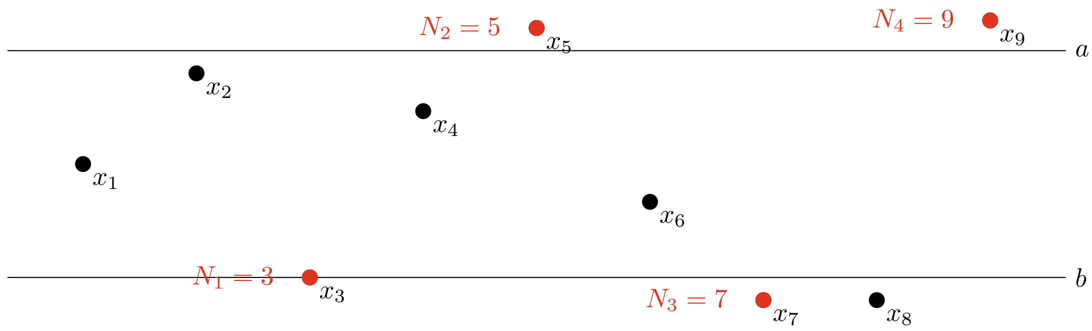

Figure 1: A sequence $\{x_i\}_{i=1}^{9}$ with two upcrossings.

These results depend crucially on **Doob's Upcrossing Lemma**. Recall that a sequence of real numbers $\{x_j\}_{j\geq 1}$ converges (may converge to $+\infty$ or $-\infty$), if and only if, for any $a, b \in \mathbb{R}$ and $a < b$, the sequence crosses from a to b only finitely many times.

Definition 7 (Upcrossings): Define, for a sequence of real numbers $\{x_j\}_{j=1}^{n}$ and $a < b$,

$$\begin{aligned} N_1(\{x_j\}_{j=1}^{n}; a, b) &\coloneqq \min\{j : 1 \leq j \leq n, x_j \leq a\}, \\ N_2 &\coloneqq \min\{j : N_1 < j \leq n, x_j \geq b\}, \\ &\vdots \\ N_{2k-1} &\coloneqq \min\{j : N_{2k-2} < j \leq n, x_j \leq a\}, \\ N_{2k} &\coloneqq \min\{j : N_{2k-1} < j \leq n, x_j \geq b\}. \end{aligned}$$

Let K be the last k such that N_{2k} is well-defined. If N_1 and N_2 are not well-defined, then $K = 0$. The number of upcrossings of $\{x_j\}_{j=1}^{n}$is defined as

$$U(\{x_j\}_{j=1}^{n}; a, b) \coloneqq K.$$

Figure 1 shows a sequence with a total of two upcrossings, $U(\{x_j\}_{j=1}^{9}; a, b) = 2$.

Proposition 1: Let $\{x_j\}_{j\geq 1}$ be a sequence of real numbers and let $U_n(a, b) \coloneqq U_n\left(\{x_j\}_{j=1}^{n}; a, b\right)$ for all $a < b$. Then $\{x_j\}_{j\geq 1}$ converges (may converge to $+\infty$ or $-\infty$) if and only if $\sup_{n\geq 1} U_n(a, b) < \infty$ for any $a < b$.

Doob's Upcrossing Lemma asserts that for a sub-martingale $\{X_j, \mathcal{F}_j\}_{j=1}^{n}$, the *expected* number of upcrossings from a to b can be bounded by above by some function of $\mathrm{E}[X_n^+]$ and a, b!

Theorem 7 (Doob's Upcrossing Lemma): Let $\{X_j, \mathcal{F}_j\}_{j=1}^{n}$ be a sub-martingale and let $a < b$ be real numbers. Let $U_n \coloneqq U(\{X_j\}_{j=1}^{n}; a, b)$. Then

$$\mathrm{E}[U_n] \leq \frac{\mathrm{E}(X_n - a)^+ - \mathrm{E}(X_1 - a)^+}{b - a} \leq \frac{\mathrm{E}\, X_n^+ + |a|}{(b - a)}.$$

Proof. First prove the case when $X_j \geq 0$ for all $j \geq 1$ and $a = 0$. Let $U_n \coloneqq U\left(\{X_j\}_{j=1}^n; 0, b\right)$. This means that $N_1\left(\{X_j\}_{j=1}^n; 0, b\right), ..., N_{2U_n}\left(\{X_j\}_{j=1}^n; 0, b\right)$ are well-defined. Define $\tilde{N}_0 \equiv 1$ and for $1 \leq j \leq n$,

$$\tilde{N}_j = \begin{cases} N_j & \text{if } j \in \{1, 2, ..., 2U_n\}, \\ n & \text{otherwise.} \end{cases}$$

One can check that $\tilde{N}_j$'s are proper stopping times with $\tilde{N}_j \leq \tilde{N}_{j+1}$ inductively. By Theorem 4,

$$\mathrm{E}\left[X_{\tilde{N}_{j+1}} - X_{\tilde{N}_j}\right] \geq 0$$

for $1 \leq j \leq n$. Therefore,

$$\begin{aligned} \mathrm{E}[X_n - X_1] &= \sum_{j=1}^{n} \mathrm{E}\left[X_{\tilde{N}_{j+1}} - X_{\tilde{N}_j}\right] \\ &= \sum_{j \text{ is odd}} \mathrm{E}\left[X_{\tilde{N}_{j+1}} - X_{\tilde{N}_j}\right] + \sum_{j \text{ is even}} \mathrm{E}\left[X_{\tilde{N}_{j+1}} - X_{\tilde{N}_j}\right] \\ &\geq \mathrm{E}[bU_n] + \sum_{j \text{ is even}} \mathrm{E}\left[X_{\tilde{N}_{j+1}} - X_{\tilde{N}_j}\right] \\ &\geq b\,\mathrm{E}[U_n], \end{aligned}$$

which proves both inequalities under the special case. Now let $\{X_j, \mathcal{F}_j\}_{j \geq n}$ be an arbitrary sub-martingale and $a < b$ be arbitrary. Define $Y_j \coloneqq (X_j - a)^+$. Then $\{Y_j, \mathcal{F}_j\}_{j=1}^n$ is a sub-martingale. Note that $U_n = U\left(\{X_j\}_{j=1}^n; a, b\right) = U\left(\{Y_j\}_{j=1}^n; 0, b - a\right)$. By the above conclusion,

$$\mathrm{E}[(X_n - a)^+] - \mathrm{E}[(X_1 - a)^+] \geq (b - a)\,\mathrm{E}[U_n].$$

The last inequality follows from the fact that $(X - a)^+ \leq X^+ + |a|$ for any random variable X. □

Lemma 1: Let $\{X_n, \mathcal{F}_n\}_{n \geq 1}$ be a sub(super)-martingale. Then

$$\sup_{n \geq 1} \mathrm{E}\,|X_n| < \infty \iff \sup_{n \geq 1} \mathrm{E}[X_n^+] < \infty \quad \left(\sup_{n \geq 1} \mathrm{E}\,|X_n| < \infty \iff \sup_{n \geq 1} \mathrm{E}[X_n^-] < \infty\right).$$

Proof. The direction ($\implies$) is obvious because $|X_n| \geq X_n^+$. We prove the opposite direction. Note that $X_n^- = X_n^+ - X_n$. Hence,

$$\begin{aligned} \sup_{n \geq 1} \mathrm{E}[X_n^-] &\leq \sup_{n \geq 1} \mathrm{E}[X_n^+] - \inf_{n \geq 1} \mathrm{E}[X_n] \\ &= \sup_{n \geq 1} \mathrm{E}[X_n^+] - \mathrm{E}[X_1] < \infty. \end{aligned}$$

□

Theorem 8 (Almost Sure Convergence of Sub-martingales): Let $\{X_n, \mathcal{F}_n\}_{n \geq 1}$ be a sub-martingale (super-martingale) such that

$$\sup_{n\geq 1} \mathrm{E}[X_n^+] < \infty \quad \left(\sup_{n\geq 1} \mathrm{E}[X_n^-] < \infty\right).$$

Then X_n converges almost surely to a finite limit X_∞ and $\mathrm{E}\,|X_\infty| < \infty$.

Proof. We first prove that

$$\limsup_{n\to\infty} X_n = \liminf_{n\to\infty} X_n$$

almost surely. Let $U_n(a,b)$ denote $U\left(\{X_j\}_{j=1}^n; a, b\right)$ and $U_\infty(a,b) = U\left(\{X_j\}_{j\geq 1}; a, b\right)$ for any $a, b \in \mathbb{R}$. It is clear that

$$U_n(a,b) \leq U_{n+1}(a,b), \;\; U_n(a,b) \longrightarrow U_\infty(a,b) \text{ as } n \to \infty.$$

Note that

$$\{\omega : U_\infty(a,b) < \infty \text{ for all } a, b \in \mathbb{Q}\} = \left\{\omega : \limsup_{n\to\infty} X_n = \liminf_{n\to\infty} X_n\right\}.$$

Fix any $a < b \in \mathbb{Q}$,

$$\begin{aligned}
\mathrm{E}[U_\infty(a,b)] &= \lim_{n\to\infty} \mathrm{E}[U_n(a,b)] && \text{(by Monotone Convergence Theorem)} \\
&\leq \frac{\sup_{n\geq 1} \mathrm{E}[X_n^+] + |a|}{b-a} && \text{(by Theorem 7)} \\
&< \infty.
\end{aligned}$$

Therefore, for any $a, b \in \mathbb{Q}$, $\mathrm{P}(U_\infty(a,b) = \infty) = 0$. And we thus have

$$\begin{aligned}
\mathrm{P}\left(U_\infty(a,b) = \infty \text{ for some } a, b \in \mathbb{Q}\right) &= \mathrm{P}\left(\bigcup_{a<b\in\mathbb{Q}} \{U_\infty(a,b) < \infty\}\right) \\
&\leq \sum_{a<b\in\mathbb{Q}} \mathrm{P}(U_\infty(a,b) < \infty) = 0,
\end{aligned}$$

which proves the assertion. We now know that $\lim_{n\to\infty} X_n = X_\infty$ exists. By Lemma 1 and Fatou's Lemma,

$$\mathrm{E}\,|X_\infty| = \mathrm{E}\left(\lim_{n\to\infty} |X_n|\right) \leq \lim_{n\to\infty} \mathrm{E}\,|X_n| \leq \sup_{n\geq 1} \mathrm{E}\,|X_n| < \infty.$$

□

Corollary 2: Let $\{X_n, \mathcal{F}_n\}_{n\geq 1}$ be a sub-martingale (super-martingale) that is bounded above (below), then X_n converges almost surely to a finite limit.

If we want convergence in L^1, more conditions are required. Recall that a collection of L^1 functions $\{f_\lambda\}_{\lambda\in\mathcal{I}}$ defined on a measured space $(\mathcal{X}, \Sigma, \mu)$ is said to be *uniformly integrable* if the following two conditions hold:

- $\sup_{\lambda\in\mathcal{I}} \int_{\mathcal{X}} |f_\lambda| \, d\mu < \infty$.
- For any $\epsilon > 0$, there exists $\delta > 0$ such that for any set A with $\mu(A) < \delta$, $\sup_{n\geq 1} \int_{\mathcal{X}} |X_n| \mathbf{1}_A \, d\mu < \epsilon$.

Lemma 2: Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ and that $\{|X_n|^p\}_{n\geq 1}$ is uniformly integrable. Suppose X_n converges in probability to a random variable X_∞. Then $X_\infty \in L^p(\mathrm{P})$ and $\{X_n\}_{n\geq 1}$ converges to X_∞ in L^p.

Proof. It suffices to prove that $\{|X_n|^p\}_{n\geq 1}$ is a Cauchy sequence since $L^p(\mathrm{P})$ is complete. Fix $\epsilon > 0$. For any $0 < n < m$, let $A(n,m) \coloneqq \{|X_n - X_\infty| > \epsilon \text{ or } |X_m - X_\infty| > \epsilon\}$.

$$\begin{aligned} \mathrm{E}\left[|X_n - X_m|^p\right] &= \mathrm{E}\left[|X_n - X_m|^p \mathbf{1}_A\right] + \mathrm{E}\left[|X_n - X_m| \mathbf{1}_{A^c}\right] \\ &\leq \mathrm{E}\left[|X_n|^p \mathbf{1}_A\right] + \mathrm{E}[|X_m|^p \mathbf{1}_A] + (2\epsilon)^p. \end{aligned}$$

By uniformly integrability, there exists $\delta > 0$ such that for any set E with $\mathrm{P}(E) < \delta$, $\mathrm{E}\left[|X_k|^p \mathbf{1}_A\right] < \epsilon$ for all k. There exists $N \in \mathbb{N}$ such that for all $n, m > N$, $\mathrm{P}(A(n,m)) < \delta$ and thus both $\mathrm{E}\left[|X_n|^p \mathbf{1}_A\right]$ and $\mathrm{E}[|X_m|^p \mathbf{1}_A]$ are smaller than ϵ. □

Theorem 9 (L^1 Convergence of Sub-martingales): Let $\{X_n, \mathcal{F}_n\}_{n\geq 1}$ be a sub-martingale (super-martingale), then the two statements are equivalent:

(i) $\{X_n\}_{n\geq 1}$ converges almost surely and in L^1 to a finite limit.

(ii) $\{X_n\}_{n\geq 1}$ is uniformly integrable.

Proof.

- (i $\implies$ ii): Indeed, for any sequence of random variables $\{X_n\}_{n\geq 1}$, convergence in L^1 implies that the sequence is uniformly integrable.
- (ii $\implies$ i): Since $\{X_n\}_{n\geq 1}$ is uniformly integrable, by Theorem 8, $\{X_n\}_{n\geq 1}$ converges almost surely to an L^1 limit X_∞. By Lemma 2, the conclusion follows.

□

Theorem 10 (L^p Convergence of Nonnegative Sub-martingales): Let $\{X_n, \mathcal{F}_n\}_{n\geq 1}$ be a non-negative sub-martingale such that $\sup_{n\geq 1} \mathrm{E}[X_n^p] < \infty$. Then $\{X_n\}_{n\geq 1}$ converges to an L^p limit X_∞ almost surely and in L^p.

Proof. Let $M_n \coloneqq \max\{X_1, ..., X_n\}$. $\{X_n^p, \mathcal{F}_n\}_{n\geq 1}$ is a nonnegative sub-martingale. By Theorem 6,

$$\mathrm{E}[M_n^p] \leq \left(\frac{p}{p-1}\right)^p \mathrm{E}[X_n^p] \leq \left(\frac{p}{p-1}\right)^p \sup_{n\geq 1} \mathrm{E}[X_n^p] < \infty.$$

Let M denote $\sup_{n\geq 1} X_n$ and $\mathrm{E}[M^p] < \infty$. Hence, for any subset $A \in \mathcal{F}$,

$$\sup_{n\geq 1} \mathrm{E}\left(X_n^p \mathbf{1}_A\right) \leq \mathrm{E}\left(M^p \mathbf{1}_A\right) \longrightarrow 0 \quad \text{as } \mathrm{P}(A) \to 0,$$

which means that $\{X_n^p\}_{n\geq 1}$ is uniformly integrable. By Theorem 8 and Lemma 2, $\{X_n\}_{n\geq 1}$ converges to an L^p limit X_∞ almost surely and in L^p. □

Corollary 3: Let $\{X_n, \mathcal{F}_n\}_{n\geq 1}$ be a martingale such that $\sup_{n\geq 1} \mathrm{E}[|X_n|^p] < \infty$. Then $\{X_n\}_{n\geq 1}$ converges to an L^p limit X_∞ almost surely and in L^p.

4 Applications

4.1 Doob Martingale and Kolmogorov's Zero-One Law

Definition 8 (Doob Martingale): Let X be an integrable random variable defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ and let $\{\mathcal{F}_k\}_{k\geq 1}$ be a filtration. Define for all $k \geq 1$,

$$X_k \coloneqq \mathrm{E}[X \mid \mathcal{F}_k].$$

One can easily see that $\{X_k, \mathcal{F}_k\}_{k\geq 1}$ is a martingale and can further check that $\{X_k\}_{k\geq 1}$ is uniformly integrable. $\{X_k, \mathcal{F}_k\}_{k\geq 1}$ is called a *Doob Margingale.*

Theorem 11 (Convergence of Doob Martingale): Let X be an integrable random variable defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$. Let $\{F_n\}_{n\geq 1}$ be a filtration and write $\mathcal{F}_\infty = \sigma(\bigcup_{n\geq 1} F_n)$. Then we have

$$\mathrm{E}[X \mid \mathcal{F}_n] \longrightarrow \mathrm{E}[X \mid \mathcal{F}_\infty] \quad \text{as} \quad n \to \infty$$

almost surely and in L^1.

Proof. Let $X_k \coloneqq \mathrm{E}[X \mid \mathcal{F}_n]$. As mentioned in Definition 8, $\{X_k, \mathcal{F}_k\}_{k\geq 1}$ forms a martingale and the sequence $\{X_k\}_{k\geq 1}$ is uniformly integrable. By Theorem 9, X_k converges to some finite limit X_∞ almost surely and in L^1. We shall prove that $X_\infty = \mathrm{E}[X \mid \mathcal{F}_\infty]$. X_∞ and $\mathrm{E}[X \mid \mathcal{F}_\infty]$ are $\mathcal{F}_\infty$-measurable. Therefore, it suffices to prove that for any $A \in \mathcal{F}_\infty$,

$$\int_A X_\infty \, d\,\mathrm{P} = \int_A \mathrm{E}[X \mid \mathcal{F}_\infty] \, d\,\mathrm{P}\,.$$

Define the collection of sets

$$\mathcal{L} = \left\{ A \in \mathcal{F}_\infty : \int_A X_\infty \, d\,\mathrm{P} = \int_A \mathrm{E}[X \mid \mathcal{F}_\infty] \, d\,\mathrm{P} \right\}.$$

One can check that $\mathcal{L}$ is a lambda system. On the other hand, it is easy to see that $\mathcal{P} = \bigcup_{k\geq 1} \mathcal{F}_k$ is a π system. For any $A_0 \in \mathcal{P}$, say $A_0 \in \mathcal{F}_n$ for some n,

$$\int_{A_0} X_k \, d\,\mathrm{P} \to \int_{A_0} X_\infty \, d\,\mathrm{P} \quad \text{as} \quad k \to \infty,$$

because X_k converges to X_∞ in L^1. But $\int_{A_0} X_k \, d\,\mathrm{P} = \mathrm{E}[X \mathbf{1}_{A_0}]$ for all $k \geq n$, implying that $\int_{A_0} X_\infty \, d\,\mathrm{P} = \mathrm{E}[X\mathbf{1}_{A_0}]$. Also, $\int_{A_0} \mathrm{E}[X \mid \mathcal{F}_\infty] \, d\,\mathrm{P} = \mathrm{E}[X\mathbf{1}_{A_0}]$. This proves that $A_0 \in \mathcal{L}$. We thus conclude $\mathcal{P} \subset \mathcal{L}$, any by Dynkin's π-λ Lemma, $\mathcal{F}_\infty \subset \mathcal{L}$. $\square$

Theorem 12 (Kolmogorov's Zero-One Law): Let $(\Omega, \mathcal{F}, \mathrm{P})$ and $\{\mathcal{F}_k\}_{k\geq 1}$ be a sequence of mutually independent σ-algebras, $\mathcal{F}_k \subset \mathcal{F}$ for all k. Define $\mathcal{G}_n = \sigma\left(\bigcup_{k\geq n} \mathcal{F}_k\right)$ and $\mathcal{T} = \bigcap_{n\geq 1} \mathcal{G}_n$. $\mathcal{T}$ is called the $\sigma-$algebra of tail events. Any event in $\mathcal{T}$ has either probability 1 or 0.

Proof. We first prove that any integrable random variable Y that is $\mathcal{T}$-measurable is almost surely constant. By Theorem 11,

$$\mathrm{E}[Y \,|\, \mathcal{F}_k] \longrightarrow \mathrm{E}[Y \,|\, \mathcal{G}_1] = Y$$

almost surely and in L^1. On the other hand, $Y \perp\!\!\!\perp \mathcal{F}_k$ for all $k \in \mathbb{N}$ and therefore

$$\mathrm{E}[Y \,|\, \mathcal{F}_k] = \mathrm{E}[Y].$$

Hence, $Y = \mathrm{E}[Y]$ almost surely. Now for any set $A \in \mathcal{T}$, $\mathbf{1}_A$ is almost surely a constant. And since $\mathbf{1}_A$ can only take value of either 1 or 0, we conclude that $\mathrm{P}(A) = \mathrm{E}[\mathbf{1}_A]$ equals either 1 or 0. □

4.2 Reversed Martingale and Strong Law of Large Numbers

Definition 9 (Reversed Martingale): Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space. Let $\{X_n, \mathcal{F}_n\}_{n\leq -1}$ be an adapted family, i.e., $\mathcal{F}_n \subset \mathcal{F}_m \subset \mathcal{F}_{-1} \subset \mathcal{F}$ for all $-\infty < n < -1$ and X_n is $\mathcal{F}_n$-measurable. $\{X_n, \mathcal{F}_n\}_{n\leq -1}$ is called a *reversed martingale* if

(i) $\mathrm{E}\,|X_n| < \infty$ for all $n \leq -1$

(ii) $\mathrm{E}[X_{n+1} \,|\, \mathcal{F}_n] = X_n$ for all $n \leq -1$.

The definitions for reversed sub-martingale and reversed super-martingale are similar.

Theorem 13 (Convergence of Reversed Martingales): Let $\{X_n, \mathcal{F}_n\}_{n\leq -1}$ be a reversed martingale. Then

$$X_n \longrightarrow \mathrm{E}[X_{-1} \,|\, \mathcal{F}_{-\infty}] \quad \text{as} \quad n \to \infty$$

almost surely and in L^1 where $\mathcal{F}_{-\infty} = \bigcap_{n\leq -1} \mathcal{F}_n$.

Proof. Fix $a < b$. Let U_n denote the number of upcrossings from a to b by $\{X_k\}_{k=n}^{-1}$. By Theorem 7,

$$\mathrm{E}[U_n] \leq \frac{\mathrm{E}(X_{-1} - a)^+}{(b-a)^+}.$$

Let U denote the number of upcrossings from a to b by the whole sequence $\{X_k\}_{k\leq -1}$. Then $U = \lim_{n\to\infty} U_n$, and by MCT,

$$\mathrm{E}[U] = \lim_{n\to\infty} \mathrm{E}[U_n] \leq \frac{\mathrm{E}(X_{-1} - a)^+}{(b-a)^+} < \infty.$$

This proves that $\mathrm{P}(U < \infty) = 1$. Since this holds for all $a < b$, using the same argument in the proof of Theorem 8, we conclude that

$$\mathrm{P}\left(\limsup_{n\to -\infty} X_n = \liminf_{n\to -\infty} X_n\right) = 1.$$

Therefore, $\{X_n\}_{n\le -1}$ converges to some limit $X_{-\infty}$ almost surely. Observe that for all $n < -1$ and $x > 0$,

$$\begin{aligned}
\mathrm{E}\left[|X_n|\mathbf{1}_{\{|X_n|>x\}}\right] &= \mathrm{E}\left[|\mathrm{E}[X_{n+1}\,|\,\mathcal{F}_n]|\,\mathbf{1}_{\{|X_n|>x\}}\right] \\
&\le \mathrm{E}\left[\mathrm{E}[|X_{n+1}|\mathbf{1}_{\{|X_{n+1}|>x\}}\,|\,\mathcal{F}_n]\right] && \text{(by Jensen's Inequality)} \\
&= \mathrm{E}[|X_{n+1}|\mathbf{1}_{\{|X_{n+1}|>x\}}] \\
&\vdots \\
&\le \mathrm{E}[|X_{-1}|\mathbf{1}_{\{|X_{-1}|>x\}}].
\end{aligned}$$

This proves that $\{X_k\}_{k\le -1}$ is uniformly integrable. By Lemma 2, X_k converges to $X_{-\infty}$ in $L^1(\mathcal{F}_{-1})$ and $X_{-\infty}\in L^1(\mathcal{F}_{-\infty})$. We now prove that indeed, $X_{-\infty} = \mathrm{E}[X_{-1}\,|\,\mathcal{F}_{-\infty}]$ almost surely. Let $A\in\mathcal{F}_{-\infty}$. On one hand,

$$\int_A X_k \, d\,\mathrm{P} \longrightarrow \int_A X_\infty \, d\,\mathrm{P} \quad \text{as} \quad k\to -\infty.$$

On the other,

$$\int_A X_k \, d\,\mathrm{P} = \int_A \mathrm{E}[X_k\,|\,\mathcal{F}_{-\infty}] \, d\,\mathrm{P} = \int_A \mathrm{E}[X_{-1}\,|\,\mathcal{F}_{-\infty}] \, d\,\mathrm{P}\,.$$

This proves that

$$\int_A X_\infty \, d\,\mathrm{P} = \int_A \mathrm{E}[X_{-1}\,|\,\mathcal{F}_{-\infty}] \, d\,\mathrm{P}\,.$$

□

Theorem 14 (Strong Law of Large Numbers for i.i.d. rv's): Let $\{X_i\}_{i\ge1}$ be a sequence of i.i.d. random variables with $\mathrm{E}\,|X_1| < \infty$. Then

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k \longrightarrow \mathrm{E}[X_1]$$

almost surely and in L^1.

Proof. Let S_n denote $\sum_{k=1}^n X_k$. Consider the sequence of σ-algebras $\{\mathcal{F}_{-k}\}_{k\ge1}$ and random variables $\{Y_{-k}\}_{k\ge1}$ defined by, for each $k\ge1$,

$$\mathcal{F}_{-k} \coloneqq \sigma(S_k, X_{k+1}, X_{k+2}, \ldots), \quad Y_{-k} \coloneqq \overline{X}_k.$$

We show that $\{Y_{-k},\mathcal{F}_{-k}\}_{k\ge1}$ is a reversed martingale. It is easy to see that $\{Y_{-k},\mathcal{F}_{-k}\}_{k\ge1}$ is an adapted sequence. By independence,

$$\mathrm{E}[Y_{-k+1}\,|\,\mathcal{F}_{-k}] = \mathrm{E}[\overline{X}_{k-1}\,|\,S_k] = \frac{1}{k-1}\sum_{i=1}^{k-1}\mathrm{E}[X_i\,|\,S_k].$$

By symmetry, for all $1\le i\le k$, $\mathrm{E}[X_i\,|\,S_k]$ are the same. Since $\sum_{i=1}^k \mathrm{E}[X_i\,|\,S_k] = S_k$, $\mathrm{E}[X_i\,|\,S_k] = \frac{1}{k}S_k = \overline{X}_k$. This shows that

$$\frac{1}{k-1}\sum_{i=1}^{k-1}\mathrm{E}[X_i\,|\,S_k] = \overline{X}_k = Y_{-k}.$$

By Theorem 13, $\overline{X}_k = Y_{-k}$ converges almost surely and in L^1. This means that $\lim_{k\to\infty} \overline{X}_k$ exists almost surely. Write $\lim_{k\to\infty} \overline{X}_k = \overline{X}_\infty$. Since $\overline{X}_\infty$ is a tail random variable, it is constant almost surely. By the L^1 convergence, $\mathrm{E}[\overline{X}_\infty] = \mathrm{E}[X_1]$, and thus $\overline{X}_\infty = \mathrm{E}[X_1]$ almost surely. □

4.3 Likelihood Ratio Test

Let $\{X_k\}_{k\geq 1}$ be a sequence of i.i.d. random variables defined on a measurable space $(\Omega, \mathcal{F})$. Suppose there are two possible underlying probability measures, P and Q, each induces a p.d.f. of $X_1, X_2, ..., X_n$ on $\mathbb{R}^n$,

$$p(x_1, ..., x_n) = \prod_{i=1}^{n} p(x_i), \quad q(x_1, ..., x_n) = \prod_{i=1}^{n} q(x_i).$$

A statistician is testing a statistical hypothesis with the following null and alternative hypotheses:

$$H_0 : \text{P is the underlying probability measure.}$$
$$H_1 : \text{Q is the underlying probability measure.}$$

He tests the hypothesis using the likelihood ratio test. Write $Z_n = q(X_1, ..., X_n)/p(X_1, ..., X_n) = \prod_{i=1}^{n}(q(X_i)/p(X_i))$. Upon observing $X_1, ..., X_n$, he

$$\begin{cases} \text{rejects } H_0 & \text{if } Z_n \geq s \\ \text{accepts } H_0 & \text{if } Z_n < s. \end{cases}$$

for some predetermined $s > 0$. Let α_n denote the probability of type 1 error (rejecting H_0 when H_0 is true), $\alpha_n \coloneqq \mathrm{P}(Z_n \geq s)$ and β_n denote the probability of type 2 error (accepting H_0 when H_0 is false), $\beta_n \coloneqq \mathrm{Q}(Z_n < s)$.

Theorem 15: The probability of type 1 error and type 2 error both converge to 0 as the sample size goes to infinity. Namely,

$$\alpha_n \longrightarrow 0, \quad \beta_n \longrightarrow 0 \quad \text{as } n \to \infty.$$

Proof. We first prove that the probability of type 1 error goes to 0. On the probability space $(\Omega, \mathcal{F}, \mathrm{P})$, as mentioned in Example 1.2, $\{Z_n, \mathcal{F}_n\}_{n\geq 1}$ is a martingale where $\mathcal{F}_n = \sigma(X_1, ..., X_n)$. Note that Z_n is bounded below by 0, and by Corollary 2, Z_n converges to a finite limit Z_∞ almost surely. Indeed, $Z_\infty = 0$ almost surely. Observe that

$$\mathrm{E}_{\mathrm{P}}[q(X_1)/p(X_1)] = \int_{\mathbb{R}} (q(x)/p(x))p(x)\, dx = 1.$$

Since the two densities p and q are assumed to be different, $q(X_1)/p(X_1)$ is non-constant, and by Jensen's Inequality,

$$\eta \coloneqq \mathrm{E}\left[\left(\frac{q(X_1)}{p(X_1)}\right)^{\frac{1}{2}}\right] < 1.$$

By Fatuo's Lemma,

$$\mathrm{E}\left[Z_\infty^{\frac{1}{2}}\right] \leq \limsup_{n\to\infty} \mathrm{E}\left[Z_n^{\frac{1}{2}}\right] = \limsup_{n\to\infty} \prod_{i=1}^{n} \mathrm{E}\left[\left(\frac{q(X_i)}{p(X_i)}\right)^{\frac{1}{2}}\right] = \limsup_{n\to\infty} \eta^n = 0.$$

This proves that $Z_\infty = 0$ P-almost surely. Therefore, $\alpha_n = \mathrm{P}(Z_n \geq s) \longrightarrow 0$ as $n \to \infty$. Now write the event

$\{Z_n < s\}$ as $\{1/Z_n > \frac{1}{s}\}$ and $Z'_n \coloneqq 1/Z_n = p(X_1, ..., X_n)/q(X_1, ..., X_n)$. Observe that $\{Z'_n, \mathcal{F}_n\}_{n\geq 1}$ forms a martingale on $(\Omega, \mathcal{F}, \mathrm{Q})$. Following the exact same reasoning, we obtain $\beta_n \longrightarrow 0$ as $n \to \infty$. $\square$

References

Grimmet, G., & Stirzaker, D. (2001). *Probability and random processes* (3rd ed.). Oxford University Press.

Athreya, K. B., & Lahiri, S. N. (2006). *Measure theory and probability theory*. Springer.

Gut, A. (2013). *Probability: A graduate course* (2nd ed.). Springer.