Doob's Inequality and Convergence in L^p, $p > 1$

김지수 (Jisu KIM)

확률론 2 (Probability Theory 2), 2025 2nd semester (fall)

Some parts of this lecture note are from the lecture notes from Prof. Alessandro Rianldo's "Advanced Probability Overview".

We first recall [1, Theorem 4.2.8] and [1, Theorem 4.2.9]:

Theorem ([1, Theorem 4.2.8])**.** *Let X_n be a (sub, super) martingale with respect to $\mathcal{F}_n$ and let $\{H_n\}$ be predictable. Suppose $(H \bullet X)_n$ is integrable (this holds in particular when each H_n is bounded).*

Then $\{(H \bullet X)_n,\ \mathcal{F}_n\}$ is a martingale.

If $H_n \geq 0$, it is a (sub, super) martingale.

Theorem ([1, Theorem 4.2.9])**.** *If N is a stopping time (정지시간) and X_n is a supermartingale, then $X_{N \wedge n}$ is a supermartingale.*

Now we prove a consequence of [1, Theorem 4.2.9].

Theorem ([1, Theorem 4.4.1])**.** *If X_n is a submartingale and N is a stopping time with $P(N \leq k) = 1$, then*

$$\mathbb{E}X_0 \leq \mathbb{E}X_N \leq \mathbb{E}X_k.$$

Remark. Let S_n be a simple random walk with $S_0 = 1$ and let $N = \inf\{n : S_n = 0\}$ (see [1, Example 4.2.13] for more details). $\mathbb{E}S_0 = 1 > 0 = \mathbb{E}S_N$, so the first inequality need not hold for unbounded stopping times. In Section 4.8 we will give conditions that guarantee $\mathbb{E}X_0 \leq \mathbb{E}X_N$ for unbounded N.

Proof. [1, Theorem 4.2.9] implies $X_{N \wedge n}$ is a submartingale, so it follows that

$$\mathbb{E}X_0 = \mathbb{E}X_{N \wedge 0} \leq \mathbb{E}X_{N \wedge k} = \mathbb{E}X_N.$$

To prove the other inequality, let $K_n = 1_{\{N<n\}} = 1_{\{N \leq n-1\}}$. K_n is predictable, so [1, Theorem 4.2.8] implies $(K \cdot X)_n = X_n - X_{N \wedge n}$ is a submartingale, and it follows that

$$\mathbb{E}X_k - \mathbb{E}X_N = \mathbb{E}(K \cdot X)_k \geq \mathbb{E}(K \cdot X)_0 = 0.$$

We will see below that [1, Theorem 4.4.1] is very useful. The first indication of this is: □

Theorem ([1, Theorem 4.4.2])**.** *Doob's Inequality. Let X_m be a submartingale, $\lambda > 0$, and*

$$\bar{X}_n = \max_{0 \leq m \leq n} X_m^+.$$

Then

$$\lambda P(\bar{X}_n \geq \lambda) \leq \mathbb{E}\left[X_n 1(\bar{X}_n \geq \lambda)\right] \leq \mathbb{E}X_n^+. \tag{1}$$

Proof. Let $A = \{\bar{X}_n \geq \lambda\}$ and $N = \inf\{m : X_m \geq \lambda \text{ or } m = n\}$, then N is a stopping time. Since $X_N \geq \lambda$ on A,

$$\lambda P(A) \leq \mathbb{E}\left[X_N 1_A\right].$$

And then [1, Theorem 4.4.1] implies $\mathbb{E}X_N \leq \mathbb{E}X_n$, and we have $X_N = X_n$ on A^c, so we have

$$\mathbb{E}\left[X_N 1_A\right] \leq \mathbb{E}\left[X_n 1_A\right].$$

And combining these gives the first inequality of (1) as

$$\lambda P(\bar{X}_n \geq \lambda) \leq \mathbb{E}\left[X_n 1(\bar{X}_n \geq \lambda)\right].$$

The second inequality of (1) is trivial, so the proof is complete. □

Example ([1, Example 4.4.3])**.** Random Walks. If we let $S_n = \xi_1 + \cdots + \xi_n$ where the ξ_m are independent and have $\mathbb{E}\xi_m = 0$, $\sigma_m^2 = \mathbb{E}\xi_m^2 < \infty$, then S_n is a martingale, so [1, Theorem 4.2.6] implies $X_n = S_n^2$ is a submartingale. If we let $\lambda = x^2$ and apply [1, Theorem 4.4.2] to X_n, we get Kolmogorov's maximal inequality ([1, Theorem 2.5.5]):

$$P\left(\max_{1 \leq m \leq n} |S_m| \geq x\right) \leq x^{-2}\mathrm{var}(S_n).$$

A consequence of [1, Theorem 4.4.2] is:

Theorem ([1, Theorem 4.4.4])**.** *L^p Maximum Inequality. If X_n is a submartingale, then for $1 < p < \infty$,*

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

Consequently, if Y_n is a martingale and $Y_n^* = \max_{0 \leq m \leq n} |Y_m|$, then

$$\mathbb{E}|Y_n^*|^p \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}|Y_n|^p.$$

(not covered in class). The second inequality follows by applying the first to $X_n = |Y_n|$. To prove the first we will work with $\bar{X}_{n\wedge M}$ rather than $\bar{X}_n$. Since $\{\bar{X}_{n\wedge M} \geq \lambda\}$ is always $\{\bar{X}_n \geq \lambda\}$ or $\emptyset$, this does not change the application of Doob's inequality [1, Theorem 4.4.2].

Using [1, Lemma 2.2.13], Doob's inequality [1, Theorem 4.4.2], Fubini's theorem, and a little calculus gives

$$\begin{aligned}
\mathbb{E}\left[(\bar{X}_{n\wedge M})^p\right] &= \int_0^\infty p\lambda^{p-1} P(\bar{X}_{n\wedge M} \geq \lambda)\, d\lambda \\
&\leq \int_0^\infty p\lambda^{p-1}\left(\lambda^{-1}\int X_n^+ 1_{\{\bar{X}_{n\wedge M} \geq \lambda\}} dP\right) d\lambda \\
&= \int X_n^+ \int_0^{\bar{X}_{n\wedge M}} p\lambda^{p-2}\, d\lambda\, dP \\
&= \frac{p}{p-1}\int X_n^+ (\bar{X}_{n\wedge M})^{p-1}\, dP.
\end{aligned}$$

If we let $q = p/(p-1)$ be the conjugate exponent to p and apply Hölder's inequality ([1, Theorem 1.6.3]), we see that the above

$$\leq \frac{p}{p-1}(\mathbb{E}|X_n^+|^p)^{1/p}(\mathbb{E}|\bar{X}_{n\wedge M}|^p)^{1/q}.$$

If we divide both sides of the last inequality by $(\mathbb{E}|\bar{X}_{n\wedge M}|^p)^{1/q}$, which is finite thanks to the $\wedge M$, then take the pth power of each side, we get

$$\mathbb{E}|\bar{X}_{n\wedge M}|^p \le \left(\frac{p}{p-1}\right)^p \mathbb{E}(X_n^+)^p.$$

Letting $M \to \infty$ and using the monotone convergence theorem gives the desired result. $\square$

Example ([1, Example 4.4.5])**.** There is no L^1 maximal inequality. Again, the counterexample is provided by [1, Example 4.2.13].

Let $S_0 = 1$ and $\{S_n,\ n \ge 1\}$ be i.i.d. symmetric simple random walk. That is, $S_n = S_{n-1} + \xi_n$ where $\xi_1, \xi_2, \cdots$ are i.i.d. with $P(\xi_i = 1) = P(\xi_i = -1) = \frac{1}{2}$. Let $N = \min\{n :\ S_n = 0\}$ and let $X_n = S_{N\wedge n}$. Then [1, Theorem 4.4.1] implies

$$\mathbb{E}X_n = \mathbb{E}S_{N\wedge n} = \mathbb{E}S_0 = 1 \quad \text{for all } n.$$

Using hitting probabilities for simple random walk from [1, Theorem 4.4.1], we have

$$P\left(\max_m X_m \ge M\right) = \frac{1}{M}. \tag{2}$$

Hence

$$\mathbb{E}\left[\max_m X_m\right] = \sum_{M=1}^{\infty} P\left(\max_m X_m \ge M\right) = \sum_{M=1}^{\infty} \frac{1}{M} = \infty.$$

The monotone convergence theorem implies that $\mathbb{E}\max_{m\le n} X_m \uparrow \infty$ as $n \uparrow \infty$.

From L^p Maximum Inequality [1, Theorem 4.4.4], we get the following:

Theorem ([1, Theorem 4.4.6])**.** *L^p Convergence Theorem (L^p 수렴정리). If X_n is a martingale with $\sup \mathbb{E}|X_n|^p < \infty$ where $p > 1$, then $X_n \to X$ a.s. and in L^p.*

Proof.

$$(\mathbb{E}X_n^+)^p \le (\mathbb{E}|X_n|)^p \le \mathbb{E}|X_n|^p,$$

so it follows from the martingale convergence theorem (마팅게일 수렴정리) [1, Theorem 4.2.11] that $X_n \to X$ a.s..

The second conclusion in L^p Maximum Inequality [1, Theorem 4.4.4] implies

$$\mathbb{E}\left(\sup_{0\le m\le n} |X_m|\right)^p \le \left(\frac{p}{p-1}\right)^p \mathbb{E}|X_n|^p \le \left(\frac{p}{p-1}\right)^p \sup \mathbb{E}|X_n|^p.$$

Letting $n \to \infty$ and using the monotone convergence theorem implies $\sup |X_n| \in L^p$. Since $|X_n - X|^p \le (2\sup|X_n|)^p$, it follows from the dominated convergence theorem that

$$\mathbb{E}|X_n - X|^p \to 0.$$

$\square$

The most important special case of the results in this lecture note occurs when $p = 2$. To treat this case, we take some results for martingales in L^2 from [2, Chapter 12].

Theorem ([1, Theorem 4.4.7])**.** *Orthogonality of Martingale Increments. Let X_n be a martingale with $\mathbb{E}X_n^2 < \infty$ for all n. If $m \le n$ and $Y \in \mathcal{F}_m$ has $\mathbb{E}Y^2 < \infty$, then*

$$\mathbb{E}\left[(X_n - X_m)Y\right] = 0,$$

and hence if $\ell < m < n$,

$$\mathbb{E}\left[(X_n - X_m)(X_m - X_\ell)\right] = 0.$$

Proof. The Cauchy–Schwarz inequality implies $\mathbb{E}|(X_n - X_m)Y| < \infty$. Using the law of total expectation, [1, Theorem 4.1.14], and the definition of a martingale,

$$\begin{aligned}\mathbb{E}\left[(X_n - X_m)Y\right] &= \mathbb{E}\left[\mathbb{E}\left[(X_n - X_m)Y \mid \mathcal{F}_m\right]\right] \\ &= \mathbb{E}\left[Y\mathbb{E}\left[X_n - X_m \mid \mathcal{F}_m\right]\right] = 0.\end{aligned}$$

□

Hence the formula

$$X_n = X_0 + \sum_{k=1}^{n}(X_k - X_{k-1})$$

expresses X_n as the sum of orthogonal terms, and Pythagoras's theorem and the law of total expectation yields

$$\begin{aligned}\mathbb{E}X_n^2 &= \mathbb{E}X_0^2 + \sum_{k=1}^{n}\mathbb{E}\left[(X_k - X_{k-1})^2\right] \\ &= \mathbb{E}X_0^2 + \sum_{k=1}^{n}\mathbb{E}\left[\mathbb{E}\left[(X_k - X_{k-1})^2|\mathcal{F}_{k-1}\right]\right]. \qquad (3)\end{aligned}$$

This is what we use for Branching Processes below, but one can also use conditional variance formula below, as how [1, Example 4.4.9] is explained in [1].

Theorem ([1, Theorem 4.4.7])**.** *Conditional Variance Formula. If X_n is a martingale with $\mathbb{E}X_n^2 < \infty$ for all n, and $m \leq n$, then*

$$\mathbb{E}\left[(X_n - X_m)^2 \mid \mathcal{F}_m\right] = \mathbb{E}\left[X_n^2 \mid \mathcal{F}_m\right] - X_m^2.$$

Remark. This is the conditional analogue of $\mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2$ and is proved in exactly the same way.

Proof. Using the linearity of conditional expectation and then [1, Theorem 4.1.14], we have

$$\begin{aligned}\mathbb{E}\left[X_n^2 - 2X_nX_m + X_m^2 \mid \mathcal{F}_m\right] &= \mathbb{E}\left[X_n^2 \mid \mathcal{F}_m\right] - 2X_m\mathbb{E}\left[X_n \mid \mathcal{F}_m\right] + X_m^2 \\ &= \mathbb{E}\left[X_n^2 \mid \mathcal{F}_m\right] - 2X_m^2 + X_m^2,\end{aligned}$$

which gives the desired result. □

Example ([1, Example 4.4.9])**.** Branching Processes. We now see the L^2 convergence of Branching Processes. To recall, we had:

Let ξ_i^n, $i, n \geq 1$, be i.i.d. nonnegative integer-valued random variables. Define a stochastic process $\{Z_n\}_{n\geq 0}$ by $Z_0 = 1$ and

$$Z_{n+1} = \begin{cases} \xi_1^{n+1} + \cdots + \xi_{Z_n}^{n+1}, & \text{if } Z_n > 0, \\ 0, & \text{if } Z_n = 0. \end{cases}$$

We suppose $\mu = \mathbb{E}\xi_i^m > 1$ and $\mathrm{var}(\xi_i^m) = \sigma^2 < \infty$. Let $X_n = Z_n/\mu^n$, then we have seen in [1, Lemma 4.3.9\] that X_n is a martingale.

Now we apply (3) to have

$$\mathbb{E}X_n^2 = \mathbb{E}X_0^2 + \sum_{m=1}^{n}\mathbb{E}\left[\mathbb{E}\left[(X_m - X_{m-1})^2|\mathcal{F}_{m-1}\right]\right]. \qquad (4)$$

To compute this, we observe

$$\begin{aligned}\mathbb{E}\left[(X_m - X_{m-1})^2 | \mathcal{F}_{m-1}\right] &= \mathbb{E}\left[(Z_m/\mu^m - Z_{m-1}/\mu^{m-1})^2 \mid \mathcal{F}_{m-1}\right] \\ &= \mu^{-2m}\mathbb{E}\left[(Z_m - \mu Z_{m-1})^2 \mid \mathcal{F}_{m-1}\right].\end{aligned}$$

It follows from [1, Theorem 4.1.2] that on $\{Z_{m-1} = k\}$,

$$\mathbb{E}\left[(Z_m - \mu Z_{m-1})^2 \mid \mathcal{F}_{m-1}\right] = \mathbb{E}\left[\left(\sum_{i=1}^{k} \xi_i^m - \mu k\right)^2 \Big| \mathcal{F}_{m-1}\right] = k\sigma^2 = Z_{m-1}\sigma^2.$$

Hence

$$\mathbb{E}\left[\mathbb{E}\left[(X_m - X_{m-1})^2 | \mathcal{F}_{m-1}\right]\right] = \mu^{-2m}\mathbb{E}\left[Z_{m-1}\sigma^2\right] = \sigma^2/\mu^{m+1},$$

since $\mathbb{E}\left[Z_{m-1}/\mu^{m-1}\right] = \mathbb{E}Z_0 = 1$. Now $\mathbb{E}X_0^2 = 1$, so applying to (4) gives

$$\mathbb{E}X_n^2 = 1 + \sigma^2 \sum_{m=2}^{n+1} \mu^{-m}.$$

This shows $\sup \mathbb{E}X_n^2 < \infty$, so $X_n \to X$ in L^2, and hence $\mathbb{E}X_n \to \mathbb{E}X$. $\mathbb{E}X_n = 1$ for all n, so $\mathbb{E}X = 1$ and $X \not\equiv 0$.

Moreover, it follows from [1, Exercise 4.3.11] that $\{X > 0\} = \{Z_n > 0 \text{ for all } n\}$.

References

[1] Rick Durrett. *Probability—theory and examples*, volume 49 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2019. Fifth edition.

[2] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.