An Analytical Introduction to Probability Theory

14. Submartingales

DASN, NTU https://personal.ntu.edu.sg/wptay/

14.1 Submartingales

Definition 14.1. *Consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a filtration $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \ldots \subset \mathscr{F}$. Suppose $X_n \in \mathscr{F}_n$ and $\mathbb{E}|X_n| < \infty$. If for all $m \geq n$, $\mathbb{E}[X_m \mid \mathscr{F}_n] \geq X_n$, then we say that $(X_n, \mathscr{F}_n)_{n\geq 0}$ is a* submartingale. *If $\mathbb{E}[X_m \mid \mathscr{F}_n] \leq X_n$, $(X_n, \mathscr{F}_n)_{n\geq 0}$ is called a* supermartingale.

The same proof as in Theorem 12.1 shows that the martingale transform w.r.t. a predictable sequence (A_n), which is bounded and non-negative, of a submartingale or supermartingale remains as a submartingale or supermartingale, respectively.

Theorem 14.1 (MTT). *If (A_n) is predictable w.r.t. $(\mathscr{F}_n)$, non-negative and bounded, and $(X_n, \mathscr{F}_n)_{n\geq 0}$ is a submartingale (supermartingale), then $(\widetilde{X}_n, \mathscr{F}_n)_{n\geq 0}$ is a submartingale (supermartingale).*

Example 14.1. *In a casino game, let X_n be the amount of money you win at time n if you had bet one dollar at each time, starting with $X_0 = 0$. Recall that a predictable sequence (A_n) is a gambling strategy so that the winnings at time n is*

$$\widetilde{X}_n = \sum_{k=1}^{n} A_k(X_k - X_{k-1}).$$

Suppose that $M_n = X_n - X_{n-1} \in \{-1, 1\}$ has distribution $\mathrm{Bern}\,(p)$. *Your friend claims that a "sure-win" strategy is to choose $A_1 = 1$ and for $n \geq 2$,*

$$A_n = \begin{cases} 2A_{n-1} & \text{if } M_{n-1} = -1, \\ 1 & \text{if } M_{n-1} = 1. \end{cases}$$

Assuming that $p > 0$, $\{M_n = 1 \text{ i.o.}\}$ has probability one. Once $M_n = 1$, the above strategy recoups all your previous losses with a winning of $-1 - 2 - \ldots - 2^{n-1} + 2^n = 1$. Therefore, this strategy seems to suggest you can always beat the house. But if M_n is a supermartingale (i.e., $\mathbb{E}[M_m \mid \mathscr{F}_n] \leq M_n$ for $m \geq n$ and this is usually true has the house incorporates some advantages), then $\widetilde{M}_n$ is also a supermartingale. This means no strategy *can beat the house, in an expectation sense. Many gamblers have gone bankrupt using the above strategy!*

Lemma 14.1. *Suppose $f : \mathbb{R} \mapsto \mathbb{R}$ is convex, $(X_n, \mathscr{F}_n)_{n\geq 0}$ is an adaptation with $\mathbb{E}|f(X_n)| < \infty$ for all $n \geq 0$. Then if either*

i) $(X_n, \mathscr{F}_n)_{n\geq 0}$ is a martingale; or *ii)* $(X_n, \mathscr{F}_n)_{n\geq 0}$ *is a submartingale and* f *is increasing,*

then $(f(X_n), \mathscr{F}_n)_{n\geq 0}$ *is a submartingale.*

Proof. To prove i), from Jensen's inequality, for $n \leq m$, we have $f(X_n) = f(\mathbb{E}[X_m \mid \mathscr{F}_n]) \leq \mathbb{E}[f(X_m) \mid \mathscr{F}_n]$. The proof of ii) is similar. □

Theorem 14.2. *A submartingale* $(X_n, \mathscr{F}_n)_{n\geq 0}$ *can be decomposed a.s. uniquely as* $X_n = Y_n + Z_n$*, where* $(Y_n, \mathscr{F}_n)_{n\geq 0}$ *is a martingale and* $(Z_n)_{n\geq 0}$ *is predictable with* $Z_0 = 0$ *and* $Z_n \leq Z_{n+1}$ *a.s.*

Proof. We first construct a decomposition as follows: Let $Z_0 = 0$,

$$Z_n = \sum_{k=1}^{n} \mathbb{E}[X_k - X_{k-1} \mid \mathscr{F}_{k-1}]$$

and $Y_n = X_n - Z_n$. By our construction, $Z_n \in \mathscr{F}_{n-1}$ is predictable and since

$$\mathbb{E}[X_k - X_{k-1} \mid \mathscr{F}_{k-1}] = \mathbb{E}[X_k \mid \mathscr{F}_{k-1}] - X_{k-1} \geq 0,$$

we have $Z_n \leq Z_{n+1}$ a.s. Furthermore, $Z_n - Z_{n-1} = \mathbb{E}[X_n \mid \mathscr{F}_{n-1}] - X_{n-1}$ and we obtain

$$\mathbb{E}[Y_n \mid \mathscr{F}_{n-1}] = \mathbb{E}[X_n - Z_n \mid \mathscr{F}_{n-1}] = \mathbb{E}[X_n \mid \mathscr{F}_{n-1}] - Z_n = X_{n-1} - Z_{n-1} = Y_{n-1},$$

showing that Y_n is a martingale.

To show uniqueness, we proceed by induction. The requirement that $Z_0 = 0$ implies that $Y_0 = X_0$ uniquely. Suppose the decomposition is unique a.s. up to X_{n-1}. Then for any decomposition $X_n = Y_n + Z_n$ meeting the criteria,

$$Z_n = \mathbb{E}[Z_n \mid \mathscr{F}_{n-1}] = \mathbb{E}[X_n - Y_n \mid \mathscr{F}_{n-1}] = \mathbb{E}[X_n \mid \mathscr{F}_{n-1}] - Y_{n-1},$$

because Y_n is a martingale. This implies that Z_n is unique a.s., and hence so is $Y_n = X_n - Z_n$. The proof is now complete. □

14.2 Doob's Inequalities

Theorem 14.3 (Doob's maximal inequality)**.** *Suppose* $(X_n, \mathscr{F}_n)_{n\geq 0}$ *is a non-negative submartingale. Let* $X_n^* = \max_{0\leq k\leq n} X_k$*. Then for all* $\lambda \geq 0$*,*

$$\lambda \mathbb{P}(X_n^* \geq \lambda) \leq \mathbb{E} X_n \mathbf{1}_{\{X_n^* \geq \lambda\}} \leq \mathbb{E} X_n. \tag{1}$$

Proof. Let $\tau = \inf\{k : X_k \geq \lambda\}$ be a stopping time. Then $\{X_n^* \geq \lambda\} = \{\tau \leq n\}$ and

$$\lambda \mathbf{1}_{\{\tau \leq n\}} \leq X_\tau \mathbf{1}_{\{\tau \leq n\}} = \sum_{k=0}^{n} X_k \mathbf{1}_{\{\tau = k\}}. \tag{2}$$

Since $X_k \le \mathbb{E}[X_n \,|\, \mathscr{F}_k] \;\forall\, k \le n$ and $\{\tau = k\} \in \mathscr{F}_k$, $\mathbb{E}X_k \mathbf{1}_{\{\tau=k\}} \le \mathbb{E}X_n \mathbf{1}_{\{\tau=k\}}$ for all $k \le n$. Taking expectations in (2), we obtain

$$\begin{aligned}\lambda\mathbb{P}(\tau \le n) &\le \mathbb{E}\sum_{k=0}^{n} X_k \mathbf{1}_{\{\tau=k\}} \\ &\le \mathbb{E}\sum_{k=0}^{n} X_n \mathbf{1}_{\{\tau=k\}} \\ &= \mathbb{E}X_n \mathbf{1}_{\{\tau\le n\}} \\ &= \mathbb{E}X_n \mathbf{1}_{\{X_n^*\ge\lambda\}} \\ &\le \mathbb{E}X_n.\end{aligned}$$

□

Corollary 14.1. *For $\lambda > 0$ and $p \ge 1$,*

$$\mathbb{P}(X_n^* \ge \lambda) \le \frac{1}{\lambda^p}\mathbb{E}X_n^p.$$

Proof. From Lemma 14.1, (X_n^p) is a non-negative submartingale. We then apply Theorem 14.3. □

Example 14.2. *Suppose X_i, $i \ge 1$ are independent with $\mathbb{E}X_i = 0$. From Example 12.3, S_n is a martingale. Doob's maximal inequality (or its corollary) recovers Kolmogorov's maximal equality:*

$$\mathbb{P}\Big(\max_{1\le k\le n} |S_k| \ge \lambda\Big) \le \frac{1}{\lambda^2}\mathbb{E}S_n^2.$$

Let $p \ge 1$. A r.v. $X \in L^p$ if $\|X\|_p = \{\mathbb{E}|X|^p\}^{1/p} < \infty$. Hölder's inequalty says that for $1 \le p < \infty$, $p + q = 1$ (i.e., $q = \dfrac{p}{p-1}$), then for any r.v.s X, Y,

$$\|XY\|_1 \le \|X\|_p \|Y\|_q.$$

Lemma 14.2. *Suppose $X, Y \ge 0$, $\mathbb{E}Y^p < \infty$ for $p > 1$ and for all $\lambda \ge 0$, we have*

$$\lambda\mathbb{P}(X \ge \lambda) \le \mathbb{E}Y\mathbf{1}_{\{X\ge\lambda\}}.$$

Then,

$$\|X\|_p \le \frac{p}{p-1}\|Y\|_p. \tag{3}$$

Proof. Let $X_n = X \wedge n$. We use the fact that

$$z^p = p\int_0^z x^{p-1}\,\mathrm{d}x = p\int_0^\infty x^{p-1}\mathbf{1}_{\{z\ge x\}}\,\mathrm{d}x$$

to obtain

$$\begin{aligned}
\mathbb{E}X_n^p &= \mathbb{E}\Big[p\int_0^\infty x^{p-1}\mathbf{1}_{\{X_n\ge x\}}\,\mathrm{d}x\Big]\\
&\overset{\text{Fubini}}{=} p\int_0^\infty x^{p-1}\mathbb{P}(X_n\ge x)\,\mathrm{d}x\\
&\le p\int_0^\infty x^{p-2}\mathbb{E}\Big[Y\mathbf{1}_{\{X\ge x\}}\Big]\,\mathrm{d}x \quad \text{since } \{X_n\ge x\}\subset\{X\ge x\}\\
&\overset{\text{Fubini}}{=} p\mathbb{E}\Big[Y\int_0^\infty x^{p-2}\mathbf{1}_{\{X\ge x\}}\,\mathrm{d}x\Big]\\
&= \frac{p}{p-1}\mathbb{E}\Big[YX^{p-1}\Big]\\
&\overset{\text{Hölder}}{\le} \frac{p}{p-1}\|Y\|_p\|X\|_p^{p-1}.
\end{aligned}$$

Therefore, from Fatou's lemma,

$$\|X\|_p^p \le \liminf_{n\to\infty}\mathbb{E}X_n^p \le \frac{p}{p-1}\|Y\|_p\|X\|_p^{p-1},$$

and we obtain (3) (the case $\|X\|_p = 0$ is trivial). □

Theorem 14.4 (Doob's L^p inequality)**.** *If $(X_n, \mathscr{F}_n)$ is a non-negative submartingale, then for all $p > 1$,*

$$\|X_n^*\|_p \le \frac{p}{p-1}\|X_n\|_p. \tag{4}$$

Proof. We apply Lemma 14.2 with $X = X_n^*$ and $Y = X_n$, together with Doob's maximal inequality (Theorem 14.3). □

Theorems 14.3 and 14.4 require $(X_n, \mathscr{F}_n)$ to be non-negative. For a general submartingale $(X_n, \mathscr{F}_n)$, we can apply these results to $(X_n^+, \mathscr{F}_n)$ since this is a non-negative submartingale from Lemma 14.1.

Theorem 14.4 holds only for $p > 1$. Indeed, there is no corresponding result for $p = 1$ as shown in the following example.

Example 14.3. *Consider the random walk in Example 13.1 with $S_0 = 0$. Take $B = 1$ and $\tau = \inf\{n : S_n = -1\}$. Let $X_n = S_{n\wedge\tau}$. Then using a result in Example 13.1 in the second equality below, we have*

$$\mathbb{E}\Big[\max_{m\ge 0} X_m\Big] = \sum_{A=1}^\infty \mathbb{P}\Big(\max_{m\ge 0} X_m \ge A\Big) = \sum_{A=1}^\infty \frac{1}{A+1} = \infty.$$

The MCT then implies that $\mathbb{E}[\max_{0\le m\le n} X_m] \to \infty$ as $n\to\infty$.

14.3 Uniform Integrability

Definition 14.2. *A collection of r.v.s* $(X_n)_{n\in N}$ *is uniformly integrable if*

$$\sup_{n\in N} \mathbb{E}|X_n|\mathbf{1}_{\{|X_n|>K\}} \to 0, \tag{5}$$

as $K \to \infty$.

If $(X_n)_{n\in N}$ is uniformly integrable, then for K sufficiently large, we have $\sup_{n\in N} \mathbb{E}|X_n|\mathbf{1}_{\{|X_n|>K\}} \le 1$ and $\sup_n \mathbb{E}|X_n| \le K+1 < \infty$ is uniformly bounded. Clearly, the converse is false.

Example 14.4. *If* $|X_n| \le Y$, $\forall\, n \in N$, *and* $\mathbb{E}Y < \infty$, *then*

$$\mathbb{E}|X_n|\mathbf{1}_{\{|X_n|>K\}} \le \mathbb{E}Y\mathbf{1}_{\{Y>K\}} \to 0,$$

as $K \to \infty$ *from DCT. Therefore* $(X_n)_{n\in N}$ *is uniformly integrable.*

Lemma 14.3. *If* $X \in L^1$, *then* $\forall\, \epsilon > 0$, $\exists\, \delta > 0$ *such that if* $\mathbb{P}(A) \le \delta$, *then* $\mathbb{E}|X|\mathbf{1}_A \le \epsilon$.

Proof. If $\mathbb{P}(A) \le \delta$, we have for all $K > 0$,

$$\mathbb{E}|X|\mathbf{1}_A \le K\mathbb{P}(A) + \mathbb{E}|X|\mathbf{1}_{\{|X|>K\}} \le K\delta + \mathbb{E}|X|\mathbf{1}_{\{|X|>K\}}.$$

Choose K sufficiently large so that $\mathbb{E}|X|\mathbf{1}_{\{|X|>K\}} \le \dfrac{\epsilon}{2}$ and set $\delta = \dfrac{\epsilon}{2K}$. Then from above, we have $\mathbb{E}|X|\mathbf{1}_A \le \epsilon$. ◻

Proposition 14.1. *Let* $X \in L^1(\Omega, \mathscr{F}, \mathbb{P})$. *Then* $\{\mathbb{E}[X \mid \mathscr{F}'] : \mathscr{F}' \text{ is a } \sigma\text{-algebra} \subset \mathscr{F}\}$ *is uniformly integrable.*

Proof. Fix $\epsilon > 0$ and choose $\delta > 0$ as in Lemma 14.3. Pick K large so that $\mathbb{E}|X|/K \le \delta$. Let $Y = \mathbb{E}[X \mid \mathscr{F}']$. From Jensen's inequality, $|Y| \le \mathbb{E}[|X| \mid \mathscr{F}']$, therefore we have

$$\begin{aligned}
\mathbb{E}\Big[|Y|\mathbf{1}_{\{|Y|>K\}}\Big] &\le \mathbb{E}\Big[\mathbb{E}[|X| \mid \mathscr{F}']\mathbf{1}_{\{\mathbb{E}[|X||\mathscr{F}']>K\}}\Big] \\
&= \mathbb{E}|X|\mathbf{1}_{\{\mathbb{E}[|X||\mathscr{F}']>K\}} \quad \text{since } \{\mathbb{E}[|X| \mid \mathscr{F}'] > K\} \in \mathscr{F}' \\
&\le \epsilon,
\end{aligned}$$

where the last inequalty follows from Lemma 14.3 as $\mathbb{P}(\mathbb{E}[|X| \mid \mathscr{F}'] > K) \le \mathbb{E}|X|/K \le \delta$. ◻

Proposition 14.2. *Suppose* $\varphi : \mathbb{R} \mapsto \mathbb{R}_+$ *is such that* $\lim_{x\to\infty} \dfrac{\varphi(x)}{x} = \infty$. *(Examples include* $\varphi(x) = x^p$, *for* $p > 1$ *and* $\varphi(x) = x \log^+ x$.*) If* $\mathbb{E}\varphi(|X_n|) \le C < \infty$, *then* $(X_n)_{n\in N}$ *is uniformly integrable.*

Proof. Let $\epsilon_K = \sup\{x/\varphi(x) : x \ge K\}$. Note that $\epsilon_K \to 0$ as $K \to \infty$ because for any $\epsilon > 0$, $\exists\, K$ sufficiently large so that $x/\varphi(x) \le \epsilon$ for all $x > K$. Then we have

$$\mathbb{E}|X_n|\mathbf{1}_{\{|X_n|>K\}} \le \epsilon_K \mathbb{E}\Big[\varphi(|X_n|)\mathbf{1}_{\{|X_n|>K\}}\Big] \le C\epsilon_K \to 0,$$

as $K \to \infty$. ◻

Lemma 14.4. *Suppose $\mathbb{E}|X_n| < \infty$ for all $n \in N$ and $\mathbb{E}|X| < \infty$, then the following are equivalent:*

(i) $X_n \to X$ *in* L^1*, i.e.,* $\mathbb{E}|X_n - X| \to 0$ *as* $n \to \infty$*.*

(ii) $(X_n)_{n\in N}$ *is uniformly integrable and* $X_n \xrightarrow{\text{p}} X$*.*

(iii) $X_n \xrightarrow{\text{p}} X$ *and* $\mathbb{E}|X_n| \to \mathbb{E}|X|$*.*

Proof. We show $(ii) \implies (i) \implies (iii) \implies (ii)$.

$(ii) \implies (i)$: $\forall\, \epsilon > 0$, $K > 0$, we have

$$\begin{aligned}
\mathbb{E}|X_n - X| &\le \epsilon + \mathbb{E}|X_n - X|\mathbf{1}_{\{|X_n - X|>\epsilon\}} \\
&\le \epsilon + \mathbb{E}|X_n|\mathbf{1}_{\{|X_n - X|>\epsilon\}} + \mathbb{E}|X|\mathbf{1}_{\{|X_n - X|>\epsilon\}} \\
&\le \epsilon + 2K\mathbb{P}(|X_n - X| > \epsilon) + \mathbb{E}|X_n|\mathbf{1}_{\{|X_n|>K\}} + \mathbb{E}|X|\mathbf{1}_{\{|X|>K\}}
\end{aligned}$$

and

$$\limsup_{n\to\infty} \mathbb{E}|X_n - X| \le \epsilon + \sup_n \mathbb{E}|X_n|\mathbf{1}_{\{|X_n|>K\}} + \mathbb{E}|X|\mathbf{1}_{\{|X|>K\}}.$$

Taking $\epsilon \to 0$ and $K \to \infty$ completes the proof.

$(i) \implies (iii)$: From Markov's inequality, for any $\epsilon > 0$, we obtain $\mathbb{P}(|X_n - X| > \epsilon) \le \epsilon\mathbb{E}|X_n - X| \to 0$ as $n \to \infty$. We also have

$$|\mathbb{E}|X_n| - \mathbb{E}|X|| \le \mathbb{E}||X_n| - |X|| \le \mathbb{E}|X_n - X| \to 0,$$

as $n \to \infty$.

$(iii) \implies (ii)$: For any $\epsilon > 0$, $\exists\, n_0$ such that $\forall\, n \ge n_0$, $\mathbb{E}|X_n| \le \mathbb{E}|X| + \epsilon/2$. Let

$$\phi_K(x) = \begin{cases} x, & \text{for } x \in [0, K-1], \\ 0, & \text{for } x > K, \\ \text{linear}, & \text{for } x \in [K-1, K]. \end{cases}$$

Then from the DCT, for K sufficiently large,

$$\mathbb{E}|X| - \mathbb{E}\phi_K(|X|) \le \epsilon.$$

Since ϕ_K is continuous, the DCT also yields $\mathbb{E}\phi_K(|X_n|) \to \mathbb{E}\phi_K(|X|)$ as $n \to \infty$. Therefore, since $x \ge \phi_K(x) + x\mathbf{1}_{\{x>K\}}$ for all $x \ge 0$, we have

$$\begin{aligned}
\mathbb{E}|X_n|\mathbf{1}_{\{|X_n|>K\}} &\le \mathbb{E}|X_n| - \mathbb{E}\phi_K(|X_n|) \\
&\le \mathbb{E}|X| - \mathbb{E}\phi_K(|X|) + \epsilon \\
&\le 2\epsilon,
\end{aligned}$$

for all n and K sufficiently large and the proof is complete. $\square$