Martingales

Last updated by Serik Sagitov: May 23, 2013

Abstract

This Stochastic Processes course is based on the book Probabilities and Random Processes by Geoffrey Grimmett and David Stirzaker. Chapters 7.7-7.8, and 12.

1 Definitions and examples

Example 1 Martingale: a betting strategy. Let X_n be the gain of a gambler doubling the bet after each loss. The game stops after the first win.

- $X_0 = 0$
- $X_1 = 1$ with probability 1/2 and $X_1 = -1$ with probability 1/2,
- $X_2 = 1$ with probability 3/4 and $X_2 = -3$ with probability 1/4,
- $X_3 = 1$ with probability 7/8 and $X_3 = -7$ with probability $1/8, \ldots,$
- $X_n = 1$ with probability $1 - 2^{-n}$ and $X_n = -2^n + 1$ with probability 2^{-n}.

Conditional expectation

$$\mathbb{E}(X_{n+1}|X_n) = (2X_n - 1)\frac{1}{2} + (1)\frac{1}{2} = X_n.$$

If N is the number of games, then $\mathbb{P}(N = n) = 2^{-n}$, $n = 1, 2, \ldots$ with $\mathbb{E}(N) = 2$ and

$$\mathbb{E}(X_{N-1}) = \mathbb{E}(1 - 2^{N-1}) = 1 - \sum_{n=1}^{\infty} 2^{n-1}2^{-n} = -\infty.$$

Definition 2 *A sequence of sigma-fields $(\mathcal{F}_n)$ such that $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_n \subset \ldots \subset \mathcal{F}$ is called a filtration. A sequence of r.v. (Y_n) is called adapted to $(\mathcal{F}_n)$ if Y_n is $\mathcal{F}_n$-measurable for all n. In this case the sequence $(Y_n, \mathcal{F}_n)$ is called a martingale if, for all $n \geq 0$,*

- $\mathbb{E}(|Y_n|) < \infty$ *(* $\mathbb{E}(Y_n^+) < \infty$, $\mathbb{E}(Y_n^-) < \infty$*),*
- $\mathbb{E}(Y_{n+1}|\mathcal{F}_n) = Y_n$ *(* $\geq Y_n$, $\leq Y_n$*).*

Definition 3 *Let (Y_n) be adapted to a filtration $(\mathcal{F}_n)$. Then $(Y_n, \mathcal{F}_n)$ is called a submartingale if*

- $\mathbb{E}(Y_n^+) < \infty$,
- $\mathbb{E}(Y_{n+1}|\mathcal{F}_n) \geq Y_n$.

Definition 4 *Let (Y_n) be adapted to a filtration $(\mathcal{F}_n)$. Then $(Y_n, \mathcal{F}_n)$ is called a supermartingale if*

- $\mathbb{E}(Y_n^-) < \infty$,
- $\mathbb{E}(Y_{n+1}|\mathcal{F}_n) \leq Y_n$.

Consider the sequence of means $m_n = \mathbb{E}(Y_n)$. We have $m_{n+1} \geq m_n$ for submartingales, $m_{n+1} \leq m_n$ for supermartingales, and $m_{n+1} = m_n$ for martingales. A martingale is both a sub- and supermartingale. If (Y_n) is a submartingale, then $(-Y_n)$ is a supermartingale.

Example 5 Consider a simple random walk $S_n = X_1 + \ldots + X_n$ with $\mathbb{P}(X_i = 1) = p$, $\mathbb{P}(X_i = -1) = q$, and $S_0 = k$. The centered $S_n - n(p - q)$ is a martingale:

$$\mathbb{E}(S_{n+1} - (n+1)(p-q)|X_1, \ldots, X_n) = S_n + \mathbb{E}(X_{n+1}) - (n+1)p = S_n - n(p-q).$$

Another martingale is $Y_n = (q/p)^{S_n}$:

$$\mathbb{E}(Y_{n+1}|X_1, \ldots, X_n) = p(q/p)^{S_n+1} + q(q/p)^{S_n-1} = (q/p)^{S_n} = Y_n$$

with $\mathbb{E}(Y_n) = \mathbb{E}(Y_0) = (q/p)^k$. It is called De Moivre's martingale.

Example 6 Stopped de Moivre's martingale. Consider the same simple random walk and suppose that it stops if it hits 0 or N which is larger than the initial state k. Denote $D_n = Y_{T \wedge n}$, where T is the stopping time of the random walk and Y_n is the de Moivre martingale. It is easy to see that D_n is also a martingale. Put $\mathbb{P}(S_T = 0) = \mathbb{P}(Y_T = 1) = p_k$ and $\mathbb{P}(S_T = N) = \mathbb{P}(Y_T = (q/p)^N) = 1 - p_k$. From $\mathbb{E}(Y_T) = \mathbb{E}(Y_0)$ (to be proved later) we derive

$$(q/p)^0 p_k + (q/p)^N (1 - p_k) = (q/p)^k \quad \Rightarrow \quad p_k = \frac{(p/q)^{N-k} - 1}{(p/q)^N - 1}$$

as long as $p \neq q$.

Example 7 Let $S_n = X_1 + \ldots + X_n$, where X_i are iid r.v. with zero means and finite variances σ^2. Then $S_n^2 - n\sigma^2$ is a martingale

$$\mathbb{E}(S_{n+1}^2 - (n+1)\sigma^2|X_1, \ldots, X_n) = S_n^2 + 2S_n\mathbb{E}(X_{n+1}) + \mathbb{E}(X_{n+1}^2) - (n+1)\sigma^2 = S_n^2 - n\sigma^2.$$

Example 8 Branching processes. Let Z_n be a branching process with $Z_0 = 1$ and the mean offspring number μ. Since $\mathbb{E}(Z_{n+1}|Z_n) = \mu Z_n$, the ratio $W_n = \mu^{-n} Z_n$ is a martingale.

In the supercritical case, $\mu > 1$, the extinction probability $\eta \in [0, 1)$ of Z_n is identified as a solution of the equation $\eta = h(\eta)$, where $h(s) = \mathbb{E}(s^X)$ is the generating function of the offspring number. The process $V_n = \eta^{Z_n}$ is also a martingale

$$\mathbb{E}(V_{n+1}|Z_1, \ldots, Z_n) = \mathbb{E}(\eta^{X_1 + \ldots + X_{Z_n}}|Z_1, \ldots, Z_n) = h(\eta)^{Z_n} = V_n.$$

Example 9 Doob's martingale. Let Z be a r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}(|Z|) < \infty$. For a filtration $(\mathcal{F}_n)$ define $Y_n = \mathbb{E}(Z|\mathcal{F}_n)$. The $(Y_n, \mathcal{F}_n)$ is a martingale: first, by Jensen's inequality,

$$\mathbb{E}(|Y_n|) = \mathbb{E}|\mathbb{E}(Z|\mathcal{F}_n)| \leq \mathbb{E}(\mathbb{E}(|Z|\,|\mathcal{F}_n)) = \mathbb{E}(|Z|),$$

and secondly

$$\mathbb{E}(Y_{n+1}|\mathcal{F}_n) = \mathbb{E}(\mathbb{E}(Z|\mathcal{F}_{n+1})|\mathcal{F}_n)) = \mathbb{E}(Z|\mathcal{F}_n) = Y_n.$$

As we show next the Doob martingale is uniformly integrable. Again due to Jensen's inequality,

$$|Y_n| = |\mathbb{E}(Z|\mathcal{F}_n)| \leq \mathbb{E}(|Z|\,|\mathcal{F}_n) =: Z_n,$$

so that $|Y_n|1_{\{|Y_n| \geq a\}} \leq Z_n 1_{\{Z_n \geq a\}}$. By the definition of conditional expectation $\mathbb{E}\big((|Z| - Z_n)1_{\{Z_n \geq a\}}\big) = 0$ and we have

$$|Y_n|1_{\{|Y_n| \geq a\}} \leq |Z|1_{\{Z_n \geq a\}}$$

which entails uniform integrability, since $\mathbb{P}(Z_n \geq a) \to 0$ by the Markov inequality.

2 Convergence in L^2

Lemma 10 *If (Y_n) is a martingale with $\mathbb{E}(Y_n^2) < \infty$, then $Y_{n+1} - Y_n$ and Y_n are uncorrelated. It follows that (Y_n^2) is a submartingale. More generally, if $J(x)$ is convex, then $J(Y_n)$ is a submartingale.*

Proof. The first assertion follows from

$$\mathbb{E}(Y_n(Y_{n+1}-Y_n)|\mathcal{F}_n) = Y_n(\mathbb{E}(Y_{n+1}|\mathcal{F}_n) - Y_n) = 0.$$

The second claim is derived as follows

$$\begin{aligned}\mathbb{E}(Y_{n+1}^2|\mathcal{F}_n) &= \mathbb{E}((Y_{n+1}-Y_n)^2 + 2S_n(Y_{n+1}-Y_n) + Y_n^2|\mathcal{F}_n)\\ &= \mathbb{E}((Y_{n+1}-Y_n)^2|\mathcal{F}_n) + Y_n^2 \geq Y_n^2.\end{aligned}$$

Notice that

$$\mathbb{E}(Y_{n+1}^2) = \mathbb{E}(Y_n^2) + \mathbb{E}((Y_{n+1}-Y_n)^2)$$

so that $\mathbb{E}(Y_n^2)$ is non-decreasing and there always exists a finite or infinite limit

$$M = \lim_{n\to\infty} \mathbb{E}(Y_n^2). \tag{1}$$

The third assertion is due to the Jensen inequality.

Lemma 11 *Doob-Kolmogorov's inequality. If (Y_n) is a martingale with $\mathbb{E}(Y_n^2) < \infty$, then for any $\epsilon > 0$*

$$\mathbb{P}(\max_{1\leq i\leq n} |Y_i| \geq \epsilon) \leq \frac{\mathbb{E}(Y_n^2)}{\epsilon^2}.$$

Proof. Let $B_k = \{|Y_1| < \epsilon, \ldots, |Y_{k-1}| < \epsilon, |Y_k| \geq \epsilon\}$. Then using a submartingale property for the second inequality we get

$$\mathbb{E}(Y_n^2) \geq \sum_{i=1}^n \mathbb{E}(Y_n^2 1_{B_i}) \geq \sum_{i=1}^n \mathbb{E}(Y_i^2 1_{B_i}) \geq \epsilon^2 \sum_{i=1}^n \mathbb{P}(B_i) = \epsilon^2 \mathbb{P}(\max_{1\leq i\leq n} |Y_i| \geq \epsilon).$$

Theorem 12 *If (Y_n) is a martingale with finite M defined by (1), then there exists a random variable Y such that $Y_n \to Y$ a.s. and in mean square.*

Proof. Step 1. For

$$A_m(\epsilon) = \bigcup_{i\geq 1} \{|Y_{m+i} - Y_m| \geq \epsilon\}$$

we will show that

$$\mathbb{P}(A_m(\epsilon)) \to 0, \quad m \to \infty \text{ for any } \epsilon > 0. \tag{2}$$

Put $S_n = Y_{m+n} - Y_m$. It is also a martingale, since

$$\mathbb{E}(S_{n+1}|S_1, \ldots, S_n) = \mathbb{E}(\mathbb{E}(S_{n+1}|\mathcal{F}_{m+n})|S_1, \ldots, S_n) = \mathbb{E}(S_n|S_1, \ldots, S_n) = S_n.$$

Apply the Doob-Kolmogorov inequality to this martingale to find that

$$\mathbb{P}(|Y_{m+i} - Y_m| \geq \epsilon \text{ for some } i \in [1, n]) \leq \epsilon^{-2}\mathbb{E}((Y_{m+n} - Y_m)^2) = \epsilon^{-2}(\mathbb{E}(Y_{m+n}^2) - \mathbb{E}(Y_m^2)).$$

Letting $n \to \infty$ we obtain $\mathbb{P}(A_m(\epsilon)) \leq \epsilon^{-2}(M - \mathbb{E}(Y_m^2))$ and hence (2).

Step 2. Show that the sequence (Y_n) is a.s. Cauchy convergent:

$$\mathbb{P}\Big(\bigcap_{\epsilon>0} \bigcup_{m\geq 1} A_m^c(\epsilon)\Big) = 1$$

which implies the existence of Y such that $Y_n \to Y$ a.s. Indeed, since $A_m(\epsilon_1) \subset A_m(\epsilon_2)$ for $\epsilon_1 > \epsilon_2$, we have

$$\mathbb{P}\Big(\bigcup_{\epsilon>0} \bigcap_{m\geq 1} A_m(\epsilon)\Big) = \lim_{\epsilon\to 0} \mathbb{P}\Big(\bigcap_{m\geq 1} A_m(\epsilon)\Big) \leq \lim_{\epsilon\to 0} \lim_{m\to\infty} \mathbb{P}(A_m(\epsilon)) = 0.$$

Step 3. Prove the convergence in mean square using the Fatou lemma

$$\begin{aligned}\mathbb{E}((Y_n - Y)^2) &= \mathbb{E}(\liminf_{m\to\infty}(Y_n - Y_m)^2) \leq \liminf_{m\to\infty} \mathbb{E}((Y_n - Y_m)^2)\\ &= \liminf_{m\to\infty} \mathbb{E}(Y_m^2) - \mathbb{E}(Y_n^2) = M - \mathbb{E}(Y_n^2) \to 0, \quad n \to \infty.\end{aligned}$$

Example 13 Branching processes. Let Z_n be a branching process with $Z_0 = 1$ and the offspring numbers having mean μ and variance σ^2. The ratio $W_n = \mu^{-n} Z_n$ is a martingale with

$$\mathbb{E}(W_n^2) = 1 + (\sigma/\mu)^2 (1 + \mu^{-1} + \ldots + \mu^{-n+1}).$$

In the supercritical case, $\mu > 1$, we have $\mathbb{E}(W_n^2) \to 1 + \frac{\sigma^2}{\mu(\mu-1)}$, and there is a r.v. W such that $W_n \to W$ a.s. and in L^2. The Laplace transform of the limit $\phi(\theta) = \mathbb{E}(e^{-\theta W})$ satisfies a functional equation $\phi(\mu\theta) = h(\phi(\theta))$.

3 Doob's decomposition

Definition 14 *The sequence $(S_n, \mathcal{F}_n)$ is called predictable if $S_0 = 0$, and S_n is $\mathcal{F}_{n-1}$-measurable for all $n \geq 1$. It is also called increasing if $\mathbb{P}(S_n \leq S_{n+1}) = 1$ for all $n \geq 0$.*

Theorem 15 *Doob's decomposition. A submartingale $(Y_n, \mathcal{F}_n)$ with finite means can be expressed in the form $Y_n = M_n + S_n$, where $(M_n, \mathcal{F}_n)$ is a martingale and $(S_n, \mathcal{F}_n)$ is an increasing predictable process (called the compensator of the submartingale). This decomposition is unique.*

Proof. We define M and S explicitly: $M_0 = Y_0$, $S_0 = 0$, and for $n \geq 0$

$$M_{n+1} - M_n = Y_{n+1} - \mathbb{E}(Y_{n+1}|\mathcal{F}_n), \qquad S_{n+1} - S_n = \mathbb{E}(Y_{n+1}|\mathcal{F}_n) - Y_n.$$

To see uniqueness suppose another such decomposition $Y_n = M'_n + S'_n$. Then

$$M'_{n+1} - M'_n + S'_{n+1} - S'_n = M_{n+1} - M_n + S_{n+1} - S_n.$$

Taking conditional expectations given $\mathcal{F}_n$ we get $S'_{n+1} - S'_n = S_{n+1} - S_n$. This in view of $S'_0 = S_0 = 0$ implies $S'_n = S_n$.

Definition 16 *Let (Y_n) be adapted to $(\mathcal{F}_n)$ and (S_n) be predictable. The sequence*

$$Z_n = Y_0 + \sum_{i=1}^{n} S_i (Y_i - Y_{i-1})$$

is called the transform of (Y_n) by (S_n).

Example 17 Such transforms are usually interpreted as gambling systems with (Y_n) being a supermartingale (the capital after n gambles each involving a unit stake). Optional skipping is one such strategy. Here the gambler either wagers a unit stake or skip the round: S_n equals either 1 or 0.

Theorem 18 *Let (Z_n) be the transform of (Y_n) by (S_n). Then*
(i) If (Y_n) is a martingale, then (Z_n) is a martingale so long as $\mathbb{E}|Z_n| < \infty$ for all n.
(ii) If (Y_n) is a submartingale and in addition $S_n \geq 0$ for all n, then (Z_n) is a submartingale so long as $\mathbb{E}(Z_n^+) < \infty$ for all n.

Proof. Both assertions follow immediately from

$$\mathbb{E}(Z_{n+1}|\mathcal{F}_n) - Z_n = \mathbb{E}(S_{n+1}(Y_{n+1} - Y_n)|\mathcal{F}_n) = S_{n+1}(\mathbb{E}(Y_{n+1}|\mathcal{F}_n) - Y_n).$$

Example 19 Optional stopping. The gambler wagers a unit stake on each play until the random time T. In this case $S_n = 1_{\{n \leq T\}}$ and $Z_n = Y_{T \wedge n}$. If S_n is predictable, then $\{T = n\} = \{S_n = 1, S_{n+1} = 0\} \in \mathcal{F}_n$, so that T is a stopping time.

Example 20 Optional starting. The gambler does not play until the $(T + 1)$-th round, where T is a stopping time. In this case $S_n = 1_{\{T \leq n-1\}}$ is a predictable sequence.

4 Hoeffding's inequality

Definition 21 *Let $(Y_n, \mathcal{F}_n)$ be a martingale. The sequence of martingale differences is defined by $D_n = Y_n - Y_{n-1}$, so that D_n is $\mathcal{F}_n$-measurable*

$$\mathbb{E}|D_n| < \infty, \qquad \mathbb{E}(D_{n+1}|\mathcal{F}_n) = 0, \qquad Y_n = Y_0 + D_1 + \ldots + D_n.$$

Theorem 22 *Hoeffding's inequality. Let $(Y_n, \mathcal{F}_n)$ be a martingale, and suppose $\mathbb{P}(|D_n| \le K_n) = 1$ for a sequence of real numbers K_n. Then for any $x > 0$*

$$\mathbb{P}(|Y_n - Y_0| \ge x) \le 2\exp\Big(- \frac{x^2}{2(K_1^2 + \ldots + K_n^2)}\Big).$$

Proof. Let $\theta > 0$.

Step 1. The function $e^{\theta x}$ is convex, therefore

$$e^{\theta d} \le \frac{1}{2}(1-d)e^{-\theta} + \frac{1}{2}(1+d)e^{\theta} \text{ for all } |d| \le 1.$$

Hence if D is a r.v. with mean 0 such that $\mathbb{P}(|D| \le 1) = 1$, then $\mathbb{E}(e^{\theta D}) \le \frac{e^{-\theta}+e^{\theta}}{2} < e^{\theta^2/2}$.

Step 2. Using the martingale differences we obtain

$$\mathbb{E}(e^{\theta(Y_n - Y_0)}|\mathcal{F}_{n-1}) = e^{\theta(Y_{n-1}-Y_0)}\mathbb{E}(e^{\theta D_n}|\mathcal{F}_{n-1}) \le e^{\theta(Y_{n-1}-Y_0)}e^{\theta^2 K_n^2/2}.$$

Take expectations and iterate to find

$$\mathbb{E}(e^{\theta(Y_n-Y_0)}) \le \mathbb{E}(e^{\theta(Y_{n-1}-Y_0)})e^{\theta^2 K_n^2/2} \le \exp\Big(\frac{\theta^2}{2}\sum_{i=1}^n K_i^2\Big).$$

Step 3. Due to the Markov inequality we have for any $x > 0$

$$\mathbb{P}(Y_n - Y_0 \ge x) \le e^{-\theta x}\mathbb{E}(e^{\theta(Y_n-Y_0)}) \le \exp\Big(- \theta x + \frac{\theta^2}{2}\sum_{i=1}^n K_i^2\Big).$$

Set $\theta = x/\sum_{i=1}^n K_i^2$ to minimize the exponent. Then

$$\mathbb{P}(Y_n - Y_0 \ge x) \le \exp\Big(- \frac{x^2}{2(K_1^2 + \ldots + K_n^2)}\Big).$$

Since $(-Y_n)$ is also a martingale, we get

$$\mathbb{P}(Y_n - Y_0 \le -x) = \mathbb{P}(-Y_n + Y_0 \ge x) \le \exp\Big(- \frac{x^2}{2(K_1^2 + \ldots + K_n^2)}\Big).$$

Example 23 Large deviations. Let X_n be iid Bernoulli (p) r.v. If $S_n = X_1 + \ldots + X_n$, then $Y_n = S_n - np$ is a martingale. Due to the Hoeffding's inequality for any $x > 0$

$$\mathbb{P}(|S_n - np| \ge x\sqrt{n}) \le 2\exp\Big(- \frac{x^2}{2(\max(p, 1-p))^2}\Big).$$

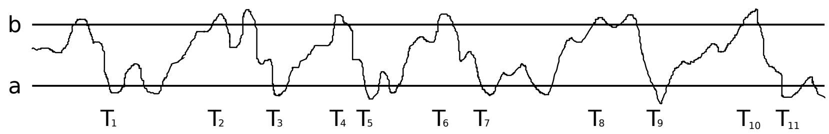
In particular, if $p = 1/2$,

$$\mathbb{P}(|S_n - n/2| \ge x\sqrt{n}) \le 2e^{-2x^2}.$$

Putting here $x = 3$ we get $\mathbb{P}(|S_n - n/2| \ge 3\sqrt{n}) \le 3 \cdot 10^{-8}$.

5 Convergence in L^1

On the figure below five uppcrossing time intervals are shown: $(T_1, T_2]$, $(T_3, T_4]$, $\ldots$, $(T_9, T_{10}]$. If for all rational intervals (a, b) the number of uppcrossings $U(a, b)$ is finite, then the corresponding trajectory has a (possibly infinite) limit.



Lemma 24 *Snell's uppcrossing inequality. Let $a < b$ and $U_n(a, b)$ is the number of uppcrossings of a submartingale $(Y_0, \ldots, Y_n)$. Then $\mathbb{E}(U_n(a, b)) \leq \frac{\mathbb{E}((Y_n - a)^+)}{b-a}$.*

Proof. Since $Z_n = (Y_n - a)^+$ forms a submartingale, it is enough to prove $\mathbb{E}(U_n(0, c)) \leq \frac{\mathbb{E}(Z_n)}{c}$, where $U_n(0, c)$ is the number of uppcrossings of the submartingale $(Z_0, \ldots, Z_n)$. Let I_i be the indicator of the event that $i \in (T_{2k-1}, T_{2k}]$ for some k. Note that I_i is $\mathcal{F}_{i-1}$-measurable, since

$$\{I_i = 1\} = \bigcup_k \{T_{2k-1} \leq i - 1\} \setminus \{T_{2k} \leq i - 1\}$$

is an event that depends on $(Y_0, \ldots, Y_{i-1})$ only. Therefore,

$$\begin{aligned}\mathbb{E}((Z_i - Z_{i-1})I_i) &= \mathbb{E}(\mathbb{E}((Z_i - Z_{i-1})I_i|\mathcal{F}_{i-1})) = \mathbb{E}(I_i(\mathbb{E}(Z_i|\mathcal{F}_{i-1}) - Z_{i-1})) \\ &\leq \mathbb{E}(\mathbb{E}(Z_i|\mathcal{F}_{i-1}) - Z_{i-1}) = \mathbb{E}(Z_i) - \mathbb{E}(Z_{i-1}).\end{aligned}$$

It remains to observe that

$$c \cdot U_n(0, c) \leq \sum_{i=1}^{n} (Z_i - Z_{i-1})I_i \Rightarrow c \cdot \mathbb{E}(U_n(0, c)) \leq \mathbb{E}(Z_n) - \mathbb{E}(Z_0) \leq \mathbb{E}(Z_n).$$

Theorem 25 *Suppose $(Y_n, \mathcal{F}_n)$ is a submartingale such that $\mathbb{E}(Y_n^+) \leq M$ for some constant M and all n. (i) There exists a r.v. Y such that $Y_n \to Y$ almost surely. In addition: (ii) the limit Y has a finite mean if $\mathbb{E}|Y_0| < \infty$, and (iii) if (Y_n) is uniformly integrable, then $Y_n \to Y$ in L^1.*

Proof. (i) Using Snell's inequality we obtain that $U(a, b) = \lim U_n(a, b)$ satisfies

$$\mathbb{E}(U(a, b)) \leq \frac{M + |a|}{b - a}.$$

Therefore, $\mathbb{P}(U(a, b) < \infty) = 1$. Since there are only countably many rationals, it follows that with probability 1, $U(a, b) < \infty$ for all rational (a, b), and $Y_n \to Y$ almost surely.

(ii) We have to check that $\mathbb{E}|Y| < \infty$ given $\mathbb{E}|Y_0| < \infty$. Indeed, since $|Y_n| = 2Y_n^+ - Y_n$ and $\mathbb{E}(Y_n|\mathcal{F}_0) \geq Y_0$, we get

$$\mathbb{E}(|Y_n| \big| \mathcal{F}_0) \leq 2\mathbb{E}(Y_n^+ \big| \mathcal{F}_0) - Y_0.$$

By Fatou's lemma

$$\mathbb{E}(|Y| \big| \mathcal{F}_0) = \mathbb{E}(\liminf_{n\to\infty} |Y_n| \big| \mathcal{F}_0) \leq \liminf_{n\to\infty} \mathbb{E}(|Y_n| \big| \mathcal{F}_0) \leq 2 \liminf_{n\to\infty} \mathbb{E}(Y_n^+ \big| \mathcal{F}_0) - Y_0,$$

and it remains to observe that $\mathbb{E}(\liminf_{n\to\infty} \mathbb{E}(Y_n^+ \big| \mathcal{F}_0)) \leq M$, again due to Fatou's lemma.

(iii) Finally, recall that given $Y_n \stackrel{\mathrm{P}}{\to} Y$, the uniform integrability of (Y_n) is equivalent to $\mathbb{E}|Y_n| < \infty$ for all n, $\mathbb{E}|Y| < \infty$, and $Y_n \stackrel{L^1}{\to} Y$.

Corollary 26 *Any martingale, submartingale or supermartingale $(Y_n, \mathcal{F}_n)$ satisfying $\sup_n \mathbb{E}|Y_n| \leq M$ converges almost surely to a r.v. with a finite limit.*

Corollary 27 *A non-negative supermartingale converges almost surely. A non-positive submartingale converges almost surely.*

Example 28 De Moivre martingale $Y_n = (q/p)^{S_n}$ is non-negative and hence converges a.s. to some limit Y. Let $p \neq q$. Since $S_n \to \infty$ for $p > q$ and $S_n \to -\infty$ for $p < q$ we have $Y = 0$. Note that Y_n does not converge in mean, since $\mathbb{E}(Y_n) = \mathbb{E}(Y_0) \neq 0$.

Example 29 Doob's martingale $Y_n = \mathbb{E}(Z|\mathcal{F}_n)$ is uniformly integrable, see Example 9. It converges a.s. and in mean to $\mathbb{E}(Z|\mathcal{F}_\infty)$, where $\mathcal{F}_\infty$ is the smallest σ-algebra containing all $\mathcal{F}_n$. There an important converse result: if a martingale $(Y_n, \mathcal{F}_n)$ converges in mean, then there exists a r.v. Z with finite mean such that $Y_n = \mathbb{E}(Z|\mathcal{F}_n)$.

6 Bounded stopping times. Optional sampling theorem

Definition 30 *A r.v T taking values in $\{0, 1, 2, \ldots\} \cup \{\infty\}$ is called a stopping time with respect to the filtration $(\mathcal{F}_n)$, if $\{T = n\} \in \mathcal{F}_n$ for all $n \geq 0$. It is called a bounded stopping time if $\mathbb{P}(T \leq N) = 1$ for some finite constant N.*

We denote by $\mathcal{F}_T$ the σ-algebra of all events A such that $A \cap \{T = n\} \in \mathcal{F}_n$ for all n.

The stopped de Moivre martingale from Example 6 is also a martingale. A general statement of this type follows next.

Theorem 31 *Let $(Y_n, \mathcal{F}_n)$ be a submartingale and let T be a stopping time. Then $(Y_{T\wedge n}, \mathcal{F}_n)$ is a submartingale. If moreover, $\mathbb{E}|Y_n| < \infty$, then $(Y_n - Y_{T\wedge n}, \mathcal{F}_n)$ is also a submartingale.*

Proof. The r.v. $Z_n = Y_{T\wedge n}$ is $\mathcal{F}_n$-measurable:

$$Z_n = \sum_{i=0}^{n-1} Y_i 1_{\{T=i\}} + Y_n 1_{\{T\geq n\}},$$

and

$$\mathbb{E}(Z_n^+) \leq \sum_{i=0}^{n} \mathbb{E}(Y_i^+) < \infty.$$

It remains to see that $Z_{n+1} - Z_n = (Y_{n+1} - Y_n) 1_{\{T>n\}}$ implies

$$\mathbb{E}(Z_{n+1} - Z_n|\mathcal{F}_n) = \mathbb{E}(Y_{n+1} - Y_n|\mathcal{F}_n) 1_{\{T>n\}} \geq 0.$$

with

$$0 \leq \mathbb{E}(Y_{n+1} - Y_n|\mathcal{F}_n) 1_{\{T>n\}} \leq \mathbb{E}(Y_{n+1} - Y_n|\mathcal{F}_n).$$

Corollary 32 *If $(Y_n, \mathcal{F}_n)$ is a martingale, then it is both a submartingale and a supermartingale, and therefore, for a given stopping time T, both $(Y_{T\wedge n}, \mathcal{F}_n)$ and $(Y_n - Y_{T\wedge n}, \mathcal{F}_n)$ are martingales.*

Theorem 33 *Optional sampling. Let $(Y_n, \mathcal{F}_n)$ be a submartingale.*

(i) If T is a bounded stopping time, then $\mathbb{E}(Y_T^+) < \infty$ and $\mathbb{E}(Y_T|\mathcal{F}_0) \geq Y_0$.

(ii) If $0 = T_0 \leq T_1 \leq T_2 \leq \ldots$ is a sequence of bounded stopping times, then $(Y_{T_j}, \mathcal{F}_{T_j})$ is a submartingale.

Proof. (i) Let $\mathbb{P}(T \leq N) = 1$ where N is a positive constant. Since $(Y_{T\wedge n})$ is a submartingale and $Y_{T\wedge N} = Y_T$, we have $\mathbb{E}(Y_T^+) < \infty$ and $\mathbb{E}(Y_T|\mathcal{F}_0) \geq Y_0$.

(ii) Consider two bounded stopping times $S \leq T \leq N$. To show that $\mathbb{E}(Y_T|\mathcal{F}_S) \geq Y_S$ observe that for $A \in \mathcal{F}_S$ we have

$$\mathbb{E}(Y_T 1_A) = \sum_{k\leq N} \mathbb{E}(Y_T 1_{A\cap\{S=k\}}) = \sum_{k\leq N} \mathbb{E}\Big(1_{A\cap\{S=k\}} \mathbb{E}(Y_T|\mathcal{F}_k)\Big),$$

and since in view of Theorem 31, $\mathbb{E}(Y_T|\mathcal{F}_k) = \mathbb{E}(Y_{T\wedge N}|\mathcal{F}_k) \geq Y_{T\wedge k}$ for all $k \leq N$, we conclude

$$\mathbb{E}(Y_T 1_A) \geq \mathbb{E}\Big(\sum_{k\leq N} 1_{A\cap\{S=k\}} Y_{T\wedge k}\Big) = \mathbb{E}\Big(\sum_{k\leq N} 1_{A\cap\{S=k\}} Y_k\Big) = \mathbb{E}(Y_S 1_A).$$

Example 34 The process Y_n is a martingale iff it is both a submartingale and a supermartingale. Therefore, according to Theorem 33 (i) we have $\mathbb{E}(Y_T|\mathcal{F}_0) = Y_0$ and $\mathbb{E}(Y_T) = \mathbb{E}(Y_0)$ for any bounded stopping time T. This martingale property is not enough for Example 6 because the ruin time is not bounded. However, see Theorem 35.

7 Unbounded stopping times

Theorem 35 *Optional stopping. Let $(Y_n, \mathcal{F}_n)$ be a martingale and T be a stopping time. Then $\mathbb{E}(Y_T) = \mathbb{E}(Y_0)$ if (a) $\mathbb{P}(T < \infty) = 1$, (b) $\mathbb{E}|Y_T| < \infty$, and (c) $\mathbb{E}(Y_n 1_{\{T>n\}}) \to 0$ as $n \to \infty$.*

Proof. From $Y_T = Y_{T\wedge n} + (Y_T - Y_n)1_{\{T>n\}}$ using that $\mathbb{E}(Y_{T\wedge n}) = \mathbb{E}(Y_0)$ we obtain

$$\mathbb{E}(Y_T) = \mathbb{E}(Y_0) + \mathbb{E}(Y_T 1_{\{T>n\}}) - \mathbb{E}(Y_n 1_{\{T>n\}}).$$

It remains to apply (c) and observe that due to the dominated convergence $\mathbb{E}(Y_T 1_{\{T>n\}}) \to 0$.

Theorem 36 *Let $(Y_n, \mathcal{F}_n)$ be a martingale and T be a stopping time. Then $\mathbb{E}(Y_T) = \mathbb{E}(Y_0)$ if (a) $\mathbb{E}(T) < \infty$ and (b) there exists a constant c such that for any n*

$$\mathbb{E}(|Y_{n+1} - Y_n|\big|\mathcal{F}_n)1_{\{T>n\}} \le c1_{\{T>n\}}.$$

Proof. Since $T \wedge n \to T$, we have $Y_{T\wedge n} \to Y_T$ a.s. It follows that

$$\mathbb{E}(Y_0) = \mathbb{E}(Y_{T\wedge n}) \to \mathbb{E}(Y_T)$$

as long as $(Y_{T\wedge n})$ is uniformly integrable. To prove the uniform integrability it is enough to verify that $\mathbb{E}(W) < \infty$, where

$$|Y_{T\wedge n}| \le |Y_0| + W, \quad W := |Y_1 - Y_0| + \ldots + |Y_T - Y_{T-1}|.$$

Indeed, since $\mathbb{E}(|Y_i - Y_{i-1}|1_{\{T\ge i\}}\big|\mathcal{F}_{i-1}) \le c1_{\{T\ge i\}}$, we have $\mathbb{E}(|Y_i - Y_{i-1}|1_{\{T\ge i\}}) \le c\mathbb{P}(T \ge i)$ and therefore

$$\mathbb{E}(W) = \sum_{i=1}^{\infty} \mathbb{E}(|Y_i - Y_{i-1}|1_{\{T\ge i\}}) \le c\mathbb{E}(T) < \infty.$$

Example 37 Wald's equality. Let (X_n) be iid r.v. with finite mean μ and $S_n = X_1 + \ldots + X_n$, then $Y_n = S_n - n\mu$ is a martingale with respect to $\mathcal{F}_n = \sigma\{X_1, \ldots, X_n\}$. Now

$$\mathbb{E}(|Y_{n+1} - Y_n|\big|\mathcal{F}_n) = \mathbb{E}|X_{n+1} - \mu| = \mathbb{E}|X_1 - \mu| < \infty.$$

We deduce from Theorem 36 that $\mathbb{E}(Y_T) = \mathbb{E}(Y_0)$ for any stopping time T with finite mean, implying that $\mathbb{E}(S_T) = \mu\mathbb{E}(T)$.

Lemma 38 *Wald's identity. Let (X_n) be iid r.v. with $M(t) = \mathbb{E}(e^{tX})$ and $S_n = X_1 + \ldots + X_n$. If T is a stopping time with finite mean such that $|S_n|1_{\{T>n\}} \le c1_{\{T>n\}}$, then*

$$\mathbb{E}(e^{tS_T} M(t)^{-T}) = 1 \text{ whenever } M(t) \ge 1.$$

Proof. Define $Y_0 = 1$, $Y_n = e^{tS_n}M(t)^{-n}$, and let $\mathcal{F}_n = \sigma\{X_1, \ldots, X_n\}$. It is clear that (Y_n) is a martingale and thus the claim follows from Theorem 36. To verify condition (b) note that

$$\mathbb{E}(|Y_{n+1} - Y_n|\big|\mathcal{F}_n) = Y_n\mathbb{E}|e^{tX}M(t)^{-1} - 1| \le Y_n\mathbb{E}(e^{tX}M(t)^{-1} + 1) = 2Y_n.$$

Furthermore, given $M(t) \ge 1$

$$Y_n = e^{tS_n}M(t)^{-n} \le e^{c|t|} \text{ for } n < T.$$

Example 39 Simple random walk S_n with $\mathbb{P}(X_i = 1) = p$ and $\mathbb{P}(X_i = -1) = q$. Let T be the first exit time of $(-a, b)$. By Lemma 38 with $M(t) = pe^t + qe^{-t}$,

$$e^{-at}\mathbb{E}(M(t)^{-T}1_{\{S_T=-a\}}) + e^{bt}\mathbb{E}(M(t)^{-T}1_{\{S_T=b\}}) = 1 \text{ whenever } M(t) \ge 1.$$

Setting $M(t) = s^{-1}$ we obtain a quadratic equation for e^t having two solutions

$$\lambda_1(s) = \frac{1 + \sqrt{1 - 4pqs^2}}{2ps}, \qquad \lambda_2(s) = \frac{1 - \sqrt{1 - 4pqs^2}}{2ps}, \quad s \in [0, 1].$$

They give us two linear equations resulting in

$$\mathbb{E}(s^T 1_{\{S_T = -a\}}) = \frac{\lambda_1^a \lambda_2^a (\lambda_1^b - \lambda_2^b)}{\lambda_1^{a+b} - \lambda_2^{a+b}}, \qquad \mathbb{E}(s^T 1_{\{S_T = b\}}) = \frac{\lambda_1^a - \lambda_2^a}{\lambda_1^{a+b} - \lambda_2^{a+b}}.$$

Summing up these two relations we get the probability generating function

$$\mathbb{E}(s^T) = \frac{\lambda_1^a (1 - \lambda_2^{a+b}) + \lambda_2^a (\lambda_1^{a+b} - 1)}{\lambda_1^{a+b} - \lambda_2^{a+b}}.$$

8 Maximal inequality

Theorem 40 *Maximal inequality.*

(i) If (Y_n) is a submartingale, then for any $\epsilon > 0$

$$\mathbb{P}(\max_{0 \le i \le n} Y_i \ge \epsilon) \le \frac{\mathbb{E}(Y_n^+)}{\epsilon}.$$

(ii) If (Y_n) is a supermartingale and $\mathbb{E}|Y_0| < \infty$, then for any $\epsilon > 0$

$$\mathbb{P}(\max_{0 \le i \le n} Y_i \ge \epsilon) \le \frac{\mathbb{E}(Y_0) + \mathbb{E}(Y_n^-)}{\epsilon}.$$

(iii) Moreover, if (Y_n) is a submartingale, then for any $\epsilon > 0$

$$\mathbb{P}(\max_{0 \le i \le n} |Y_i| \ge \epsilon) \le \frac{2\mathbb{E}(Y_n^+) - \mathbb{E}(Y_0)}{\epsilon}.$$

In particular, if (Y_n) is a martingale, then

$$\mathbb{P}(\max_{0 \le i \le n} |Y_i| \ge \epsilon) \le \frac{\mathbb{E}|Y_n|}{\epsilon}.$$

Proof. (i) If (Y_n) is a submartingale, then (Y_n^+) is a non-negative submartingale with finite means and

$$T := \min\{n : Y_n \ge \epsilon\} = \min\{n : Y_n^+ \ge \epsilon\}.$$

By Theorem 31, $\mathbb{E}(Y_{T \wedge n}^+) \le \mathbb{E}(Y_n^+)$. Therefore,

$$\mathbb{E}(Y_n^+) \ge \mathbb{E}(Y_{T \wedge n}^+) = \mathbb{E}(Y_T^+ 1_{\{T \le n\}}) + \mathbb{E}(Y_n^+ 1_{\{T > n\}}) \ge \mathbb{E}(Y_T^+ 1_{\{T \le n\}}) \ge \epsilon \mathbb{P}(T \le n),$$

implying the first stated inequality as $\{T \le n\} = \{\max_{0 \le i \le n} Y_i \ge \epsilon\}$.

Furthermore, since $\mathbb{E}(Y_{T \wedge n}^+ 1_{\{T > n\}}) = \mathbb{E}(Y_n^+ 1_{\{T > n\}})$, we have

$$\mathbb{E}(Y_n^+ 1_{\{T \le n\}}) \ge E(Y_{T \wedge n}^+ 1_{\{T \le n\}}) = \mathbb{E}(Y_T^+ 1_{\{T \le n\}}) \ge \epsilon \mathbb{P}(T \le n).$$

Using this we get a stronger inequality

$$\mathbb{P}(A) \le \frac{\mathbb{E}(Y_n^+ 1_A)}{\epsilon}, \text{ where } A = \{\max_{0 \le i \le n} Y_i \ge \epsilon\}. \tag{3}$$

(ii) If (Y_n) is a supermartingale, then by Theorem 31 the second assertion follows from

$$\mathbb{E}(Y_0) \ge \mathbb{E}(Y_{T \wedge n}) = \mathbb{E}(Y_T I_{\{T \le n\}}) + \mathbb{E}(Y_n I_{\{T > n\}}) \ge \epsilon \mathbb{P}(T \le n) - \mathbb{E}(Y_n^-).$$

(iii) Let $\epsilon > 0$. If (Y_n) is a submartingale, then $(-Y_n)$ is a supermartingale so that according to (ii),

$$\mathbb{P}(\min_{0 \le i \le n} Y_i \le -\epsilon) \le \frac{\mathbb{E}(Y_n^+) - \mathbb{E}(Y_0)}{\epsilon}.$$

Combine this with (i) to get the asserted inequality.

Corollary 41 *Doob-Kolmogorov's inequality. If* (Y_n) *is a martingale with finite second moments, then* (Y_n^2) *is a submartingale for any* $\epsilon > 0$

$$\mathbb{P}(\max_{1\le i\le n} |Y_i| \ge \epsilon) = \mathbb{P}(\max_{1\le i\le n} Y_i^2 \ge \epsilon^2) \le \frac{\mathbb{E}(Y_n^2)}{\epsilon^2}.$$

Corollary 42 *Kolmogorov's inequality. Let* (X_n) *are iid r.v. with zero means and finite variances* (σ_n^2)*, then for any* $\epsilon > 0$

$$\mathbb{P}(\max_{1\le i\le n} |X_1 + \ldots + X_i| \ge \epsilon) \le \frac{\sigma_1^2 + \ldots + \sigma_n^2}{\epsilon^2}.$$

Theorem 43 *Convergence in* L^r*. Let* $r > 1$*. Suppose* $(Y_n, \mathcal{F}_n)$ *is a martingale such that* $\mathbb{E}(|Y_n|^r) \le M$ *for some constant* M *and all* n*. Then* $Y_n \to Y_\infty$ *in* L^r*, where* Y_∞ *is the a.s. limit of* Y_n *as* $n \to \infty$*.*

Proof. Combining Corollary 26 and Lyapunov's inequality we get the a.s. convergence $Y_n \to Y_\infty$. To prove $Y_n \stackrel{L^r}{\to} Y_\infty$, we observe first that

$$\mathbb{E}\big((\max_{0\le i\le n} |Y_i|)^r\big) \le \mathbb{E}\big((|Y_0| + \ldots + |Y_n|)^r\big) < \infty.$$

Now using (3) we obtain (writing $A(x) = \{\max_{0\le i\le n} |Y_i| \ge x\}$)

$$\begin{aligned}\mathbb{E}\big((\max_{0\le i\le n} |Y_i|)^r\big) &= \int_0^\infty r x^{r-1} \mathbb{P}\big(\max_{0\le i\le n} |Y_i| > x\big) dx \\ &\le \int_0^\infty r x^{r-2} \mathbb{E}\big(|Y_n| 1_{A(x)}\big) dx = \mathbb{E}\big(|Y_n| \int_0^\infty r x^{r-2} 1_{A(x)} dx\big) = \frac{r}{r-1} \mathbb{E}\Big[|Y_n| (\max_{0\le i\le n} |Y_i|)^{r-1}\Big].\end{aligned}$$

By Hölder's inequality,

$$\mathbb{E}\Big[|Y_n| (\max_{0\le i\le n} |Y_i|)^{r-1}\Big] \le \Big[\mathbb{E}(|Y_n|^r)\Big]^{1/r} \Big[\mathbb{E}\big((\max_{0\le i\le n} |Y_i|)^r\big)\Big]^{(r-1)/r}$$

and we conclude

$$\mathbb{E}\big((\max_{0\le i\le n} |Y_i|)^r\big) \le \Big(\frac{r}{r-1}\Big)^r \mathbb{E}(|Y_n|^r) \le \Big(\frac{r}{r-1}\Big)^r M.$$

Thus by monotone convergence $\mathbb{E}\big(\sup_n |Y_n|^r\big) < \infty$ and (Y_n^r) is uniformly integrable, implying $Y_n \stackrel{L^r}{\to} Y_\infty$.

9 Backward martingales

Definition 44 *Let* $(\mathcal{G}_n)$ *be a decreasing sequence of* σ*-algebras and* (Y_n) *be a sequence of adapted r.v. The sequence* $(Y_n, \mathcal{G}_n)$ *is called a backward or reversed martingale if, for all* $n \ge 0$*,*

- $\mathbb{E}(|Y_n|) < \infty$,
- $\mathbb{E}(Y_n | \mathcal{G}_{n+1}) = Y_{n+1}$.

Theorem 45 *Let* $(Y_n, \mathcal{G}_n)$ *be a backward martingale. Then* Y_n *converges to a limit* Y_∞ *almost surely and in mean.*

Proof. The sequence $Y_n = \mathbb{E}(Y_0 | \mathcal{G}_n)$ is uniformly integrable, see the proof in Example 9. Therefore, it suffices to prove a.s. convergence. Applying Lemma 24 to the martingale $(Y_n, \mathcal{G}_n), \ldots, (Y_0, \mathcal{G}_0)$ we obtain $\mathbb{E}(U_n(a,b)) \le \frac{\mathbb{E}((Y_0 - a)^+)}{b-a}$ for the number $U_n(a,b)$ of $[a,b]$ uppcrossings by $(Y_n, \ldots, Y_0)$. Now let $n \to \infty$ and repeat the proof of Theorem 25 to get the required a.s. convergence.

Theorem 46 *Strong LLN. Let* $X_1, X_2, \ldots$ *be iid random variables defined on the same probability space. Then*

$$\frac{X_1 + \ldots + X_n}{n} \stackrel{\text{a.s.}}{\to} \mu$$

for some constant μ *iff* $\mathbb{E}|X_1| < \infty$*. In this case* $\mu = \mathbb{E}X_1$ *and* $\frac{X_1 + \ldots + X_n}{n} \stackrel{L^1}{\to} \mu$*.*

Proof. Set $S_n = X_1 + \ldots + S_n$ and let $\mathcal{G}_n = \sigma(S_n, S_{n+1}, \ldots)$, then

$$\mathbb{E}(S_n|\mathcal{G}_{n+1}) = \mathbb{E}(S_n|S_{n+1}) = n\mathbb{E}(X_1|S_{n+1}).$$

On the other hand,

$$S_{n+1} = \mathbb{E}(S_{n+1}|S_{n+1}) = (n+1)\mathbb{E}(X_1|S_{n+1}).$$

We conclude that S_n/n is a backward martingale, and according to the Backward Martingale Convergence Theorem there exists Y such that $S_n/n \to Y$ a.s. and in mean. By Kolmogorov's zero-one law, Y is almost surely constant, and hence $Y = \mathbb{E}(X_1)$ a.s.

The converse. If $S_n/n \overset{\text{a.s.}}{\to} \mu$, then $X_n/n \overset{\text{a.s.}}{\to} 0$ by the theory of convergent real series. Indeed, from $(a_1 + \ldots + a_n)/n \to \mu$ it follows that

$$\frac{a_n}{n} = \frac{a_1 + \ldots + a_{n-1}}{n(n-1)} + \frac{a_1 + \ldots + a_n}{n} - \frac{a_1 + \ldots + a_{n-1}}{n-1} \to 0$$

Now, in view of $X_n/n \overset{\text{a.s.}}{\to} 0$, the second Borell-Cantelli lemma gives

$$\sum_n \mathbb{P}(|X_n| \geq n) < \infty,$$

since otherwise $\mathbb{P}(n^{-1}|X_n| \geq 1 \text{ i.o.}) = 1$.