18.175 Theory of Probability
Fall 2008

Section 14

Optional stopping. Inequalities for martingales.

Consider a sequence of σ-algebras $(\mathcal{B}_n)_{n\geq 0}$ such that $\mathcal{B}_n \subseteq \mathcal{B}_{n+1}$. Integer valued r.v. $\tau \in \{1,2,\ldots\}$ is called a *stopping time* if $\{\tau \leq n\} \in \mathcal{B}_n$. Let us denote by $\mathcal{B}_\tau$ a σ-algebra of the events B such that

$$\{\tau \leq n\} \cap B \in \mathcal{B}_n, \ \forall n \geq 1.$$

If (X_n) is adapted to $(\mathcal{B}_n)$ then random variables such as X_τ or $\sum_{k=1}^{\tau} X_k$ are measurable on $\mathcal{B}_\tau$. For example,

$$\{X_\tau \in A\} = \bigcup_{n\geq 1} \{\tau = n\} \cap \{X_n \in A\} = \bigcup_{n\geq 1} \Big(\{\tau \leq n\} \setminus \{\tau \leq n-1\} \bigcap \{X_n \in A\}\Big) \in \mathcal{B}_\tau.$$

Theorem 31 *(Optional stopping) Let $(X_n, \mathcal{B}_n)$ be a martingale and $\tau_1, \tau_2 < \infty$ be stopping times such that*

$$\mathbb{E}|X_{\tau_2}| < \infty, \ \lim_{n\to\infty} \mathbb{E}|X_n|\mathrm{I}(n \leq \tau_2) = 0. \tag{14.0.1}$$

Then on the event $\{\tau_1 \leq \tau_2\}$

$$\mathbb{E}(X_{\tau_2}|\mathcal{B}_{\tau_1}) = X_{\tau_1}.$$

More precisely, for any set $A \in \mathcal{B}_{\tau_1}$,

$$\mathbb{E}X_{\tau_2}\mathrm{I}_A\mathrm{I}(\tau_1 \leq \tau_2) = \mathbb{E}X_{\tau_1}\mathrm{I}_A\mathrm{I}(\tau_1 \leq \tau_2).$$

If $(X_n, \mathcal{B}_n)$ is a submartingale then equality is replaced by $\geq$.

Remark. If stopping times τ_1, τ_2 are bounded then (14.0.1) is satisfied. As the next example shows, without some control of the stopping times the statement is not true.

Example. Consider an i.i.d. sequence (X_n) such that

$$\mathbb{P}(X_n = \pm 2^n) = \frac{1}{2}.$$

If $\mathcal{B}_n = \sigma(X_1, \ldots, X_n)$ then $(S_n, \mathcal{B}_n)$ is a martingale. Let $\tau_1 = 1$ and $\tau_2 = \min\{k \geq 1, S_k > 0\}$. Clearly, $S_{\tau_2} = 2$ because if $\tau_2 = k$ then

$$S_{\tau_2} = S_k = -2 - 2^2 - \ldots - 2^{k-1} + 2^k = 2.$$

However,

$$2 = \mathbb{E}(S_{\tau_2}|\mathcal{B}_1) \neq S_{\tau_1} = X_1.$$

The second condition in (14.0.1) is violated since $\mathbb{P}(\tau_2 = n) = 2^{-n}$ and

$$\mathbb{E}|S_n|I(n \le \tau_2) = 2\mathbb{P}(\tau_2 = n) + (2^{n+1} - 2)\mathbb{P}(n+1 \le \tau_2) = 2 \not\to 0.$$

Proof of Theorem 31. Consider a set $A \in \mathcal{B}_{\tau_1}$. We have,

$$\begin{aligned}
\mathbb{E}X_{\tau_2}\mathrm{I}_A\mathrm{I}(\tau_1 \le \tau_2) &= \sum_{n\ge 1}\mathbb{E}X_{\tau_2}\mathrm{I}\big(A \cap \{\tau_1 = n\}\big)\mathrm{I}(n \le \tau_2) \\
&\overset{(*)}{=} \sum_{n\ge 1}\mathbb{E}X_n\mathrm{I}\big(A \cap \{\tau_1 = n\}\big)\mathrm{I}(n \le \tau_2) = \mathbb{E}X_{\tau_1}\mathrm{I}_A\mathrm{I}(\tau_1 \le \tau_2).
\end{aligned}$$

To prove (*) it is enough to prove that for $A_n = A \cap \{\tau_1 = n\} \in \mathcal{B}_n$,

$$\mathbb{E}X_{\tau_2}\mathrm{I}_{A_n}\mathrm{I}(n \le \tau_2) = \mathbb{E}X_n\mathrm{I}_{A_n}\mathrm{I}(n \le \tau_2). \tag{14.0.2}$$

We can write

$$\begin{aligned}
\mathbb{E}X_n\mathrm{I}_{A_n}\mathrm{I}(n \le \tau_2) &= \mathbb{E}X_n\mathrm{I}_{A_n}\mathrm{I}(\tau_2 = n) + \mathbb{E}X_n\mathrm{I}_{A_n}\mathrm{I}(n+1 \le \tau_2) \\
&= \mathbb{E}X_{\tau_2}\mathrm{I}_{A_n}\mathrm{I}(\tau_2 = n) + \mathbb{E}X_n\mathrm{I}_{A_n}\mathrm{I}(n+1 \le \tau_2) \\
\Big\{\text{ since } \{n+1 \le \tau_2\} = \{\tau_2 \le n\}^c \in \mathcal{B}_n, &\text{ by martingale property}\Big\} \\
&= \mathbb{E}X_{\tau_2}\mathrm{I}_{A_n}\mathrm{I}(\tau_2 = n) + \mathbb{E}X_{n+1}\mathrm{I}_{A_n}\mathrm{I}(n+1 \le \tau_2) \\
\Big\{\text{ by induction }\Big\} &= \sum_{n\le k<m}\mathbb{E}X_{\tau_2}\mathrm{I}_{A_n}\mathrm{I}(\tau_2 = k) + \mathbb{E}X_m\mathrm{I}_{A_n}\mathrm{I}(m \le \tau_2) \\
&= \mathbb{E}X_{\tau_2}\mathrm{I}_{A_n}\mathrm{I}(n \le \tau_2 < m) + \mathbb{E}X_m\mathrm{I}_{A_n}\mathrm{I}(m \le \tau_2).
\end{aligned}$$

By (14.0.1), the last term

$$\big|\mathbb{E}X_m\mathrm{I}_{A_n}\mathrm{I}(m \le \tau_2)\big| \le \mathbb{E}|X_m|\mathrm{I}(m \le \tau_2) \to 0 \text{ as } m \to \infty.$$

Since

$$X_{\tau_2}\mathrm{I}_{A_n}\mathrm{I}(n \le \tau_2 \le m) \to X_{\tau_2}\mathrm{I}_{A_n}\mathrm{I}(n \le \tau_2) \text{ as } m \to \infty$$

and $\mathbb{E}|X_{\tau_2}| < \infty$, by dominated convergence theorem,

$$\mathbb{E}X_{\tau_2}\mathrm{I}_{A_n}\mathrm{I}(n \le \tau_2 < m) \to \mathbb{E}X_{\tau_2}\mathrm{I}_{A_n}\mathrm{I}(n \le \tau_2).$$

This proves (14.0.2).

□

Theorem 32 *(Doob's inequality) If $(X_n, \mathcal{B}_n)$ is a submartingale then for $Y_n = \max_{1\le k\le n} X_k$ and $M > 0$*

$$\mathbb{P}\Big(Y_n \ge M\Big) \le \frac{1}{M}\mathbb{E}X_n I(Y_n \ge M) \le \frac{1}{M}\mathbb{E}X_n^+. \tag{14.0.3}$$

Proof. Define a stopping time

$$\tau_1 = \begin{cases} \min\{k : X_k \ge M, k \le n\} & \text{if such } k \text{ exists,} \\ n & \text{otherwise.} \end{cases}$$

Let $\tau_2 = n$ so that $\tau_1 \le \tau_2$. By Theorem 31,

$$\mathbb{E}(X_n|\mathcal{B}_{\tau_1}) = \mathbb{E}(X_{\tau_2}|\mathcal{B}_{\tau_1}) \ge X_{\tau_1}.$$

Let us apply this to the set $A = \{Y_n = \max_{1\le k\le n} X_n \ge M\}$ which belongs to $\mathcal{B}_{\tau_1}$ because

$$A \cap \{\tau_1 \le k\} = \Big\{\max_{1\le i\le k} X_i \ge M\Big\} \in \mathcal{B}_k.$$

On the event A, $X_{\tau_1} \geq M$ and, therefore,

$$\mathbb{E}X_n \mathrm{I}_A = \mathbb{E}X_{\tau_2}\mathrm{I}_A \geq \mathbb{E}X_{\tau_1}\mathrm{I}_A \geq M\mathbb{E}\mathrm{I}_A = M\mathbb{P}(A).$$

On the other hand, $\mathbb{E}X_n \mathrm{I}_A \leq \mathbb{E}X_n^+$ and this finishes the proof.

□

As a corollary we obtain the *second Kolmogorov's inequality.* If (X_i) are independent and $\mathbb{E}X_i = 0$ then $S_n = \sum_{1\leq i\leq n} X_i$ is a martingale and S_n^2 is a submartingale. Therefore,

$$\mathbb{P}\Big(\max_{1\leq k\leq n} |S_k| \geq M\Big) = \mathbb{P}\Big(\max_{1\leq k\leq n} S_k^2 \geq M^2\Big) \leq \frac{1}{M^2}\mathbb{E}S_n^2 = \frac{1}{M^2}\sum_{1\leq k\leq n} \mathrm{Var}(X_k).$$

Exercises.
1. Show that for any random variable Y, $\mathbb{E}|Y^p| = \int_0^\infty pt^{p-1}\mathbb{P}(|Y| \geq t)dt$.
2. Let X, Y be two non-negative random variables such that for every $t > 0$, $\mathbb{P}(Y \geq t) \leq t^{-1}\int XI(Y \geq t)d\mathbb{P}$. For any $p > 1$, $\|f\|_p = (\int |f|^p d\mathbb{P})^{1/p}$ and $1/p + 1/q = 1$, show that $\|Y\|_p \leq q\|X\|_p$.
3. Given a non-negative submartingale $(X_n, \mathcal{B}_n)$, let $X_n^* := \max_{j\leq n} X_j$ and $X^* := \max_{j\geq 1} X_j$. Prove that for any $p > 1$ and $1/p + 1/q = 1$, $\|X^*\|_p \leq q\sup_n \|X_n\|_p$. *Hint:* use exercise 2 and Doob's maximal inequality.

□

Doob's upcrossing inequality. Let $(X_n, \mathcal{B}_n)_{n\geq 1}$ be a submartingale. Given two real numbers $a < b$ we will define a sequence of stopping times (τ_n) when X_n is crossing a downward and b upward as in figure 14.1. Namely, we define

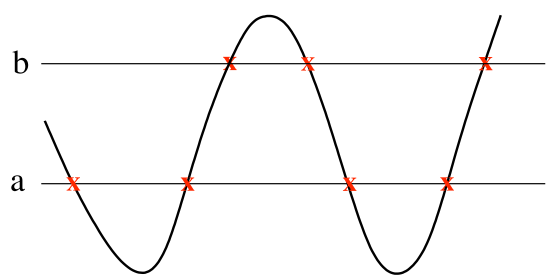

Figure 14.1: Stopping times of level crossings.

$$\tau_1 = \min\{n \geq 1, X_n \leq a\}, \ \ \tau_2 = \min\{n > \tau_2 : X_n \geq b\}$$

and, by induction, for $k \geq 2$

$$\tau_{2k-1} = \min\{n > \tau_{2k-2}, X_n \leq a\}, \ \ \tau_{2k} = \min\{n > 2k-1, X_n \geq b\}.$$

Define

$$\nu(a, b, n) = \max\{k : \tau_{2k} \leq n\}$$

- the number of upward crossings of $[a, b]$ before time n.

Theorem 33 *(Doob's upcrossing inequality) We have,*

$$\mathbb{E}\nu(a, b, n) \leq \frac{\mathbb{E}(X_n - a)^+}{b - a}. \tag{14.0.4}$$

Proof. Since $x \to (x - a)^+$ is increasing convex function, $Z_n = (X_n - a)^+$ is also a submartingale. Clearly,

$$\mu_X(a, b, n) = \nu_Z(0, b - a, n)$$

which means that it is enough to prove (14.0.4) for nonnegative submartingales. From now on we can assume that $0 \leq X_n$ and we would like to show that

$$\mathbb{E}\nu(0, b, n) \leq \frac{\mathbb{E}X_n}{b}.$$

Let us define a sequence of r.v.s

$$\eta_j = \begin{cases} 1, & \tau_{2k-1} < j \le \tau_{2k} \text{ for some } k \\ 0, & \text{otherwise,} \end{cases}$$

i.e. η_j is the indicator of the event that at time j the process is crossing $[0, b]$ upward. Define $X_0 = 0$. Then

$$b\nu(0,b,n) \le \sum_{j=1}^{n} \eta_j (X_j - X_{j-1}) = \sum_{j=1}^{n} \mathrm{I}(\eta_j = 1)(X_j - X_{j-1}).$$

The event

$$\{\eta_j = 1\} = \bigcup_k \{\tau_{2k-1} < j \le \tau_{2k}\} = \bigcup_k \overbrace{\Big\{\tau_{2k-1} \le j-1\Big\}}^{\in \mathcal{B}_{j-1}} \Big\backslash \overbrace{\Big\{\tau_{2k} \le j-1\Big\}}^{\in \mathcal{B}_{j-1}}{}^{c} \in \mathcal{B}_{j-1}$$

i.e. the fact that at time j we are crossing upward is determined completely by the sequence up to time $j-1$. Then

$$\begin{aligned} b\mathbb{E}\nu(0,b,n) &\le \sum_{j=1}^{n} \mathbb{E}\mathbb{E}\Big(\mathrm{I}(\eta_j = 1)(X_j - X_{j-1})\Big|\mathcal{B}_{j-1}\Big) = \sum_{j=1}^{n} \mathbb{E}\mathrm{I}(\eta_j = 1)\mathbb{E}(X_j - X_{j-1}|\mathcal{B}_{j-1}) \\ &= \sum_{j=1}^{n} \mathbb{E}\mathrm{I}(\eta_j = 1)(\mathbb{E}(X_j|\mathcal{B}_{j-1}) - X_{j-1}) \le \sum_{j=1}^{n} \mathbb{E}(X_j - X_{j-1}) = \mathbb{E}X_n, \end{aligned}$$

where in the last inequality we used that $(X_j, \mathcal{B}_j)$ is a submartingale, $\mathbb{E}(X_j|\mathcal{B}_{j-1}) \ge X_{j-1}$, which implies that

$$\mathrm{I}(\eta_j = 1)(\mathbb{E}(X_j|\mathcal{B}_{j-1}) - X_{j-1}) \le \mathbb{E}(X_j|\mathcal{B}_{j-1}) - X_{j-1}.$$

This finishes the proof.

□