Markov Processes and Martingales

Károly Simon

Department of Stochastics
Institute of Mathematics
Budapest University of Technology and Economics
www.math.bme.hu/~simonk

File A 2025

Martingales, the definition

Definition 1.1 (Filtered space)

Here we follow the Williams' book. [16] A filtered space is $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $\{\mathcal{F}_n\}_{n=0}^{\infty}$ is a filtration. This means:

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \cdots \subset \mathcal{F}$$

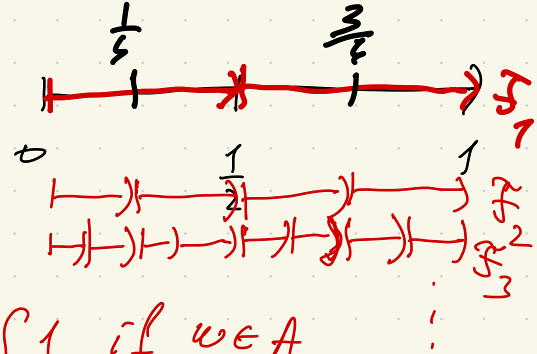
is an increasing sequence of sub σ-algebras of $\mathcal{F}$. Put

$$\mathcal{F}_\infty := \sigma\left(\bigcup_n \mathcal{F}_n\right) \subset \mathcal{F}. \tag{1}$$

$$(\Omega, \mathcal{F}, \mathbb{P}) = \left([0,1), \underset{\uparrow}{\mathcal{R}}, \overset{\nwarrow}{\mathbb{P}}\right)$$

Leb. meas.

Borel σ-alg. on [0,1)



$$\mathcal{F}_\infty = \sigma\left(\bigcup_{n=1}^{\infty} \mathcal{F}_n\right)$$

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{otherwise} \end{cases}$$

$$(g_1 \circ g_2)^{-1}(V) = g_2^{-1} \circ \underbrace{g_1^{-1}(V)}_{\in \mathcal{R}} \in \mathcal{R}$$

If g_1, g_2 are Borel measurable functions then $g_1 \circ g_2$ is also

Let C be the triadic Cantor set:

$\# C = \mathfrak{c}$ (the cardinality of the set of real numbers)

$\mathrm{Leb}(C) = 0 \Rightarrow \forall H \subset C$, H is Leb. measurable & $\mathrm{Leb}(H) = 0$. $\#\{\text{subsets of } C\} = 2^{\mathfrak{c}} > \mathfrak{c}$

$$\underbrace{\overset{\text{cont.}}{g} \circ \underset{\text{Leb. meas.}}{f}}_{\text{is Leb. measurable}}$$

The reason that we use filtration so often is

Theorem 1.2

Given the r.v. $X_1, \ldots, X_n$ and Y on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We define $\mathcal{F} := \sigma(X_1, \ldots, X_n)$. Then

$$(2) \quad Y \in \mathcal{F} \iff \exists g : \mathbb{R}^n \to \mathbb{R}, \text{ Borel s.t. } Y(\omega) = g\left(X_1(\omega), \ldots, X_n(\omega)\right).$$

This means that if $X_1, \ldots, X_n$ are outcomes of an experiment then the value of Y is predictable based on we know the values of $X_1, \ldots, X_n$ iff $Y \in \mathcal{F}$, where $Y \in \mathcal{F}$ means that Y is $\mathcal{F}$-measurable.

When we say simply "process" in this talk, we mean "Discrete time stochastic process".

Definition 1.3 (Adapted process)

We say that the process $M = \{M_n\}_{n=0}^{\infty}$ is adapted to the filtration $\{\mathcal{F}_n\}$ if $M_n \in \mathcal{F}_n$.

Definition 1.4

Let $M = \{M_n\}_{n=0}^{\infty}$ be an adaptive process to the filtration $\{\mathcal{F}_n\}$. We say that X is a martingale if

$\mathcal{F}_{n-1}$

M

- (i) $\mathbb{E}\left[|M_n|\right] < \infty, \forall n$
- (ii) $\mathbb{E}\left[M_n|\mathcal{F}_n\right] = M_{n-1}$ a.s. for $n \geqslant 1$

X is supermartingale if we substitute (ii) with

$$\mathbb{E}\left[M_n|\mathcal{F}_{n-1}\right] \leqslant M_{n-1} \text{ a.s. } n \geqslant 1.$$

Finally, M is a submartingale if we substitute (ii) with

$$\mathbb{E}\left[M_n|\mathcal{F}_{n-1}\right] \geqslant M_{n-1} \text{ a.s. } n \geqslant 1.$$

Remark 1.5

(a) If $M_0 \in L^1$ then the process $M_n - M_0$ is a martingale (respectively submartingale, supermartingale) iff so is $M = \{M_n\}$. (This follows from the definition immediately.)

(b) Assume that $M = \{M_n\}$ is a supermartingale. Then by the tower property for $m < n$ we have

$$\mathbb{E}\left[X_n | \mathcal{F}_m\right] \leqslant X_m. \tag{3}$$

Remark 1.6

In some cases there is another process $X = \{X_n\}$ such that $M_n = f(X_n, n)$ for some function f (like $M_n = X_n^2 - n$). Let $\mathcal{F}_n := \sigma(X_0, \ldots, X_n, M_0)$. Then we say that M is a martingale w.r.t. X if M is a martingale w.r.t. the filtration $\mathcal{F}_n$.

Example 1.7

Let $X_1, X_2, \dots$ be independent L^1 r.v. (this means that $\forall k,\ \mathbb{E}\left[|X_k|\right] < \infty$) with zero mean (that is $\forall k,\ \mathbb{E}\left[X_k\right] = 0$). Let

$S_{n-1} \in \mathcal{F}_{n-1}$

$$S_0 = 0 \text{ and } S_n := X_1 + \cdots + X_n,$$

$$\mathcal{F}_0 = \{\emptyset, \Omega\}, \quad \mathcal{F}_n := \sigma\{X_1, \dots, X_n\}.$$

$$(4) \quad \mathbb{E}\left[S_n|\mathcal{F}_{n-1}\right] = \mathbb{E}\left[S_{n-1}|\mathcal{F}_{n-1}\right] + \mathbb{E}\left[X_n|\mathcal{F}_{n-1}\right]$$

$S_{n-1} \cdot \mathbb{E}[1|\mathcal{F}_{n-1}]$

$$= \underline{S_{n-1}} + \mathbb{E}\left[X_n\right] = S_{n-1}.$$

$$\mathbb{E}[M_{n+1}|\mathcal{F}_n] = \mathbb{E}[X_1\cdots X_n X_{n+1}|\mathcal{F}_n] = X_1\cdots X_n \cdot \underbrace{\mathbb{E}[X_{n+1}|\cancel{\mathcal{F}_n}]}_{1} = M_n.$$

Example 1.8

$X_1\cdots X_n \in \mathcal{F}_n$

(i) Let $X_1, X_2, \ldots$ be independent non-negative r.v. with $\mathbb{E}[X_k] = 1$, $\forall k$. Let $M_1 := 1$, $\mathcal{F}_n$ as in Example 1.7. Let $M_n := X_1 \cdots X_n$. Then $M = \{M_n\}$ is a martingale.

(ii) Given a r.v. $\{X_n\}_{n=1}^{\infty}$ and Y with $\mathbb{E}[|Y|] < \infty$. Then

$$M_n := \mathbb{E}[Y|X_1, \ldots, X_n],$$

is a martingale, called Doob martingale.

Example 1.9 (Exponential Martingale)

Let $Y = \{Y_n\}_{n=1}^{\infty}$ be iid with moment generating function finite at some $\theta \neq 0$: $M(\theta) = \mathbb{E}\left[e^{\theta Y_1}\right] < \infty$. We write $S_n := S_0 + Y_1 \cdots + Y_n$. Then

$$\frac{e^{\theta S_0}\cdot e^{\theta Y_1}\cdot\cdots e^{\theta Y_n}}{\underbrace{M(\theta)\cdots M(\theta)}_{n}} = \frac{e^{\theta S_n}}{M^n(\theta)} = M_n := \frac{\exp(\theta S_n)}{M^n(\theta)}$$

$\boxed{X_i}$

is a martingale w.r.t. Y. Namely, let $X_i := \frac{\exp(\theta Y_i)}{M(\theta)}$. Then $\mathbb{E}\left[Y_i\right] = 1$. So, we apply Example 1.8 (i).

We proved the following convergence theorem (which is also [5, Theorem 5.2.9]) in the course Stochastic Processes. This will be a consequence of some more general convergence theorems that we learn letter in this course.

Theorem 1.10 (Convergence Theorem for non-negative supermartingales)

Let $X_n \geqslant 0$ *be a supermartingale. Then there exists a r.v.* X *s.t.* $X_n \to X$ *a.s. and* $\mathbb{E}\left[X\right] \leqslant \mathbb{E}\left[X_0\right]$.

Functions of MC

Remark 2.1

Given a Markov chain $X = (X_n)$ with transition probability matrix $\mathbf{P} = (p(x,y))_{x,y}$. We are also give a function $f : S \times \mathbb{N} \to \mathbb{R}$ satisfying

$$f(x,n) = \sum_{y \in S} p(x,y) f(y, n+1). \tag{5}$$

$f(X_n, n)$

Then $M_n = \cancel{f(X_n)}$ is a martingale w.r.t. X. (We verified this in the Stochastic Processes course. See [3, Theorem 5.5].)

$$\mathbb{E}[M_{n+1}|\mathcal{F}_n] = \mathbb{E}[f(X_{n+1}, n+1)|\mathcal{F}_n] \overset{\text{by } \{X_n\} \text{ is a MC}}{=} \sum_y p(X_n, y)\cdot f(y, n+1) \overset{\text{by def. of } f}{=} f(X_n, n) \overset{\text{by def of } M_n}{=} M_n$$

Functions of MC (cont.)

Given a Markov chain $X = (X_n)$ with transition probability matrix $\mathbf{P} = (p(x, y))_{x,y}$.

Functions of MC (cont.)

Definition 2.2 (P-harmonic functions)

For an $f : S \to \mathbb{R}$:

(6) $$Pf(x) := \sum_{y \in S} p(x,y) f(y).$$

We say that such an f is harmonic if

(i) $\sum_{y \in S} p(x,y)|f(y)| < \infty, \ \forall x \in S$ and

(ii) $\forall x \in S, \quad \overset{f}{\cancel{h}}(x) = P\overset{f}{\cancel{h}}(x)$

if we replace (ii) with $\forall x, \ f(x) \leqslant Pf(x)$ then f is subharmonic.

f is called superharmonic if $-f$ is subharmonic. It follows from Remark 2.1 that

Theorem 2.3

Let $X = (X_n)$ *be a Markov chain with transition probability matrix* $\mathbf{P} = (p(x,y))_{x,y}$ *and let* h *be a* $\mathbf{P}$*-harmonic function. Then* $h(X_n)$ *is a Martingale w.r.t.* X.

Example 2.4

Let $X_1, X_2, \ldots$ be iid with

$$\mathbb{P}\left(X_i = 1\right) = p \text{ and } \mathbb{P}\left(X_i = -1\right) = 1 - p,$$

$p \in (0, 1)$, $p \neq 0.5$. Let $S_n := X_1 + \cdots + X_n$. Then

$$M_n := \left(\frac{1-p}{p}\right)^{S_n} \tag{7}$$

is a martingale. Namely, $h(x) = ((1 - p)/p)^x$ is harmonic.

Example 2.5 (Simple Symmetric Random Walk)

Let $Y_1, Y_2, \dots$ be iid with

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = 1/2,$$

We write $S_n := S_0 + Y_1 + \cdots + Y_n$. Then $M_n := S_n^2 - n$ is a martingale. Namely, $f(x, n) = x^2 - n$ satisfies (5).

Theorem 2.6

Let h be a subharmonic function for the Markov chain $X = (X_n)$. Then $M_k := h(X_k)$ is a submartingale.

Polya's Urn,

One can find a nice account with more details at `http://www.math.uah.edu/stat/urn/Polya.html` or click here

Given an urn with initially contains: $r > 0$ red and $g > 0$ green balls.

(a) draw a ball from the urn randomly,
(b) observe its color,
(c) return the ball to the urn along with c new balls of the same color.

- If $c = 0$ this is sampling with replacement.
- If $c = -1$ sampling without replacement.

Polya's Urn, (cont.)

From now we assume that $c \geqslant 1$. After the n-th draw and replacement step is completed:

- the number of green balls in the urn is: G_n.
- the number of red balls in the urn is: R_n.
- the fraction of green balls in the urn is X_n.
- Let $Y_n = 1$ if the n-th ball drawn is green. Otherwise $Y_n := 0$.
- Let $\mathcal{F}_n$ be the filtration generated by $Y = (Y_n)$.

Polya's Urn, (cont.)

Claim 1

X_n is a martingale w.r.t. $\mathcal{F}_n$.

Proof Assume that

$$R_n = i \text{ and } G_n = j$$

Then

$$\mathbb{P}\left(X_{n+1} = \frac{j+c}{i+j+c}\right) = \frac{j}{i+j},$$

and

$$\mathbb{P}\left(X_{n+1} = \frac{j}{i+j+c}\right) = \frac{i}{i+j}.$$

Polya's Urn, (cont.)

Hence

$$\mathbb{E}\left[X_{n+1}|\mathcal{F}_n\right] = \frac{j+c}{i+j+c}\cdot\frac{j}{i+j} + \frac{j}{i+j+c}\cdot\frac{i}{i+j} = \frac{j}{i+j} = X_n. \tag{8}$$

□

Corollary 3.1

There exists an X_∞ s.t. $X_n \to X_\infty$ a.s..

Polya's Urn, (cont.)

This is immediate from Theorem 1.10.
In order to find the distribution of $X\infty$ observe that:

- The probability $p_{n,m}$ of getting green on the first m steps and getting red in the next $n-m$ steps is the same as the probability of drawing altogether m green and $n-m$ red balls in any particular redescribed order.
- $$p_{n,m} = \prod_{k=0}^{m-1} \frac{g+kc}{g+r+kc} \cdot \prod_{\ell=0}^{n-m-1} \frac{r+\ell c}{g+r+(m+\ell)c}$$

Polya's Urn, (cont.)

If $c = g = r = 1$ then

$$\mathbb{P}(G_n = \overset{m}{\cancel{2m+1}}) = \binom{n}{m} \frac{m!(n-m)!}{(n+1)!} = \frac{1}{n+1}.$$

That is X_∞ is uniform on $(0,1)$. In the general case X_∞ has density

$$\frac{\Gamma((g+r)/c)}{\Gamma(g/c)\Gamma(r/c)} x^{(g/c)-1}(1-x)^{(r/x)-1}.$$

That is the distribution of X_∞ is $\mathrm{Beta}\left(\frac{g}{c}, \frac{r}{c}\right)$

Games

Let $\{X_n\}_{n=1}^{\infty}$ be an adaptive process to the filtration $\{\mathcal{F}_n\}_{n=1}^{\infty}$.

That is given the probability space $(\Omega, \mathcal{F}, P)$, $\mathcal{F}_0 < \mathcal{F}_1 < \dots < \mathcal{F}_n < \dots \mathcal{F}$ sub-σ-alebras, $X_n \in \mathcal{F}_n$.

Imagine that somebody plays games at times $k = 1, 2, \ldots$. Let $X_k - X_{k-1}$ be the net winnings per unit stake in game n.

In the martingale case

$$\mathbb{E}\left[X_n - X_{n-1} | \mathcal{F}_{n-1}\right] = 0, \quad \text{the game is fair.}$$

In the supermartingale case

$$\mathbb{E}\left[X_n - X_{n-1} | \mathcal{F}_{n-1}\right] \leqslant 0, \quad \text{the game is unfavorable.}$$

Intiutively, we can think of X_n as our wealth after game n if we play with unit stake always.

Games (cont.)

Definition 4.1

Given a process $C = (C_n)$. We say that:

(i) C is previsible or predictable if

$$\forall n \geqslant 1, \quad C_n \in \mathcal{F}_{n-1}.$$

(ii) C is bounded if $\exists K$ such that $\forall n, \forall \omega, \; |C_n(\omega)| < K$.

(iii) C has bounded increments if $\exists K$ s.t. $\forall n \geqslant 1, \forall \omega \in \Omega, \; |C_{n+1}(\omega) - C_n(\omega)| < K$

Games (cont.)

C_n is the player's stake at time n which is decided based upon the history of the game up to time $n-1$. The winning on game n is $C_n(X_n - X_{n-1})$. The total winning after n game is

$$Y_n := \sum_{1\leqslant k\leqslant n} C_k(X_k - X_{k-1}) =: (C \bullet X)_n. \tag{9}$$

By definition:

$$(C \bullet X)_0 = 0.$$

Clearly,

$$Y_n - Y_{n-1} = C_n(X_n - X_{n-1}).$$

Games (cont.)

We say that

$C \bullet X$ is the martingale transform of X by C.

Games (cont.)

Theorem 4.2 (You cannot beat the system)

Given $C = (C_n)_{n=1}^{\infty}$ *satisfying:*

- *(a)* $C_n \geqslant 0$ *for all* n *(otherwise the player would be the Casino),*
- *(b)* C *is previsible (that is* $C_n \in \mathcal{F}_{n-1}$*),*
- *(c)* C *is bounded.*

Then $C \bullet X$ *is a supermartingale (martingale) if* X *is a supermartingale (martingale) respectively.*

Games (cont.)

Proof.

$$\mathbb{E}\left[Y_n - Y_{n-1} | \mathcal{F}_{n-1}\right] \quad = \quad C_n \mathbb{E}\left[X_n - X_{n-1} | \mathcal{F}_{n-1}\right] \quad \underset{\sim}{<} \quad 0. \tag{10}$$

□

Theorem 4.3

Assume that C is a bounded and previsible process and X is a martingale then $C \bullet X$ is a $martingale$ which is null at 0.

Proof

Let $\{X_n\}$ be a supermartingale. As we saw on slide 30:

$$Y_n - Y_{n-1} = C_n(X_n - X_{n-1}).$$

Using that $C_n \in \mathcal{F}_{n-1}$; bounded

$$\mathbb{E}[Y_n - Y_{n-1} \mid \mathcal{F}_{n-1}] = \mathbb{E}[C_n \cdot (X_n - X_{n-1}) \mid \mathcal{F}_{n-1}] =$$

$\{X_n\}$ supermartingale

$$= C_n \cdot \mathbb{E}[X_n - X_{n-1} \mid \mathcal{F}_{n-1}] \le 0$$

So, $\{Y_n\}$ is super-martingale.

Games (cont.)

Theorem 4.4

In the previous two theorems the boundedness can be replaced by $C_n \in L^2, \forall n$ *if* $X_n \in L^2, \forall n$.

The proofs of the one but last theorem is obvious. The proof of the last theorem immediately follows from (f) on slide **??** of file "Some basic facts from probability theory".

1. Martingales, the definitions
2. Martingales that are functions of Markov Chains
3. Polya Urn
4. Games, fair and unfair
5. Stopping Times
6. Stopped martingales

Stopping Times, definitions

Stopping Times, definitions (cont.)

Definition 5.1

A map $T : \Omega \to \{0, 1, \dots, \infty\}$ is called stopping time if

$$\{T \leqslant n\} = \{\omega : T(\omega) \leqslant n\} \in \mathcal{F}_n, \quad n \leqslant \infty. \tag{11}$$

equivalent definition:

$$\{T = n\} = \{\omega : T(\omega) = n\} \in \mathcal{F}_n, \quad n \leqslant \infty. \tag{12}$$

We say that the stopping time T is bounded if $\exists K$ s.t. $T(\omega) < K$ holds for all $\omega \in \Omega$.

Stopping Times, definitions (cont.)

E.g. T is the time when we stop plying the game. We can decided at time n if we stop at that moment based on the history up to time n.

Stopping Times, definitions (cont.)

Example 5.2

Given a process (X_n) which is adapted to the filtration $\{\mathcal{F}_n\}$, further given a Borel set B. Let

$$T := \inf \{n \geqslant 0 : X_n \in B\} .$$

By convention: $\inf \emptyset := \infty$. Then

$$\{T \leqslant n\} = \bigcup_{k \leqslant n} \{T = k\} \in \mathcal{F}_n.$$

Doubling Strategy

(Z_n) i.i.d. $\mathbb{P}(Z_n=-1)=\mathbb{P}(Z_n=1)=\frac{1}{2}$. If $Z_n=1$ the gambler wins & $Z_n=-1$ the gambler loses his stake at game n. Let X_n be the net winnig if per unit stake every game after n game. So, $Z_n=X_n-X_{n-1}$ is the net winning per unit stake at game n. Let $\mathcal{F}_n:=\sigma(X_1,..,X_n)$.

The strategy is: first stake is $C_1=1$. The player doubles stakes after he losses and quits when he wins for the first time. We describe this with the previously introduced terminology: $C_1:=1$, and $C_n=2^{k-1}$ if $k = \min\{n-\ell : C_\ell=1\}$.

	$n=1.$	$2.$	$3.$	$4.$	$5.$
bet C_n	1	2	4	8	16
outcome	L	L	L	L	W
profit	-1	-3	-7	-15	1

The player initial wealth is S.

After $n-1$ losses I lost $2^{n-1}-1$ and the nth stake $=2^{n-1}$. So to play nth game $\boxed{S\geq 2^n-1}$

Let $T:=\inf\{n\geq 1: X_n=1 \text{ or } \underline{S<2^n-1}\}$.

I cannot play nth game

Observe that after k consecutive losses we lost $\sum_{1}^{k} 2^{k-1} = 2^k-1$

Case A $\boxed{S<\infty}$. Then T time

The expected winning: $\mathbb{E}[1\!\!1_{\{T\leq N_0\}} - (2^{N_0}-1)\cdot 1\!\!1_{\{N_0<T\}}]$

$$=\mathbb{P}(N_0\geq T)\cdot 1-(2^{N_0}-1)\cdot\underbrace{\mathbb{P}(N_0<T)}_{1/2^{N_0}} = \left(1-\frac{1}{2^{N_0}}\right)\cdot 1-\frac{2^{N_0}-1}{2^{N_0}}=0.$$

If $S=\infty$ then $X_T=1$ a.s. So the expected winning $=1$.

Stopping Times, definitions (cont.)

Lemma 5.3

Assume that T is a stopping time w.r.t. the filtration $\{\mathcal{F}_n\}$. Let

$$C_n^T := \mathbb{1}_{n \leqslant T}.$$

Then C_n^T is previsible. That is

$$C_n^T \in \mathcal{F}_{n-1}. \tag{13}$$

We play the game with stake one until time T (including time T). The we stop.

Stopping Times, definitions (cont.)

Proof.

$\{C_n^T = 0\} = \{T \leqslant n-1\} \in \mathcal{F}_{n-1}$. □

Stopped martingales

Let T be a stopping time for an $\{\mathcal{F}_n\}$ filtration. For a process $X = (X_n)$ we write X^T for the process stopped at T:

$$X_n^T(\omega) := X_{T(\omega)\wedge n}(\omega),$$

where $a \wedge b := \min\{a, b\}$. $X_{T\wedge n} = \sum_{k=0}^{n-1} 1\!\!1_{\{T=k\}} \cdot X_k + 1\!\!1_{\{T>n-1\}} \cdot X_n$

Assume that Kazér always bets $1\$$ and stops playing at time T. Then Kazmér's stake process is:

$$C_n^{(T)} = \mathbb{1}_{n\leqslant T} \qquad (14)$$

all functions $\left.\begin{matrix} 1\!\!1_{\{T=k\}}, X_k, \ \forall k \leq n-1, \\ 1\!\!1_{\{T>n-1\}}, X_n \end{matrix}\right\} \in \mathcal{F}_n$

So, $X_{T\wedge n} \in \mathcal{F}_n$.

Stopped martingales (cont.)

In Lemma 5.3 we proved that $C^{(T)}$ is previsible (the notion "previsible" was defined on slide # 29).

By (9), Kazmér's winning's process:

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

That is $C^{(T)} \bullet X = X^T - X_0.$

$$\sum_{k=1}^{n} C^{(T)}_k \cdot (X_k - X_{k-1}) = (X_{T\wedge n} - \cancel{X_{T\wedge n-1}}) + (\cancel{X_{T\wedge n-1}} - \cancel{X_{T\wedge n-2}}) + \cancel{\cdot} \cdots$$

$$+ \cancel{X_1} - \underline{\underline{X_0}}$$

So, by Theorems 4.2 and 4.3 we obtain

Stopped martingales (cont.)

Theorem 6.1

Let T be a stopping time

(i)

$$X \text{ supermartingale} \implies X^T \text{ supermartingale.}$$

So, in this case $\forall n, \mathbb{E}\left[X_{T\wedge n}\right] \leqslant \mathbb{E}\left[X_0\right]$

(ii)

$$X \text{ martingale} \implies X^T \text{ martingale.}$$

So, in this case $\forall n, \mathbb{E}\left[X_{T\wedge n}\right] = \mathbb{E}\left[X_0\right]$

Stopped martingales (cont.)

Proof

$= \mathbb{1}_{\{n \leq T\}}$

We define $C_n^{(T)}$ as in (14). Clearly, $C^{(T)} \geqslant 0$ and bounded. As we saw in Lemma 5.3, $C^{(T)}$ is previsible. So, we can apply Theorem 4.2 for

Stopped martingales (cont.)

$$C^T_k = \mathbb{1}_{\{k \le T\}}$$

telescopic sum

(15)
$$\boxed{(C^T \bullet X)_n} = \sum_{k=1}^{n} C^T_k \cdot (X_k - X_{k-1})$$

$$= \begin{cases} X_n - X_0, & \text{on } \{T \geqslant n\}; \\ \sum_{k=1}^{T} (X_k - X_{k-1}) = X_T - X_0, & \text{on } \{T < n\}. \end{cases}$$

$$= \boxed{X_{T \wedge n} - X_0}.$$

$$\sum_{k=1}^{n} (X_k - X_{k-1}) = (X_1 - X_0) + (X_2 - X_1) + (X_3 - X_2) + \dots + (X_n - X_{n-1}) = X_n - X_0.$$

Stopped martingales (cont.)

That is by Theorem 4.2 we get that $X_{T\wedge n} - X_0$ is a supermartingale (martingale) if (X_n) is a supermartingale (martingale) respectively. Which yields the assertion of the theorem. ■

Remark 6.2

It can happen for a martingale X that

(16) $$\mathbb{E}\left[X_{\underset{T}{\cancel{n}}}\right] \neq \mathbb{E}\left[X_0\right].$$

$\mathbb{E}[X_n] = \mathbb{E}[X_0]$ if $\{X_n\}$ is a martingale

Stopped martingales (cont.)

The most popular counter example uses the Simple Symmetric Random Walk (SSRW). First we recall its definition and a few of its most important properties.

Stopped martingales (cont.)

Example 6.3 (Simple Symmetric Random Walk (SSRW))

The Simple Symmetric Random Walk (SSRW) on $\mathbb{Z}$ is $S = (S_n)_{n=0}^{\infty}$, where

$$S_n = X_0 + X_1 + \cdots + X_n, \tag{17}$$

where $X_0 = 0$ and $X_1, X_2, \ldots$ are iid with $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = \frac{1}{2}$.

We have seen that

Stopped martingales (cont.)

Lemma 6.4 (SSRW)

The Simple Symmetric Random Walk on $\mathbb{Z}$ is

(i) *Null recurrent,*

(ii) *martingale.*

The second part follows from Example 1.7. We proved that SSRW is null recurrent in the course Stochastic processes. To give an example where (16) happens:

Stopped martingales (cont.)

Example 6.5

$S = (S_n)$ be the SSRW and let $T := \inf\{n : S_n = 1\}$. Then by Theorem 6.1, $\mathbb{E}[\overset{S}{\cancel{X}}_{T\wedge n}] = \mathbb{E}[\overset{S}{\cancel{X}}_0]$. However,

$$\mathbb{E}[\overset{S}{\cancel{X}}_T] = 1 \neq 0 = \overset{S}{\cancel{X}}_0 = \mathbb{E}[\overset{S}{\cancel{X}}_0].$$

Stopped martingales (cont.)

Question 1

Let X be a martingale and let T be a stopping time. Under which conditions can we say that

$$\mathbb{E}\left[X_T\right] = \mathbb{E}\left[X_0\right]? \tag{18}$$

Theorem 6.6 (Doob's Optional Stopping Theorem)

Let X be a supermartingale and T be a stopping time. If any of the following conditions holds

- *(i) T is bounded.*
- *(ii) X is bounded and $T < \infty$ a.s..*
- *(iii) $\mathbb{E}[T] < \infty$ and X has bounded increments.*

then

(a) $X_T \in L^1$ *and* $\mathbb{E}(X_T) \leqslant \mathbb{E}[X_0]$.

(b) *If* X *is a* $martingale$ *then* $\mathbb{E}(X_T) = \mathbb{E}[X_0]$.

Proof.

By Thm: 6.1 $\forall n, \quad X_{T\wedge n} \in L^1$ and $\mathbb{E}\left[X_{T\wedge n} - X_0\right] \leqslant 0.$

If (i) holds then $\exists N$ s.t. $T \leqslant N$. Then for $n = N$, we have $X_{T\wedge n} = X_T$. Hence (a) follows.

If (ii) holds then $\lim\limits_{n\to\infty} X_{T\wedge n} = X_T$. So, by Dominated Convergence Theorem: $\lim\limits_{n\to\infty} \mathbb{E}\left[X_{n\wedge T}\right] = \mathbb{E}\left[X_T\right]$. On the other hand, by Theorem 6.1, $\mathbb{E}\left[X_{T\wedge n}\right] \leqslant \mathbb{E}\left[X_0\right]$.

If (iii) holds The answer comes from Dom. Conv. Thm.
$|X_{T\wedge n} - X_0| = \left|\sum\limits_{k=1}^{T\wedge n}(X_k - X_{k-1})\right| \leqslant KT < \infty$. If X is a martingale, apply everything above also for $-X$. □

Corollary 6.7

Assume that

(a) $M = (M_n)$ *is a martingale.*

(b) $\exists K_1$ s.t. $\forall n$, $|M_n - M_{n-1}| < K_1$,

(c) $C = \{C_n\}$ *is a previsible process with* $|C_n(\omega)| < K_2$, $\forall \omega, \forall n$.

(d) T *is a stopping time with* $\mathbb{E}[T] < \infty$.

Then

$$\mathbb{E}[(C \bullet M)_T] = 0. \tag{19}$$

Proof.

Put together Theorem 4.3 and Theorem 6.6. ☐

A corollary of the Optional Stopping Theorem is:

Theorem 6.8

Assume that

(i) $\xcancel{M} = (\xcancel{M}_n)$ *is a non-negative supermartingale,*

(ii) T *is a stopping time s.t.* $T < \infty$ *a.s..*

Then $\mathbb{E}[X_T] \leqslant \mathbb{E}[X_0]$.

Proof

We know that $\lim_{n\to\infty} X_{T\wedge n} = X_T$ a.s. and $X_{T\wedge n} \geqslant 0$. So we can apply Fatou Lemma :

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

Proof cont.

On the other hand, by Theorem 6.1 the left hand side is smaller than or equal to $\mathbb{E}\left[X_0\right]$

.

Awaiting for the (almost) inevitable

In order to apply the previous theorems we need a machinery to check if $\mathbb{E}\left[T < \infty\right]$ a.s. holds.

Theorem 6.9

Assume that $\exists N \in \mathbb{N}, \varepsilon > 0$ *s.t.* $\forall n \in \mathbb{N}$,

$$\mathbb{P}\left(T \leqslant n + N | \mathcal{F}_n\right) > \varepsilon, \quad \textit{a.s.} \tag{20}$$

then

$$\mathbb{E}\left[T\right] < \infty.$$

Awaiting for the (almost) inevitable (cont.)

Proof.

We apply (20) for $n = (k-1)N$. Then the assertion follows by mathematical induction from Homework **??**. □

ABRACADABRA

The following exercise is named as "Tricky exercise" in Williams' book [16, p.45].

Problem 6.10 (Monkey at the typewriter)

Assume that a monkey types on a typewriter. He types only capital letters and he chooses equally likely any of the 26 letters of the English alphabet independently of everything. What is the expected number of letters he needs to type until the word "ABRACADABRA" appears in his typing for the first time?

The same problem formulated in a more formal way:

ABRACADABRA (cont.)

Problem 6.11 (Monkey at the typewriter)

Let $X_1, X_2, \dots$ *be i.i.d. r.v. taking values from the set* $\text{Alphabet} := \{A, B, \dots, Z\}$ *of cardinality* 26*. We assume that the distribution of* X_k *is uniform. Let* T *be*

$$T := \min \{n + 10 : (X_n, X_{n+1}, \dots, X_{n+10}) = (A, B, R, A, C, A, D, A, B, R, A)\} \tag{21}$$

Find $\mathbb{E}[T] = ?$

ABRACADABRA (cont.)

We associate a players in a Casino to the monkey:

Example 6.12 (Players associated to the monkey)

Imagine that for every $\ell = 1, 2, \dots,$ on the ℓ-th day a new gambler arrives in a Casino. He bets:

$1\$$ on the event: " $X_\ell = A$" .

If he loses he leaves. If he wins he receives $26\$$. Then he bets his

$26\$$ on the event: " $X_{\ell+1} = B$" .

If he loses he leaves. If he wins then he receives $26^2\$$ and then he bets all of his

$26^2\$$ on the event: "$\ell + 2$-th letter will be R"

and so on until he loses or gets ABRACADABRA.

Now for every j we define a previsible process

$C^{(j)} = \left\{C_n^{(j)}\right\}$. Namely, let $C_n^{(j)}$ be the bet of gambler j on day n:

$$C_{j+k}^{(j)} := \begin{cases} 0, & \text{if } k < 0; \\ 1, & \text{if } k = 0; \\ 26^k, & \text{if } X_j, \ldots, X_{j+k-1} \text{ were correct; } 1 \leqslant k \leqslant 11 \\ 0, & \text{otherwise,} \end{cases}$$

where $X_j, \ldots, X_{j+k-1}$ correct means that they are the first k letters of ABRACADABRA. For every j, the value of $C_n^{(j)}$ depends only on $X_1, \ldots, X_{n-1}$. So, for every j the process $C^{(j)} = \left\{C_n^{(j)}\right\}$ is previsible.

The definition of $M_k^{(j)}$ Fix a $j \geqslant 1$. The net winning of player j after the HIS first day (day j of the game) is either

- $-1\$$ if monkey did not type A on day j of the game,
- $(26-1)\$$ if monkey typed letter A on the j-th of the game.

Similarly, k days after that player j entered the game (this is the $j+k-1$-th day of the game) the net winning of player j is either $(26^k-1)\$$ or $-1\$$. This net winning comes from the amount the Casino paid to the player by the end of his k-th day in the game (which is the $k+j-1$-th day of the game) minus $1\$$ (which is the player's initial investment).

We denote this net winning of player j after HIS k-th day in the game by $M_k^{(j)}$.

Remember that we have fixed a j. For this j we define $\mathcal{F}_k^{(j)} := \sigma(X_j, \ldots, X_{j+k-1})$.

Claim 2

For every j, $M_k^{(j)}$ is a martingale w.r.t. $\mathcal{F}_k^{(j)}$ with

(22) $$\mathbb{E}\left[M_k^{(j)}\right] = 0.$$

Proof of the Claim. Then $M_k^{(j)} \in \mathcal{F}_k^{(j)}$ and $-1 \leqslant M_k^{(j)} \leqslant 26^k$. That is $M_k^{(j)}$ is bounded, in particular $M_k^{(j)} \in L^1$. If $M_k^{(j)} \neq -1$ then $M_k^{(j)} = 26^k - 1$. Conditioned on this:

$$M_{k+1}^{(j)} = \begin{cases} 26^{k+1} - 1, & \text{with probability } 1/26; \\ -1, & \text{with probability } 25/26. \end{cases}$$

Then

$$\mathbb{E}\left[M_{k+1}^{(j)} | M_k^{(j)} \cancel{\neq} -1\right] = 26^{k+1} \cdot 1/26 - 1 = M_k^{(j)}. \tag{23}$$

$$\frac{1}{26}\cdot\left(26^{k+1}-1\right)+\frac{25}{26}(-1)=\frac{1}{26}\cdot\left(26^{k+1}\overbrace{-1-25}^{-26}\right)=\overbrace{\frac{26}{26}\left(26^{k}-1\right)}^{M_k^{(j)}}$$

On the other hand, if $M_k^{(j)} = -1$ then also $M_{k+1}^{(j)} = -1$. So

(24) $$\mathbb{E}\left[M_{k+1}^{(j)} | M_k^{(j)} \neq 26^k\right] = -1 = M_k^{(j)}$$

Putting these together we obtain that

(25) $$\mathbb{E}\left[M_{k+1}^{(j)} | \mathcal{F}_k\right] = M_k^{(j)}.$$

Hence, $\mathbb{E}\left[M_k^{(j)}\right] = \mathbb{E}\left[M_1^{(j)}\right] = 0$. The last equality follows from an immediate computation. ■

Now we apply Doob's optional stopping theorem for the stopping time T defined in (21) and for a martingale $S = (S_n)$ to be defined below.

The definition of $S = (S_n)$ Let S_n be the cumulative net winning (may be negative) of all gamblers together up to (and including) time n:

$$S_n := \sum_{j=1}^{n} M_n^{(j)}. \tag{26}$$

By (22) we have

$$\forall n, \quad \mathbb{E}[S_n] = 0. \tag{27}$$

Then S_n is the finite sum of martingales, so S_n is a martingale itself w.r.t. the filtration: $\{\sigma(X_1, \dots, X_n)\}_n$. Actually we verify in the following two Claims that both parts of condition (iii) of Theorem 6.6 hold.

Claim 3

$\mathbb{E}[T] < \infty$.

Proof of the Claim The Claim follows from Theorem 6.9 with the following substitutions:

$$N = 11, \quad \varepsilon = (1/26)^{11},$$

Namely, whatever happens now, the probability is at least $(1/26)^{11}$ that in the next 11 steps the monkey gets ABRACADABRA. ■

Claim 4

There exists a finite J such that $|S_n - S_{n-1}| < J$.

Proof of the Claim

By definition, $|S_n - S_{n-1}|$ is less than the maximum amount J that the Casino can possibly pay to all of the players together on any particular single day. We prove below that J is finite. This implies that $S = (S_n)$ has bounded (by J) increments.

By Claims 3 and 4 both parts of condition (iii) of Theorem 6.6 hold. Hence by this Theorem and ~~(22)~~ (27) we get

$$\mathbb{E}\left[S_T\right] = \mathbb{E}\left[S_1\right] = 0. \tag{28}$$

The computation of J The worst day for the Casiono, that is the day when the total amount that the Casino pays to all the players together is at its maximum is clearly the last day of the game, that is day T.

So, J is the amount the Casino pays on the day when the monkey first completed the typing of the word "ABRACADABRA". This is by definition day T. To compute J note that there are exactly three players who get payment on day T. Namely,

- The one who arrived on day T. (He had to bet for A). He gets $26\$$ from the Casino.
- The one who arrived on day $T-3$ has made 4 successful bets. So, he gets $26^4\$$ from the Casino on day T.
- The player who arrived on day $T-10$ gets $26^{11}\$$ from the Casino.

So, the total amount that the Casino pays on day T is

$$\boxed{J := 26 + 26^4 + 26^{11}}.$$

Observe that whatever the Casino paid to the players on any day $n < T$ they immediately bet it on day $n+1 \leqslant T$. So, the Casino got it back. In this way, the total amount that the Casino ever paid to the players is just J. By definitions, this means that

$$\boxed{S_T = J - T}.$$

From this and from (28) we obtain that

$$\mathbb{E}\left[T\right] = J = 26 + 26^4 + 26^{11}.$$

This solves the Monkey at typewriter problem. ■

[1] P. BILLINGSLEY
Convergence of probability measures
Wiley, 1968

[2] B. DRIVER
Analysis tools with examples
Lecturenotes, 2012. Click here.

[3] R. DURRETT
Essentials of Stochastic Processes, Second edition
Springer, 2012. Click here

[4] R. DURRETT
Probability: Theory with examples, 4th edition
Cambridge University Press, 2010.

[5] R. DURRETT
Probability: Theory and Examples
Click here

[6] D.H. FREMLIN
Measure Theory Volume I
Click here

[7] D.H. FREMLIN
Measure Theory Volume II
Click here

[8] O. VAN GAANS
Probability measures on metric spaces
Click here

[9] S. Karlin, H.M. Taylor
A first course in stochastic processes
Academic Press, New York, 1975

[10] S. Karlin, H.M. Taylor
Sztochasztikus Folyamatok
Gondolat, Budapest, 1985

[11] S. Karlin, H.M. Taylor
A second course in stochastic processes
, Academic Press, 1981

[12] G. Lawler
Intoduction to Stochastic Processes
Chapmann & Hall 1995.

[13] D.A. Levin, Y. Peres,E.L. Wilmer
Markov chains and mixing times
American Mathematical Society, 2009.

[14] P. Mattila Geometry of sets and measure in Euclidean spaces. Cambridge, 1995.

[15] S. Ross
A First Course in Probability, 6th ed.
Prentice Hall, 2002

[16] D. Williams
Probability with Martingales
Cambridge 2005