Contents

6 Martingales

6a Basic definitions

First, an example.

Monsters[1] of type A have masses $a_1, a_2, \dots, a_m$; monsters of type B — $b_1, b_2, \dots, b_n$ In the first fight, a_1 eats b_1 with probability $\frac{a_1}{a_1+b_1}$, gets the mass $a_1 + b_1$ and then fights b_2; or b_1 eats a_1 and then fights a_2; and so on.[2]

6a1 Proposition. The monsters of type A win with probability $\frac{A}{A+B}$ where $A = a_1 + \dots + a_m$ and $B = b_1 + \dots + b_n$.

Now, the basic definitions.

6a2 Definition. (a) A filtration on a probability space $(\Omega, \mathcal{F}, P)$ is an increasing sequence of sub-σ-algebras: $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{F}$;

(b) An adapted (to the given filtration) process is a sequence $(X_0, X_1, \dots)$ of random variables such that for each k, X_k is $\mathcal{F}_k$-measurable.

A filtered probability space is a probability space endowed with a filtration.

Assume for now that Ω is (finite or) countable (the discrete framework).

6a3 Definition. An adapted process $(X_n)_n$ such that $\forall n\ \mathbb{E}\,|X_n| < \infty$ is

(a) a martingale, if $\forall n\ \mathbb{E}\big(X_{n+1} \,\big|\, \mathcal{F}_n\big) = X_n$ a.s.;

(b) a supermartingale, if $\forall n\ \mathbb{E}\big(X_{n+1} \,\big|\, \mathcal{F}_n\big) \le X_n$ a.s.;

(c) a submartingale, if $\forall n\ \mathbb{E}\big(X_{n+1} \,\big|\, \mathcal{F}_n\big) \ge X_n$ a.s.

We generalize it to arbitrary Ω as follows.[3]

---

[1] Maybe, banks...

[2] This is much simpler than a "special gladiator game" of: K.S. Kaminsky, E.M. Luks, P.I. Nelson (1984) "Strategy, nontransitive dominance and the exponential distribution". Austral. J. Statist. **26**:2, 111–118.

[3] See also Sect. 7c.

6a4 Definition. An adapted process $(X_n)_n$ such that $\forall n\ \mathbb{E}\,|X_n| < \infty$ is

(a) a martingale, if $\forall n\ \ \forall Y \in L_\infty(\mathcal{F}_n)\ \ \mathbb{E}\,\big((X_{n+1} - X_n)Y\big) = 0$;

(b) a supermartingale, if $\forall n\ \ \forall Y \in L_\infty^+(\mathcal{F}_n)\ \ \mathbb{E}\,\big((X_{n+1} - X_n)Y\big) \le 0$;

(c) a submartingale, if $\forall n\ \ \forall Y \in L_\infty^+(\mathcal{F}_n)\ \ \mathbb{E}\,\big((X_{n+1} - X_n)Y\big) \ge 0$.

6a5 Exercise. In the discrete framework, Definitions 6a3 and 6a4 are equivalent.

Prove it.

6a6 Exercise. Replacing in Definition 6a4 $X_{n+1} - X_n$ with $X_{n+k} - X_n$ for $k = 1, 2, \dots$ we get an equivalent definition.

Prove it.

In particular, $Y = 1\!\!1$ gives $\mathbb{E}\,X_n = \mathbb{E}\,X_0$ (a necessary condition).

Assume again the discrete framework.

Every $X \in L_1$ leads to a martingale $M_n = \mathbb{E}\big(X\,\big|\,\mathcal{F}_n\big)$. "Accumulating data", "revising prediction"... Locally it is the general form of a martingale; globally — not.

Here is an explanation of the terms "supermartingale" and "submartingale". Let $(X_n)_{n=1}^N$ be adapted; introduce a martingale $M_n = \mathbb{E}\big(X_N\,\big|\,\mathcal{F}_n\big)$; then:

if (X_n) is a martingale then $X_n = M_n$;

if (X_n) is a supermartingale then $X_n \ge M_n$;

if (X_n) is a submartingale then $X_n \le M_n$.

Conditional Jensen inequality gives: if (M_n) is a martingale and f is convex ("sublinear") then $\big(f(M_n)\big)$ is a submartingale.

6a7 Example. The one-dimensional simple random walk (S_n) is a martingale; (S_n^2) is a submartingale.

Functions on a tree...

Proof of 6a1. Denote by M_n the total mass of A-monsters at time n, then $(M_n)_n$ is a martingale ($\mathcal{F}_n$ being the whole past...) since $b_\ell \cdot \frac{a_k}{a_k+b_\ell} + (-a_k) \cdot \frac{b_\ell}{a_k+b_\ell} = 0$. Thus, $\mathbb{E}\,M_{m+n} = \mathbb{E}\,M_0 = A$; we note that M_{m+n} takes on two values only, 0 and $A + B$. $\square$

6b Gambling strategy, martingale transform, stopping

First, an example.

Let $(S_n)_n$ be the simple random walk, and $T = \inf\{n : |S_n| = 10\}$ (be it finite or infinite).

6b1 Proposition. $T < \infty$ a.s., and $\mathbb{E}\, T = 100$.

Now, the theory.

6b2 Definition. (a) A previsible (with respect to the given filtration) process is a sequence $(C_1, C_2, \dots)$ of random variables such that for each k, C_k is $\mathcal{F}_{k-1}$-measurable.[1]

(b) Given a previsible process $(C_n)_n$ and adapted process $(X_n)_n$ (on the same filtered probability space), we define an adaped process $C \bullet X$ by[2]

$$\begin{aligned} &(C \bullet X)_0 = 0\,, \\ &(C \bullet X)_n = (C \bullet X)_{n-1} + C_n (X_n - X_{n-1})\,. \end{aligned}$$

Thus,

$$\begin{aligned} (C \bullet X)_n &= C_1(X_1 - X_0) + C_2(X_2 - X_1) + \dots + C_n(X_n - X_{n-1}) = \\ &= -C_1 X_0 - (C_2 - C_1) X_1 - \dots - (C_n - C_{n-1}) X_{n-1} + C_n X_n\,. \end{aligned}$$

6b3 Proposition. Let $(M_n)_n$ be a martingale, $(C_n)_n$ previsible, and $C_n(M_n - M_{n-1}) \in L_1$ for all n. Then $C \bullet M$ is a martingale.

Proof. Let $Y \in L_\infty(\mathcal{F}_n)$, then $\mathbb{E}\,\big(((C \bullet M)_{n+1} - (C \bullet M)_n) Y\big) = \mathbb{E}\,\big(C_{n+1}(M_{n+1} - M_n) Y\big)$, and it vanishes if $C_{n+1} Y \in L_\infty$; otherwise apply it to $Y_k = Y \cdot \mathbb{1}_{[-k,k]}(C_{n+1})$ and note that $\sup_k |C_{n+1}(M_{n+1} - M_n) Y_k| \le |C_{n+1}(M_{n+1} - M_n) Y| \le \|Y\|_\infty \cdot \underbrace{|C_{n+1}(M_{n+1} - M_n)|}_{\in L_1}$. □

A sufficient condition: $\forall n \;\; C_n \in L_\infty$.

An important special case:

$$(C^\tau)_n = \mathbb{1}_{n \le \tau} = \begin{cases} 1 & \text{for } n \le \tau, \\ 0 & \text{for } n > \tau, \end{cases}$$

where τ is a stopping time as defined below.

6b4 Definition. A stopping time is a map $\tau : \Omega \to \{0, 1, 2, \dots\} \cup \{\infty\}$ such that $\{\tau \le n\} \in \mathcal{F}_n$ for all n.

[1] In discrete time it look strange, but in continuous time it does not…

[2] Imagine that C_n is the number of your shares of a stock at time n, and X_n is the share price at time n…

Note that

$$\{(C^\tau)_n = 0\} = \{\tau < n\} = \{\tau \le n-1\} \in \mathcal{F}_{n-1} .$$

In terms of a tree, $\{\tau > n\}$ is just a subtree...
The corresponding martingale transform is the stopped process,

$$(C^\tau \bullet X)_n = X_{\tau \wedge n} - X_0 .$$

6b5 Corollary. If $(M_n)_n$ is a martingale and τ a stopping time then the stopped process $(M_{\tau\wedge n})_n$ is also a martingale.

Proof of 6b1. The process M defined by $M_n = S_n^2 - n$ is a martingale (think, why). On the other hand, T is a stopping time (think, why). Thus, the stopped process $M_{T\wedge n}$ is a martingale. Therefore $\mathbb{E}\, M_{T\wedge n} = 0$, that is, $\mathbb{E}\, S^2_{T\wedge n} = \mathbb{E}\,(T \wedge n)$. We get $\mathbb{E}\,(T\wedge n) \le 100$ for all n, and therefore $T < \infty$ a.s., $\mathbb{E}\, T \le 100$; thus S_T is well-defined, $S_{T\wedge n} \to S_T$ a.s., and the bounded convergence theorem gives $\mathbb{E}\, S^2_{T\wedge n} \to \mathbb{E}\, S^2_T = 100$; therefore $\mathbb{E}\, T = 100$. $\square$

By the way, $\mathbb{E}\, S_T = 0$; $\mathbb{P}\big(S_T = -10\big) = 0.5 = \mathbb{P}\big(S_T = 10\big)$.

Recall 3b15: from $\sup_n |S_n| = \infty$ (a.s.) by Kolmogorov's 0-1 law we got $\inf_n S_n = -\infty$ and $\sup_n S_n = \infty$ a.s.

However, do not think that $\mathbb{E}\, M_\tau$ must vanish! Think about $\tau = \min\{n : S_n = +10\}$.[1]

6c Positive martingales

First, an example.

Let Z_n be the size of n-th generation (be it the number of animals, neutrons, or men of a given family). Assume that $Z_0 = 1$ always, and each member of the n-th generation produces a random number of offsprings (members of the next generation): either 2 (with probability p) or 0 (with probability $1-p$). That is, conditionally, given $Z_0, \dots, Z_n$, the distribution of $Z_{n+1}/2$ is binomial,

$$\mathbb{P}\big(Z_{n+1} = 2k \,\big|\, Z_0, \dots, Z_n\big) = \binom{Z_n}{k} p^k (1-p)^{Z_n - k} .$$

This is called the simple branching, or Galton-Watson, process.

For $p \le 0.5$ the process extincts a.s.:

6c1 Proposition. For $p \le 0.5$, $\mathbb{P}\big(\exists n \;\, Z_n = 0\big) = 1$.

[1] Do you like to get rich this way? :-)

For $p > 0.5$ the process either extincts or grows exponentially:

6c2 Proposition. For $p > 0.5$ the limit

$$M_\infty = \lim_{n\to\infty} \frac{Z_n}{(2p)^n}$$

exists and is finite almost surely, and

$$\mathbb{P}\big(\exists n \; Z_n = 0\big) = \mathbb{P}\big(M_\infty = 0\big) = \frac{1-p}{p}, \quad \mathbb{E}\, M_\infty = 1\,.$$

Now, the theory.

Let $(M_n)_n$ be a positive martingale, that is, $M_n \ge 0$ a.s. for every n.

Given $0 < a < b < \infty$, we define stopping times

$$\sigma_1 = \inf\{n \ge 0 : M_n \le a\}\,, \quad \tau_1 = \inf\{n > \sigma_1 : M_n \ge b\}\,,$$
$$\sigma_2 = \inf\{n > \tau_1 : M_n \le a\}\,, \quad \tau_2 = \inf\{n > \sigma_2 : M_n \ge b\}\,,$$

and so on. (As usual, $\inf \emptyset = \infty$.) Now we define the (random) number of upcrossings:

$$U = \sup\{k : \tau_k < \infty\}\,; \quad U : \Omega \to \{0,1,2,\dots\} \cup \{\infty\}\,.$$

6c3 Proposition (Dubins's inequality)**.**

$$\mathbb{P}\big(U \ge k\big) \le \Big(\frac{a}{b}\Big)^k \quad \text{for } k = 0,1,2,\dots$$

Proof. It is sufficient to prove that $\mathbb{P}\big(\tau_k < \infty\big) \le \frac{a}{b}\mathbb{P}\big(\sigma_k < \infty\big)$. We have

$$\mathbb{E}\, M_{\sigma_k \wedge n} = \mathbb{E}\, M_{\tau_k \wedge n} = \mathbb{E}\, M_0\,;$$
$$\underbrace{\mathbb{E}\, M_{\sigma_k \wedge n}}_{=\mathbb{E}\, M_0} = \underbrace{\mathbb{E}\,(M_{\sigma_k}; \sigma_k \le n)}_{\le a\mathbb{P}\,(\sigma_k \le n)} + \mathbb{E}\,(M_n; \sigma_k > n)\,;$$
$$\underbrace{\mathbb{E}\, M_{\tau_k \wedge n}}_{=\mathbb{E}\, M_0} = \underbrace{\mathbb{E}\,(M_{\tau_k}; \tau_k \le n)}_{\ge b\mathbb{P}\,(\tau_k \le n)} + \mathbb{E}\,(M_n; \tau_k > n)\,;$$
$$\mathbb{E}\,(M_n; \tau_k > n) - \mathbb{E}\,(M_n; \sigma_k > n) = \mathbb{E}\,(M_n; \sigma_k \le n < \tau_k) \ge 0\,;$$
$$a\mathbb{P}\,(\sigma_k \le n) \ge b\mathbb{P}\,(\tau_k \le n)\,;$$

take $n \to \infty$. $\square$

The same holds for supermartingales.

6c4 Theorem. Every positive martingale converges a.s. to an integrable random variable.

Proof. By Dubins's inequality, the martingale $(M_n)_n$ cannot cross (a,b) infinitely many times. Almost surely, for all *rational* $a<b$, it crosses (a,b) finitely many times, which excludes the case $\liminf M_n < a < b < \limsup M_n$. It means that $\liminf M_n = \limsup M_n$ a.s. Integrability of the limit follows from Fatou lemma. $\square$

We turn to branching. Let $(Z_n)_n$ be the simple branching process introduced in Sect. 6c.

Given $\mathcal{F}_n$ we have $\frac12 Z_{n+1} \sim \operatorname{Binom}(Z_n,p)$, thus $\mathbb{E}\big(Z_{n+1}\,\big|\,\mathcal{F}_n\big) = 2pZ_n$, which shows that $M_n = \frac{1}{(2p)^n} Z_n$ is a (positive) martingale. By 6c4, $M_n \to M_\infty$ a.s., $\mathbb{E}\, M_\infty \le 1$.

Proof of 6c1. Case $p<0.5$: we have $\sup_n \frac{1}{(2p)^n} Z_n < \infty$, that is, $Z_n = O\big((2p)^n\big)$, a.s., which ultimately excludes the case $Z_n \ge 1$; extinction.

Case $p=0.5$: $Z_n \to M_\infty$ a.s.; we have to prove that $M_\infty = 0$ a.s. Assuming the contrary we take $k>0$ such that $\mathbb{P}\big(M_\infty = k\big) > 0$, then

$$\begin{gathered} 1\!\!1_{Z_n=k} \to 1\!\!1_{M_\infty=k}\,; \\ 1\!\!1_{Z_n=k,Z_{n+1}=k} \to 1\!\!1_{M_\infty=k}\,; \\ \mathbb{P}\big(Z_n=k\big) \to \mathbb{P}\big(M_\infty=k\big)\,; \\ \mathbb{P}\big(Z_n=k, Z_{n+1}=k\big) \to \mathbb{P}\big(M_\infty=k\big)\,; \\ \mathbb{P}\big(Z_{n+1}=k\,\big|\,Z_n=k\big) \to 1\,, \end{gathered}$$

that is, the distribution $\operatorname{Binom}(k,0.5)$ is concentrated at $0.5k$, — a contradiction. $\square$

More detailed information on the branching process can be obtained using the generating functions

$$f_n(\theta) = \mathbb{E}\,\theta^{Z_n}\,.$$

We have $f_0(\theta)=\theta$; $f_1(\theta) = p\theta^2+1-p$; $\mathbb{E}\big(\theta^{Z_{n+1}}\,\big|\,Z_n=k\big) = \big(f_1(\theta)\big)^k$ (think, why); thus $f_{n+1}(\theta) = \mathbb{E}\,\big(f_1(\theta)\big)^{Z_n} = f_n\big(f_1(\theta)\big)$, that is,

$$f_n = f_1 \circ \dots \circ f_1 \quad (n \text{ times}).$$

Iterations for $f_n(0) = \mathbb{P}\big(Z_n=0\big)$ converge (draw a picture!) to the first root of the equation $f_1(\theta)=\theta$. Taking into account that $f_1(1)=1$ we solve the equation easily: $\theta = (1-p)/p$. We get

$$\mathbb{P}\big(Z_n=0\big) \to \frac{1-p}{p} = \mathbb{P}\big(\exists n\ Z_n=0\big)\,.$$

In order to prove 6c2 it remains to prove that $\mathbb{E}\,M_\infty = 1$ and $\mathbb{P}\big(M_\infty = 0\big) = \mathbb{P}\big(\exists n\ Z_n = 0\big)$. In order to prove the former it is sufficient to prove that $M_n \to M_\infty$ in L_1, or in L_2, or just convergence of M_n in L_2 (to whatever).

6c5 Lemma. If a martingale $(M_n)_n$ satisfies $\mathbb{E}\, M_n^2 < \infty$ for all n then random variables[1] $M_{n+1} - M_n$ are mutually orthogonal.

Proof. We have $\mathbb{E}\big((M_{n+1} - M_n)Y\big) = 0$ for all $Y \in L_\infty(\mathcal{F}_n)$ and therefore (by approximation) for all $Y \in L_2(\mathcal{F}_n)$; apply it to $Y = M_{k+1} - M_k$ for $k < n$. $\square$

Thus, $\|M_{n+k} - M_n\|^2 = \|M_{n+k} - M_{n+k-1}\|^2 + \dots + \|M_{n+1} - M_n\|^2$; convergence of $(M_n)_n$ in L_2 is equivalent to convergence of $\sum \|M_{n+1} - M_n\|^2$, that is, to $\sup_n \|M_n\|^2 < \infty$. Here is the conclusion.

6c6 Proposition. A positive[2] martingale bounded in L_2 converges both in L_2 and almost surely.

In order to prove "the former" it remains to prove that $(M_n)_n$ is bounded in L_2. We have

$$
\begin{gathered}
\operatorname{Var} Z_1 = 4p(1-p)\,; \\
\operatorname{Var}\big(Z_{n+1} \,\big|\, Z_n = k \big) = k \operatorname{Var} Z_1 = 4p(1-p)k\,; \\
\operatorname{Var}\big(Z_{n+1} \,\big|\, Z_n \big) = 4p(1-p) Z_n\,; \\
\operatorname{Var}\big(M_{n+1} \,\big|\, M_n \big) = \frac{1}{\big((2p)^{n+1}\big)^2} \operatorname{Var}\big(Z_{n+1} \,\big|\, Z_n \big) = \frac{4p(1-p)(2p)^n}{(2p)^{2n+2}} M_n\,; \\
\operatorname{Var} M_{n+1} = \mathbb{E} \operatorname{Var}\big(M_{n+1} \,\big|\, M_n \big) + \operatorname{Var} \mathbb{E}\big(M_{n+1} \,\big|\, M_n \big)\,; \\
\operatorname{Var} M_{n+1} - \operatorname{Var} M_n = \mathbb{E} \operatorname{Var}\big(M_{n+1} \,\big|\, M_n \big) = \frac{4p(1-p)}{(2p)^{n+2}}\,.
\end{gathered}
$$

But why $\mathbb{P}\big(M_\infty = 0\big) = \mathbb{P}\big(Z_n \to 0\big)$? ("$\ge$" is evident.) Is the event $1 \le Z_n = o\big((2p)^n\big)$ negligible for $p > 0.5$?

First, $\mathbb{P}\big(M_\infty = 0 \,\big|\, \mathcal{F}_n\big) = \big(\mathbb{P}\big(M_\infty = 0\big)\big)^{Z_n}$ (independent subtrees...).

Second, $\mathbb{P}\big(M_\infty = 0 \,\big|\, \mathcal{F}_n\big) \to 1\!\!1_{M_\infty = 0}$ a.s., as we'll see in 7d1.

Thus, on the event $1 \le Z_n = o\big((2p)^n\big)$ (if it is not negligible) we have $\big(\mathbb{P}\big(M_\infty = 0\big)\big)^{Z_n} \to 1$, therefore $\mathbb{P}\big(M_\infty = 0\big) = 1$ in contradiction to $\mathbb{E}\, M_\infty = 1$.

The proof of 6c2 is thus finished (except for one claim postponed to Sect. 7, the "second" above).

(In fact, $\mathbb{P}\big(M_\infty = 0\big) = 1$ if and only if $\mathbb{E}\,(X \ln X) = \infty$...)

[1] So-called martingale differences.

[2] The same holds for non-positive martingales, as we'll see in 7c6.