MARTINGALE INEQUALITIES

In this section we prove the fundamental martingale inequalities.

We first establish Doob's optional sampling theorem which shows that the (super-, sub-)martingale property holds at bounded stopping times.

Theorem 10.1 [Doob's optional stopping theorem] *Let* (X_n) *be a martingale (resp. super-martingale, resp. sub-martingale), and* $S \leq T$ *two* bounded stopping times. *Then* $\mathbb{E}[X_T|\mathcal{F}_S] = X_S$ *(resp.* $\mathbb{E}[X_T|\mathcal{F}_S] \leq X_S$*, resp.* $\mathbb{E}[X_T|\mathcal{F}_S] \geq X_S$ *).*

Proof. [*The proof is not examinable.*] We only need to prove the case for super-martingales. Thus $X = (X_n)$ is a super-martingale. Since S, T are bounded, so there is $N \in \mathbb{N}$ such that $S \leq T \leq N$. Since

$$\{X_S \in G\} \cap \{T = n\} = \{X_n \in G\} \cap \{T = n\} \in \mathcal{F}_n$$

for every n, so X_S is $\mathcal{F}_S$-measurable. Similarly X_T is $\mathcal{F}_T$-measurable. Moreover

$$\mathbb{E}[|X_T|] = \sum_{j=0}^{N} \mathbb{E}\left[|X_j| 1_{\{T=j\}}\right] \leq \sum_{j=0}^{N} \mathbb{E}[|X_j|],$$

so X_T is integrable. Similarly X_S is integrable too.

To show that $\mathbb{E}[X_T|\mathcal{F}_S] \leq X_S$, we only need to prove that

$$\mathbb{E}[X_T : A] \leq \mathbb{E}[X_S : A] \text{ for every } A \in \mathcal{F}_S$$

or equivalently we need to show that for each $A \in \mathcal{F}_S$, we have $\mathbb{E}[X_T - X_S : A] \leq 0$.

Let $A \in \mathcal{F}_S$. Since $S \leq T \leq N$ and $X_T - X_S = 0$ on $\{S = T\}$, we thus have

$$\mathbb{E}[X_T - X_S : A] = \mathbb{E}[X_T - X_S : A \cap \{S < T\}].$$

Now we use the typical technique via stopping times. Write $\{S < T\}$ as disjoint union according to the values. Since $S < T \leq N$, S takes only possible values $0, \cdots, N-1$, so that

$$A \cap \{S < T\} = \bigcup_{j=0}^{N-1} A \cap \{S = j\} \cap \{T > j\}$$

is the disjoint union. Since $A \in \mathcal{F}_S$, so by definition $A \cap \{S = j\} \in \mathcal{F}_j$, and also $\{T > j\} = \{T \leq j\}^c \in \mathcal{F}_j$ for $j = 0, \cdots, N-1$, so that

$$A_j \equiv A \cap \{S = j\} \cap \{T > j\} \in \mathcal{F}_j.$$

Hence, as $X_S = X_j$ on $\{S = j\}$ for $j = 0, \cdots, N-1$, we have

$$\mathbb{E}\left[X_T - X_S : A\right] = \mathbb{E}\left[X_T - X_S : \bigcup_{j=0}^{N-1} A_j\right] = \sum_{j=0}^{N-1} \mathbb{E}\left[X_T - X_j : A_j\right].$$

1) If $0 \le T - S \le 1$, then $X_T = X_{j+1}$ and $X_S = X_j$ on A_j for $j = 0, \cdots, N-1$, and therefore

$$\mathbb{E}\left[X_T - X_S : A\right] = \sum_{j=0}^{N-1} \mathbb{E}\left[X_{j+1} - X_j : A_j\right].$$

However, X is a super-martingale and $A_j \in \mathcal{F}_j$, so that $\mathbb{E}\left[X_{j+1} : A_j\right] \le \mathbb{E}\left[X_j : A_j\right]$. That is, $\mathbb{E}\left[X_{j+1} - X_j : A_j\right] \le 0$ for $j = 0, \cdots, N-1$, and therefore $\mathbb{E}\left[X_T - X_S : A\right] \le 0$ for every $A \in \mathcal{F}_S$.

2) In general, let $R_j = T \wedge (S + j)$, $j = 1, \cdots, n$. Then R_j are stopping times, and $S \le R_1 \le \cdots \le R_n = T$. Moreover $R_1 - S \le 1$ and $R_{j+1} - R_j \le 1$ for $1 \le j \le N-1$. Let $A \in \mathcal{F}_S$. Then $A \in \mathcal{F}_{R_j}$ as $S \le R_j$. Therefore by applying the first case to R_j we obtain

$$\mathbb{E}\left[X_S : A\right] \ge \mathbb{E}\left[X_{R_1} : A\right] \ge \cdots \ge \mathbb{E}\left[X_T : A\right]$$

so that $\mathbb{E}\left[1_A X_S\right] \ge \mathbb{E}\left[1_A X_T\right]$. The proof is complete. ∎

Let us first deduce several easy but important consequences from Doob's optional stopping theorem.

Corollary 10.2 *Let* $X = (X_n)$ *be a* super-martingale.

1) If $T \ge S$ *are two* bounded stopping times, *then* $\mathbb{E}\left[X_T\right] \le \mathbb{E}\left[X_S\right]$.

2) If T *is a stopping time, then* $\mathbb{E}\left[X_{T\wedge n}\right] \le \mathbb{E}\left[X_{T\wedge m}\right]$ *for any* $n \ge m$, *where* $X_{T\wedge n} = X_T$ *on* $\{T \le n\}$ *and* $X_{T\wedge n} = X_n$ *on* $\{T > n\}$.

Similar results hold for sub-martingales.

Proof. For 1) we have $\mathbb{E}\left[X_T|\mathcal{F}_S\right] \le \mathbb{E}\left[X_S\right]$, then taking expectations both sides we obtain $\mathbb{E}\left(X_T\right) \le \mathbb{E}\left(X_S\right)$. 2) follows 1) as $T \wedge m$ and $T \wedge n$ are both stopping times, and bounded by n. ∎

Corollary 10.3 *If* $X = (X_n)$ *is a super-martingale, and* T *is a stopping time, then*

$$\mathbb{E}\left[|X_{T\wedge n}|\right] \le \mathbb{E}\left[X_0\right] + 2\mathbb{E}\left[X_n^-\right] \quad \forall n \in \mathbb{Z}_+.$$

If in addition $\sup_n \mathbb{E}\left[|X_n|\right] < \infty$, *then*

$$\mathbb{E}\left[|X_T| 1_{\{T<\infty\}}\right] \le 3 \sup_n \mathbb{E}\left[|X_n|\right].$$

Proof. Since (X_n) is a super-martingale, so its negative part (X_n^-) is a sub-martingale, hence $\mathbb{E}[X_{T\wedge n}] \leq \mathbb{E}[X_0]$ and $\mathbb{E}\left[X_{T\wedge n}^-\right] \leq \mathbb{E}[X_n^-]$. Since

$$|X_{T\wedge n}| = X_{T\wedge n}^+ + X_{T\wedge n}^- = X_{T\wedge n} + 2X_{T\wedge n}^-$$

we therefore have

$$\begin{aligned} \mathbb{E}[|X_{T\wedge n}|] &= \mathbb{E}[X_{T\wedge n}] + 2\mathbb{E}\left[X_{T\wedge n}^-\right] \\ &\leq \mathbb{E}[X_0] + 2\mathbb{E}\left[X_n^-\right] \end{aligned}$$

which is the first inequality. It follows that

$$\mathbb{E}\left[|X_{T\wedge n}|1_{\{T<\infty\}}\right] \leq 3\sup_n \mathbb{E}[|X_n|] \qquad (10.1)$$

for every n. Since

$$|X_T|1_{\{T<\infty\}} = \lim_{n\to\infty} |X_{T\wedge n}|1_{\{T<\infty\}}$$

and applying Fatou's lemma to $|X_{T\wedge n}|1_{\{T<\infty\}}$, we obtain

$$\mathbb{E}\left[|X_T|1_{\{T<\infty\}}\right] = \mathbb{E}\left[\lim_{n\to\infty} |X_T|1_{\{T<\infty\}}\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[|X_{T\wedge n}|1_{\{T<\infty\}}\right] \leq 3\sup_n \mathbb{E}[|X_n|]$$

where the last inequality follows from (10.1). ∎

Theorem 10.4 (Stopped super-martingales are super-martingales) *Suppose $X = (X_n)$ is a super-martingale, and suppose T is a stopping time, then the stopped process $X^T = (X_{T\wedge n})$ is again a super-martingale. A similar result holds for martingales and sub-martingales.*

Proof. Since $|X_{T\wedge n}| \le \sum_{j=0}^{n} |X_j|$, so $X_{T\wedge n}$ is integrable for every $n \in \mathbb{Z}$. For $n \ge m$ we have

$$\begin{aligned}\mathbb{E}\left[X_{T\wedge n}|\mathcal{F}_m\right] &= \sum_{k=0}^{m} \mathbb{E}\left[X_k 1_{\{T=k\}}|\mathcal{F}_m\right] + \mathbb{E}\left[X_{T\wedge n} 1_{\{T>m\}}|\mathcal{F}_m\right] \\ &= \sum_{k=0}^{m} X_k 1_{\{T=k\}} + \mathbb{E}\left[X_{T\wedge n} 1_{\{T>m\}}|\mathcal{F}_m\right] \\ &= \sum_{k=0}^{m} X_k 1_{\{T=k\}} + 1_{\{T>m\}}\mathbb{E}\left[X_{T\wedge n} 1_{\{T>m\}}|\mathcal{F}_m\right]. \end{aligned} \tag{10.2}$$

where we have used the fact that $\{T > m\} \in \mathcal{F}_m$. Let $S = T 1_{\{T>m\}} + \infty 1_{\{T\le m\}}$. Then S is a stopping time. In fact, if $k \le m$, then $\{S = k\} = \emptyset$, and if $k > m$, then

$$\{S = k\} = \{T = k\} \cap \{T > m\} \in \mathcal{F}_k$$

as $\{T > m\} \in \mathcal{F}_m \subseteq \mathcal{F}_k$. By definition $S \wedge n \ge m$, and

$$X_{S\wedge n} = X_{T\wedge n} 1_{\{T>m\}} + X_n 1_{\{T\le m\}}.$$

Hence, by applying Doob's stoping theorem to X and bounded stopping times $S \wedge n \ge m$ we obtain

$$\mathbb{E}\left[X_{S\wedge n}|\mathcal{F}_m\right] \le X_m$$

that is

$$\mathbb{E}\left[X_{T\wedge n} 1_{\{T>m\}} + X_n 1_{\{T\le m\}}|\mathcal{F}_m\right] \le X_m.$$

Since $\{T \le m\} \in \mathcal{F}_m$, it follows that

$$\mathbb{E}\left[X_{T\wedge n} 1_{\{T>m\}}|\mathcal{F}_m\right] + 1_{\{T\le m\}}\mathbb{E}\left[X_n|\mathcal{F}_m\right] \le X_m.$$

Thus, by multiplying both sides by $1_{\{T>m\}}$, we have

$$1_{\{T>m\}}\mathbb{E}\left[X_{T\wedge n} 1_{\{T>m\}}|\mathcal{F}_m\right] \le X_m 1_{\{T>m\}}. \tag{10.3}$$

Putting together (10.2) with (10.3) we obtain that

$$\mathbb{E}\left[X_{T\wedge n}|\mathcal{F}_m\right] \le \sum_{k=0}^{m} X_k 1_{\{T=k\}} + X_m 1_{\{T>m\}} = X_{T\wedge m}$$

which means that X^T is again a super-martingale. ■

Corollary 10.5 *Let T be a* finite *stopping time.*

1) If $X = (X_n)$ is a non-negative *super-martingale, then $\mathbb{E}[X_T] \le \mathbb{E}[X_0]$.*

2) If $X = (X_n)$ is a super-martingale, and there is an integrable random variable ξ such that $|X_n| \le \xi$ almost everywhere on Ω for all n, then $\mathbb{E}[X_T] \le \mathbb{E}[X_0]$.

Proof. 1) In fact, since T is finite, $X_{T\wedge n} \to X_T$ as $n \to \infty$. By Fatou's lemma we have

$$\mathbb{E}[X_T] \leq \liminf_{n\to\infty} \mathbb{E}[X_{T\wedge n}] \leq \mathbb{E}[X_0]$$

which completes the proof.

2) This time we apply the Dominated Convergence Theorem to $\{X_{T\wedge n}\}$ to obtain $\mathbb{E}[X_T] = \lim_{n\to\infty} \mathbb{E}[X_{T\wedge n}]$. ∎

Corollary 10.6 *Let T be a* finite *stopping time, and $X = (X_n)$ be a super-martingale. Let $\xi = \sup_{n=1,2,\cdots} |X_n - X_{n-1}|$. Suppose ξT is integrable, i.e. $\mathbb{E}[\xi T] < \infty$, then $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$. In particular, if the sequence $|X_n - X_{n-1}| \leq L$ for every n, where L is a constant, and if T is an* integrable *stopping time, then $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$.*

Proof. For every n, we have

$$|X_{T\wedge n}| = \left| X_0 + \sum_{k=1}^{n\wedge T} (X_k - X_{k-1}) \right| \leq |X_0| + \xi T.$$

Since $|X_0| + \xi T$ is integrable, and $X_{T\wedge n} \to X_0$ almost everywhere, by Lebesgue's Dominated Convergence Theorem, $\mathbb{E}[X_T] = \lim_{n\to\infty} \mathbb{E}[X_{T\wedge n}]$ which yields the conclusion. ∎

In order to establish a general result such as 2) in Corollary 10.5, the concept of uniform integrability may be useful. For example, we have the following

Corollary 10.7 *Let T be a* finite *stopping time, and $X = (X_n)$ be a super-martingale. Suppose $\{X_{T\wedge n} : n = 0, 1, 2, \cdots\}$ is uniformly integrable, then $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$.*

The proof is exactly the same as that of 2), Corollary 10.5. In fact, since $X_{T\wedge n} \to X_T$ and $\{X_{T\wedge n}\}$ is uniformly integrable, by Theorem 8.4, $\mathbb{E}[X_T] = \lim_{n\to\infty} \mathbb{E}[X_{T\wedge n}]$.

It is therefore useful to introduce the following definition.

Definition 10.8 *Let $X = (X_n)_{n\in\mathbb{Z}_+}$ be an adapted sequence of real random variables on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_n, \mathbb{P})$. Let $\mathcal{T}$ denote the collection of all* finite *$(\mathcal{F}_n)$*-stopping times. *Then we say $X = (X_n)$ is of* class D, *if the family $\{X_T : T \in \mathcal{T}\}$ is uniformly integrable.*

Next we derive the main martingale inequalities, as applications of Doob's optional stopping theorem. Let us introduce a notation first.

If $(X_n)_{n\in\mathbb{Z}_+}$ is a sequence of real random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, for each $n \in \mathbb{Z}_+$, set $X_n^*(\omega) = \max_{k\leq n} X_k(\omega)$ for $\omega \in \Omega$. Then (X_n^*) is called the sequence of running maximal of (X_n). It is obvious that each X_n^* is a random variable. If $(X_n)_{n\in\mathbb{Z}_+}$ is an adapted sequence on the filtered space $(\Omega, \mathcal{F}, \mathcal{F}_n, \mathbb{P})$, then so is its running maximal.

Theorem 10.9 *[*Doob's maximal inequality *for sub-martingales] If $Y = (Y_n)$ is a* sub-martingale, *then*

$$\mathbb{P}\left[\sup_{k\leq n} Y_k \geq \lambda\right] \leq \frac{1}{\lambda}\mathbb{E}\left[Y_n : \sup_{k\leq n} Y_k \geq \lambda\right] \tag{10.4}$$

for any $\lambda > 0$ and for every $n = 0, 1, 2, \cdots$.

Proof. [*The proof is not examinable.*] Let $T = \inf\{k : Y_k \geq \lambda\}$. Then T is a stopping time, and

$$\{T = j\} = \{Y_0 < \lambda, \cdots, Y_{j-1} < \lambda, Y_j \geq \lambda\}. \tag{10.5}$$

Therefore

$$\mathbb{P}\left[\sup_{k\leq n} Y_k \geq \lambda\right] = \mathbb{P}[T \leq n] = \sum_{j=0}^{n} \mathbb{P}[T = j].$$

By (10.5)

$$\mathbb{P}[T = j] = \mathbb{P}[Y_j \geq \lambda, T = j] \leq \mathbb{E}\left[\frac{Y_j}{\lambda} : T = j\right] = \frac{1}{\lambda}\mathbb{E}[Y_j : T = j],$$

since Y is a sub-martingale and $\{T = j\} \in \mathcal{F}_j$, so that for $j \leq n$ we have

$$\mathbb{P}[T = j] \leq \frac{1}{\lambda}\mathbb{E}[Y_j : T = j] \leq \frac{1}{\lambda}\mathbb{E}[Y_n : T = j]$$

and therefore

$$\begin{aligned}\mathbb{P}\left[\sup_{k\leq n} Y_k \geq \lambda\right] &= \sum_{j=0}^{n} \mathbb{P}[T = j] \leq \frac{1}{\lambda}\sum_{j=0}^{n} \mathbb{E}[Y_n : T = j] \\ &= \frac{1}{\lambda}\mathbb{E}[Y_n : T \leq n] = \frac{1}{\lambda}\mathbb{E}\left[Y_n : \sup_{k\leq n} Y_k \geq \lambda\right]\end{aligned}$$

which completes the proof. ■

As a consequence, we have the following important martingale inequality.

Corollary 10.10 *[*Doob's maximal inequality *for martingales] If $M = (M_n)$ is a martingale, then*

$$\mathbb{P}\left[\sup_{k\leq n} |M_k| \geq \lambda\right] \leq \frac{1}{\lambda}\mathbb{E}\left[|M_n| : \sup_{k\leq n} |M_k| \geq \lambda\right] \tag{10.6}$$

for any $\lambda > 0$ and $n = 0, 1, \cdots$.

Proof. Since M is a martingale, so $(|M_n|)$ is a (non-negative) sub-martingale, and (10.6) follows from (10.5) immediately. ■

There is a slightly different version of the maximal inequality for super-martingales.

Theorem 10.11 *[*Doob's maximal inequality *for super-martingales] If $X = (X_n)$ is a* super-martingale, *then*

$$\mathbb{P}\left[\sup_{k\leq n} X_k \geq \lambda\right] \leq \frac{1}{\lambda}\left(\mathbb{E}[X_0] - \mathbb{E}\left[X_n : \sup_{k\leq n} X_k \leq \lambda\right]\right) \tag{10.7}$$

for any $\lambda > 0$, $n \in \mathbb{Z}_+$, and

$$\mathbb{P}\left[\sup_{k\leq n} |X_k| \geq \lambda\right] \leq \frac{1}{\lambda}\left(\mathbb{E}[X_0] + 2\mathbb{E}\left[X_n^-\right]\right) \tag{10.8}$$

for all $\lambda > 0$, where $X_n^- = \max\{0, -X_n\}$ which is a sub-martingale.

Proof. [*The proof is not examinable.*] Let $R = \inf\{k \geq 0 : X_k \geq \lambda\}$ and $T = R \wedge n$. Then T is a bounded stopping time. Since X is a super-martingale, so that, applying Doob's optional theorem to stopping times T and $S = 0$, one has $\mathbb{E}[X_0] \geq \mathbb{E}[X_T]$, hence

$$\begin{aligned}\mathbb{E}[X_0] &\geq \mathbb{E}\left[X_T : \sup_{k\leq n} X_k \geq \lambda\right] + \mathbb{E}\left[X_T : \sup_{k\leq n} X_k < \lambda\right] \\ &\geq \lambda\mathbb{P}\left[\sup_{k\leq n} X_k \geq \lambda\right] + \mathbb{E}\left[X_n : \sup_{k\leq n} X_k < \lambda\right]\end{aligned}$$

here for the second inequality we have used the fact that on $\left\{\sup_{k\leq n} X_k \geq \lambda\right\}$, $R \leq n$, so that $X_T = X_R \geq \lambda$, which is equivalent to (10.7).

Now we prove the second estimate. Since $X = (X_n)$ is a super-martingale, $(-X_n)$ is a sub-martingale, so that

$$\begin{aligned}\mathbb{P}\left[\inf_{k\leq n} X_k \leq -\lambda\right] &= \mathbb{P}\left[\sup_{k\leq n}(-X_k) \geq \lambda\right] \\ &\leq \frac{1}{\lambda}\mathbb{E}\left[-X_n : \inf_{k\leq n} X_k \leq -\lambda\right].\end{aligned}$$

together with (10.7) we deduce that

$$\begin{aligned}\mathbb{P}\left[\sup_{k\leq n}|X_k| \geq \lambda\right] &= \mathbb{P}\left[\sup_{k\leq n} X_k \geq \lambda, \text{ or } \inf_{k\leq n} X_k \leq -\lambda\right] \\ &\leq \mathbb{P}\left[\sup_{k\leq n} X_k \geq \lambda\right] + \mathbb{P}\left[\inf_{k\leq n} X_k \leq -\lambda\right] \\ &\leq \frac{1}{\lambda}\mathbb{E}[X_0] - \frac{1}{\lambda}\mathbb{E}[X_n : X_n^* \leq \lambda] + \frac{1}{\lambda}\mathbb{E}\left[-X_n : \inf_{k\leq n} X_k \leq -\lambda\right] \\ &\leq \frac{1}{\lambda}\left(\mathbb{E}[X_0] + 2\mathbb{E}\left[X_n^-\right]\right)\end{aligned}$$

which is the last inequality. ■

The following result plays a key role in proving the strong law of large numbers, which is a strong version of the elementary Markov inequality.

Theorem 10.12 *[*Kolmogorov's inequality*] Let (X_n) be a martingale and $X_N \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ where N is a positive integer. Then for any $\lambda > 0$*

$$\mathbb{P}\left[\sup_{k\leq N}|X_k| \geq \lambda\right] \leq \frac{1}{\lambda^2}\mathbb{E}\left[X_N^2\right]. \tag{10.9}$$

Proof. By Jensen's inequality, for any $k \leq N$

$$\mathbb{E}\left[X_k^2\right] = \mathbb{E}\left(\mathbb{E}[X_N|\mathcal{F}_k]\right)^2 \leq \mathbb{E}\left[X_N^2\right] < \infty.$$

[That is (X_n) is a square integrable martingale up to N]. Therefore (X_k^2) $(k = 0, 1, \cdots, N)$ is a sub-martingale (up to time N). Applying Doob's maximal inequality one obtains

$$\mathbb{P}\left[\sup_{k\leq n} X_k^2 \geq \lambda^2\right] \leq \frac{1}{\lambda^2}\mathbb{E}\left[X_n^2 : \sup_{k\leq n} X_k^2 \geq \lambda^2\right] \leq \frac{1}{\lambda^2}\mathbb{E}\left[X_n^2\right]$$

for all $n \leq N$. ■

Example 10.13 *Let (X_n) be independent and square integrable. Then $S_n = \sum_{k=0}^{n}(X_k - \mu_k)$ where $\mu_k = \mathbb{E}[X_k]$ is a martingale. Moreover*

$$\mathbb{E}\left[S_n^2\right] = \mathbb{E}\left[\sum_{k=0}^{n}(X_k - \mu_k)\right]^2 = \sum_{k=0}^{n}\sigma_k^2$$

where $\sigma_k^2 = var(X_k)$. According to Kolmogorov's inequality

$$\mathbb{P}\left[\sup_{k\le n}\left|\sum_{l=0}^{k}(X_l - \mu_l)\right| \ge \lambda\right] \le \frac{1}{\lambda^2}\sum_{k=0}^{n}\sigma_k^2$$

for any $\lambda > 0$.

Doob's maximal inequality is a tail estimate for the distribution of the running maximum of a martingale, thus can be used to estimate the L^p-norm, which is the context of Doob's L^p-inequality.

Let us begin with an elementary lemma which follows from Fubini's theorem directly.

Lemma 10.14 *Suppose ρ is right-continuous, increasing on $(0,\infty)$ and $\rho(0+) = 0$, and ξ is a non-negative random variable on $(\Omega, \mathcal{F}, \mathbb{P})$, then*

$$\rho(\xi) = \rho(\xi) - \rho(0+) = \int_{(0,\xi]} m_\rho(d\lambda) \quad \text{on } \{\xi > 0\}$$

$$\begin{aligned}
\mathbb{E}[\rho(\xi) : \xi > 0] &= \mathbb{E}\left[\int_{(0,\xi]} m_\rho(d\lambda) : \xi > 0\right] = \mathbb{E}\left[\int_{(0,\infty)} 1_{\{\lambda\le\xi\}} m_\rho(d\lambda)\right] \\
&= \int_{\Omega\times(0,\infty)} 1_{\{\xi\ge\lambda\}} m_\rho(d\lambda) d\mathbb{P} = \int_{(0,\infty)} \mathbb{P}[\xi \ge \lambda]\, m_\rho(d\lambda),
\end{aligned}$$

where $m_\rho(d\lambda)$ is the Lebesgue-Stieltjes measure defined by ρ on $(0,\infty)$, so that $m_\rho((s,t]) = \rho(t) - \rho(s)$ for any $t \ge s \ge 0$.

Theorem 10.15 *[*Doob's L^p-inequality*] 1) If (X_n) is a* non-negative sub-martingale*, then, for any $p > 1$*

$$\mathbb{E}\left[(X_n^*)^p\right] \le \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[(X_n)^p\right]. \tag{10.10}$$

2) Suppose (X_n) is a martingale and $p > 1$,

$$\mathbb{E}\left[\max_{k\le n}|X_k|^p\right] \le \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|X_n|^p\right]. \tag{10.11}$$

In particular, for $p > 1$,

$$\|X_n^*\|_p \le q\,\|X_n\|_p$$

where $\frac{1}{p} + \frac{1}{q} = 1$, $\|\cdot\|_p$ denotes the L^p-norm, and $X_n^* = \max_{k\le n} X_k$ is the running maximum.

Proof. If (X_n) is a martingale, then $(|X_n|)$ is a sub-martingale, so (10.11) follows from (10.10).

If $X = (X_n)_{n\geq 0}$ is a sub-martingale, and X_n are non-negative, then, by Doob's maximal inequality

$$\mathbb{P}[X_n^* \geq \lambda] \leq \frac{1}{\lambda}\mathbb{E}[X_n; X_n^* \geq \lambda]$$

for any $\lambda > 0$. If ρ is right continuous, increasing and $\rho(0+) = 0$ on $(0, \infty)$, by Lemma 10.14

$$\begin{aligned}
\mathbb{E}[\rho(X_n^*) : X_n^* > 0] &= \mathbb{E}\left[\int_{(0,X_n^*]} m_\rho(d\lambda) : X_n^* > 0\right] \\
&= \int_{(0,\infty)} \mathbb{P}[X_n^* \geq \lambda]\, m_\rho(d\lambda) \\
&\leq \int_{(0,\infty)} \frac{1}{\lambda}\mathbb{E}[X_n : X_n^* \geq \lambda]\, m_\rho(d\lambda) \\
&= \int_{(0,\infty)} \left\{\frac{1}{\lambda}\int_{\{X_n^*\geq\lambda\}} X_n d\mathbb{P}\right\} m_\rho(d\lambda) \\
&= \mathbb{E}\left[X_n\left(\int_{(0,X_n^*]} \frac{1}{\lambda} m_\rho(d\lambda)\right) : X_n^* > 0\right].
\end{aligned}$$

where the inequality above follows from the maximal inequality.

Let $p > 1$, and $\rho(\lambda) = \lambda^p$. Then $\rho'(\lambda) = p\lambda^{p-1}$ and $m_\rho(d\lambda) = p\lambda^{p-1}1_{(0,\infty)}d\lambda$. Applying the previous estimate to $\rho(\lambda) = \lambda^p$, we obtain that

$$\begin{aligned}
\mathbb{E}[(X_n^*)^p] &= \mathbb{E}[(X_n^*)^p : X_n^* > 0] \\
&\leq \mathbb{E}\left[X_n\left(\int_0^{X_n^*} \frac{1}{\lambda} p\lambda^{p-1} d\lambda\right) : X_n^* > 0\right] \\
&= \frac{p}{p-1}\mathbb{E}\left[X_n (X_n^*)^{p-1}\right].
\end{aligned}$$

For the term on the right-hand side, we apply the the Holder inequality

$$\int_\Omega |fg| d\mu \leq \|f\|_p \|g\|_q$$

to $f = X_n$ and $g = (X_n^*)^{p-1}$, $\frac{1}{p} + \frac{1}{q} = 1$, so that

$$\mathbb{E}[(X_n^*)^p] \leq \frac{p}{p-1}\left(\mathbb{E}[(X_n)^p]\right)^{\frac{1}{p}} \left(\mathbb{E}\left[(X_n^*)^{(p-1)q}\right]\right)^{\frac{1}{q}}.$$

Since $\frac{1}{q} = \frac{p-1}{p}$, $(p-1)q = p$, so after simplification,

$$\mathbb{E}[(X_n^*)^p] \leq \frac{p-1}{p}\left(\mathbb{E}[(X_n)^p]\right)^{\frac{1}{p}} \left(\mathbb{E}[(X_n^*)^p]\right)^{1-\frac{1}{p}}.$$

If $0 < \mathbb{E}[(X_n^*)^p] < \infty$, then by dividing both sides of the previous inequality by $\left(\mathbb{E}[(X_n^*)^p]\right)^{1-\frac{1}{p}}$ and taking p-th power both sides, to obtain

$$\mathbb{E}[(X_n^*)^p] \leq \left(\frac{p-1}{p}\right)^p \mathbb{E}[(X_n)^p]. \tag{10.12}$$

If $\mathbb{E}[(X_n^*)^p] = 0$, then $X_n^* = 0$ almost surely, so that $X_n = 0$ a.e. too, then above inequality trivially true.

Finally we prove the general case where $\mathbb{E}[(X_n^*)^p]$ my be infinity. Note that, since (X_n) is a non-negative sub-martingale, and for every constant $K > 0$, $\varphi(t) = t \wedge K$ for $t \geq 0$ and $\varphi(t) = 0$ for $t \leq 0$ is increasing and convex, so that $X_n \wedge K = \varphi(X_n)$ $(n = 0, 1, \cdots)$ is a non-negative and *bounded* sub-martingale, so that we can apply (10.12) to the non-negative sub-martingale $(X_n \wedge K)$ for each $K > 0$, to obtain that

$$\mathbb{E}[(X_n^* \wedge K)^p] \leq \left(\frac{p-1}{p}\right)^{\frac{1}{p}} \mathbb{E}[(X_n \wedge K)^p] \tag{10.13}$$

for every $K = 1, 2, \cdots$. Clearly $0 \leq X_n^* \wedge K \uparrow X_n^*$ and $0 \leq X_n \wedge K \uparrow X_n$ for every n, as $K \uparrow \infty$, so that by sending $K \to \infty$ in (10.13) and applying MCT to both sides, we finally get that

$$\mathbb{E}[(X_n^*)^p] \leq \left(\frac{p-1}{p}\right)^{\frac{1}{p}} \mathbb{E}[(X_n)^p] \tag{10.14}$$

which completes the proof. ■

The Doob's inequality implies that, for a martingale $X = (X_n)$, the L^p-norm of X_n^* and the L^p-norm of X_n are equivalent as long as $p > 1$, and

$$\|X_n\|_{L^p} \leq \|X_n^*\|_{L^p} \leq \frac{p}{p-1} \|X_n\|_{L^p}.$$

In particular, for a martingale (X_n), X_n^* is p-th integrable if and only if X_n is p-th integrable for every $p > 1$.

Doob's L^p-inequality does not apply to the case $p = 1$, as in this case $q = \infty$ which gives the infinity upper bound. That is to say, the L^1-norm of the terminal value of a martingale does not in general control the L^1-norm of its running maximal.

Exercise 10.16 *Prove that $\log x \leq x/e$ for all $x > 0$, hence prove that*

$$a \log^+ b \leq a \log^+ a + \frac{b}{e}. \tag{10.15}$$

Consider $h(t) = \log t - \frac{t}{e}$ for $t > 0$. Then $h(t) \to -\infty$ as $t \downarrow 0$ or $t \uparrow \infty$, so h achieves its maximum in $(0, \infty)$. Since $h'(t) = \frac{1}{t} - \frac{1}{e}$ has unique root $t = e$, e is the maximum of h. Therefore $h(t) \leq h(e) = 0$ for all $t > 0$, that is, $\log t \leq \frac{t}{e}$.

Now

$$\begin{aligned}\log^+(at) &= \max\{0, \log(at)\} = \max\{0, \log a + \log t\} \\ &\leq \max\left\{0, \log^+ a + \frac{t}{e}\right\} = \log^+ a + \frac{t}{e},\end{aligned}$$

Setting $t = \frac{b}{a}$ we obtain the inequality (10.15).

Theorem 10.17 *If (X_n) is a non-negative sub-martingale, then*

$$\mathbb{E}\left[\max_{k \leq n} X_k\right] \leq \frac{e}{e-1}\left(1 + \mathbb{E}\left[X_n \log^+ X_n\right]\right). \tag{10.16}$$

Proof. [*The proof is not examinable.*] We have seen from the proof of Doob's L^p-inequality

$$\mathbb{E}\left[\rho(X_n^*) : X_n^* > 0\right] \leq \mathbb{E}\left[X_n \int_{(0,X_n^*]} \frac{1}{\lambda} m_\rho(d\lambda) : X_n^* > 0\right].$$

where now $\rho(\lambda) = (\lambda - 1)^+$ which is a continuous increasing function with support on $[1, \infty)$. Therefore

$$\begin{aligned}
\mathbb{E}\left[\rho(X_n^*)\right] = \mathbb{E}\left[\rho(X_n^*) : X_n^* > 0\right] &\leq \mathbb{E}\left[X_n \int_1^{X_n^*} \frac{1}{\lambda} d\lambda : X_n^* \geq 1\right] \\
&= \mathbb{E}\left[X_n \log^+ X_n^*\right] \\
&\leq \mathbb{E}\left[X_n \log^+ X_n\right] + \frac{1}{e}\mathbb{E}\left[X_n^*\right].
\end{aligned}$$

where we have used the inequality

$$X_n \log X_n^* \leq X_n \log^+ X_n + \frac{X_n^*}{e}.$$

On the other hand

$$\begin{aligned}
\mathbb{E}\left[X_n^*\right] &= \mathbb{E}\left[X_n^* 1_{\{X_n^* \geq 1\}}\right] + \mathbb{E}\left[X_n^* 1_{\{X_n^* < 1\}}\right] \\
&\leq \mathbb{E}\left[\rho(X_n^*) 1_{\{X_n^* \geq 1\}}\right] + \mathbb{E}\left[1_{\{X_n^* > 1\}}\right] + \mathbb{E}\left[X_n^* 1_{\{X_n^* < 1\}}\right] \\
&\leq \mathbb{E}\left[\rho(X_n^*)\right] + 1.
\end{aligned}$$

Together with the previous estimate one thus deduces that

$$\mathbb{E}\left[X_n^*\right] \leq 1 + \mathbb{E}\left[X_n \log^+ X_n\right] + \frac{\mathbb{E}\left[X_n^*\right]}{e}$$

which yields the L^1-estimate. ∎