Balint Toth: Martingale Theory

Filtrations, Martingales,

$(\Omega, \mathcal{F}, \mathbb{P})$ probability space

discrete time stochastic process $=$
sequence of random variables
$X_0, X_1, X_2, \ldots$ jointly defined on $(\Omega, \mathcal{F}, \mathbb{P})$

time $n = 0, 1, 2, \ldots$

Definitions $\boxed{\text{Filtration}}$ of σ-algebras on $(\Omega, \mathcal{F}, \mathbb{P})$:

$\mathcal{F}_n : n \geq 0$

$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_n \subseteq \cdots \subseteq \mathcal{F}$

increasing sequence of (sub) σ-algebras

Meaning: $\mathcal{F}_n$ contains all information available at time n

$$\mathcal{F}_\infty := \sigma\Big(\bigcup_{n<\infty} \mathcal{F}_n\Big) \subseteq \mathcal{F} \quad \left(\begin{smallmatrix}\text{not necessarily}\\ = \mathcal{F}\end{smallmatrix}\right)$$

Natural (own) filtration of a stoch. process $Y_0, Y_1, Y_2, \dots$

$$\mathcal{F}_k := \sigma(Y_0, Y_1, \dots, Y_k)$$

The process $(X_k)_{k\geq 0}$ is **adapted** to the filtration $(\mathcal{F}_k)_{k\geq 0}$ if

$$(\forall k) \quad X_k \text{ is } \mathcal{F}_k\text{-measurable}$$

<u>Example</u> $\mathcal{F}_k = \sigma(Y_0, Y_1, \dots, Y_k)$, $k \geq 0$

$(\forall n)$ $\varphi_n : \mathbb{R}^{n+1} \to \mathbb{R}$ measurable

$$X_n := \varphi_n(Y_0, Y_1, \dots, Y_n), \quad n \geq 0$$

then $(X_n)_{n\geq 0}$ is adapted to the filtration $(\mathcal{F}_n)_{n\geq 0}$.

$$(\underset{\substack{\uparrow\\ \text{sample space}}}{\Omega}, \underset{\substack{\uparrow\\ \sigma\text{-alg}}}{\mathcal{F}}, \underbrace{(\mathcal{F}_n)_{n\geq 0}}_{\text{filtration}}, \underset{\substack{\uparrow\\ \text{probab. measure}}}{\mathbb{P}}) = \text{filtered probability space}$$

<u>Definition</u>: $\boxed{\text{martingale}}$, $\boxed{\text{supermartingale}}$, $\boxed{\text{submartingale}}$

$(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$ filtered probab sp.

A stochastic process $(X_n)_{n\geq 0}$ is a martingale / supermartingale / submartingale on $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$ if the following hold:

(i) $(X_n)_{n\geq 0}$ is adapted to the filtration $(\mathcal{F}_n)_{n\geq 0}$

(ii) $(\forall n\geq 0)$ $\mathbb{E}(|X_n|) < \infty$

(iii) $(\forall n\geq 0)$:

$\mathbb{E}(X_{n+1} \mid \mathcal{F}_n) = X_n$ a.s. martingale

$\mathbb{E}(X_{n+1} \mid \mathcal{F}_n) \leq X_n$ a.s. supermartingale

$\mathbb{E}(X_{n+1} \mid \mathcal{F}_n) \geq X_n$ a.s. submartingale

$(X_n)_{n\geq 0}$ supermart $\Leftrightarrow$ $(-X_n)_{n\geq 0}$ submart

$(X_n)_{n\geq 0}$ martingale $\Leftrightarrow$ $(X_n)_{n\geq 0}$ supermart. and submart

$V_n := X_n - X_{n-1}$, $n\geq 1$:

the martingale / submart. / supermart difference process

Proposition If $(X_n)_{n\geq 0}$ is mart/supermart/submart, then

$$\forall\, 0 \leq m \leq n < \infty: \quad \mathbb{E}(X_n \mid \mathcal{F}_m) = X_m .$$

Proof Straightforward. Apply tower and induction.

Examples:

Ex 1 $(\Omega, \mathcal{F}, \mathbb{P})$,

$\xi_1, \xi_2, \xi_3 \ldots$ independent random variables

$$\mathbb{E}(|\xi_j|) < \infty, \quad \mathbb{E}(\xi_j) = 0$$

$$\mathcal{F}_0 := \{\emptyset, \Omega\}, \quad \mathcal{F}_n := \sigma(\xi_1, \ldots, \xi_n), \quad n \geq 1$$

$X_n := \sum_{k=1}^{n} \xi_k$ is a martingale

(somewhat trivial example...)

indeed (i) adoptedness ✓

(ii) $\mathbb{E}(|X_k|) \le \mathbb{E}\left(\sum_{j=1}^{k} |\xi_j|\right) =$

$= \sum_{j=1}^{k} \mathbb{E}(|\xi_j|) < \infty$

(iii) $\mathbb{E}(X_{n+1} | \mathcal{F}_n) = \mathbb{E}(X_n + \xi_{n+1} | \mathcal{F}_n) =$

$\dots = X_n + \underbrace{\mathbb{E}(\xi_{n+1} | \mathcal{F}_n)}_{=0} = X_n$

Ex

(2) $(\Omega, \mathcal{F}, \mathbb{P})$

$(\xi_j)_{j=1}^{\infty}$ independent random variables

$\mathbb{E}(|\xi_j|) < \infty, \quad \mathbb{E}(\xi_j) = 1$

$\mathcal{F}_0 := \{\emptyset, \Omega\}, \quad \mathcal{F}_n := \sigma(\xi_1, \dots, \xi_n), \quad n \ge 1$

$$X_0 := 1, \quad X_n := \prod_{j=1}^{n} \xi_j$$

$(X_n)_{n \geq 0}$ is $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$ – martingale

(i) adaptedness ✓

(ii) $$\mathbb{E}(|X_n|) = \mathbb{E}\left(\prod_{j=1}^{n} |\xi_j|\right)$$

$$\overset{\text{indep}}{=} \prod_{j=1}^{n} \mathbb{E}(|\xi_j|) < \infty$$

(iii) $$\mathbb{E}(X_{n+1} | \mathcal{F}_n) = \mathbb{E}(X_n \cdot \xi_{n+1} | \mathcal{F}_n)$$

$$= X_n \cdot \underbrace{\mathbb{E}(\xi_{n+1} | \mathcal{F}_n)}_{=1} = X_n$$

Ex 3 $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$ filtered

$$Z \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P}) \qquad (\mathbb{E}(|Z|) < \infty)$$

Let $X_n := \mathbb{E}(Z \mid \mathcal{F}_n)$

"reveal more-and-more information about Z"

Then $(X_n)_{n\geq 0}$ is an $(\mathcal{F}_n)_{n\geq 0}$-martingale

(i) adapted by def.

(ii)

$$\mathbb{E}(|X_n|) = \mathbb{E}\left(\left|\mathbb{E}(Z \mid \mathcal{F}_n)\right|\right) \underset{}{\overset{\text{Jensen}}{\leq}} \mathbb{E}\left(\mathbb{E}(|Z| \mid \mathcal{F}_n)\right) \underset{\text{tower}}{=} \mathbb{E}(|Z|) < \infty$$

(iii)

$$\mathbb{E}(X_{n+1} \mid \mathcal{F}_n) = \mathbb{E}\left(\mathbb{E}(Z \mid \mathcal{F}_{n+1}) \mid \mathcal{F}_n\right) \underset{\text{tower}}{=} \mathbb{E}(Z \mid \mathcal{F}_n)$$

We'll see later that in this example

$$X_n \xrightarrow[\text{as } n\to\infty]{} X_\infty := \mathbb{E}(Z \mid \mathcal{F}_\infty)$$

Ex

4 Let η_n be a Markov chain on the finite state space $S = \{1, 2, \dots, N\}$

$$P_{kl} = \mathbb{P}(\eta_{n+1} = l \mid \eta_n = k), \quad k, l \in S$$

the stochastic matrix of transition probabilities

$(\Omega, \mathcal{F}, \mathbb{P})$ = probab space of the process $(\eta_n)_{n\geq 0}$

$$\mathcal{F}_n := \sigma(\eta_1, \dots, \eta_n)$$

$f: S \to \mathbb{R}$ a function on the state space

$$X_0 := 0, \quad X_n := \sum_{m=1}^{n} \left(f(\eta_m) - Pf(\eta_m) \right)$$

$$Pf(k) = \sum_l P_{kl} f(l)$$

(i) adapted ✓

(ii) $\mathbb{E}(|X_n|) < \infty$: all is finite

(iii) $X_n = \sum_{m=1}^{n} V_m \ ; \ V_m = f(\eta_m) - Pf(\eta_{m-1})$

$$\mathbb{E}(V_{n+1} \mid \mathcal{F}_n) = \mathbb{E}\left(f(\eta_{n+1}) - Pf(\eta_n) \mid \mathcal{F}_n\right)$$

$$= \underbrace{\mathbb{E}\left(f(\eta_{n+1}) \mid \mathcal{F}_n\right)}_{= Pf(\eta_n)} - Pf(\eta_n) = 0 \ ✓$$

Note: by rearrangement

$$X_n = \sum_{m=0}^{n-1} (1-P) f(\eta_m) + f(\eta_n) - f(\eta_0)$$

Dynkin-martingale

Ex 5 $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$

$(X_n)_{n\geq 0}$ be an $(\mathcal{F}_n)_{n\geq 0}$ martingale

$\varphi: \mathbb{R} \to \mathbb{R}$ convex function. (measurable)

Assume that $(\forall n)$ $\mathbb{E}(|\varphi(X_n)|) < \infty$.

Then $Y_n := \varphi(X_n)$ is $(\mathcal{F}_n)_{n\geq 0}$ - submart.

(i) ✓ (ii) ✓ (by assumption)

$$\mathbb{E}(Y_{n+1} | \mathcal{F}_n) = \mathbb{E}(\varphi(X_{n+1}) | \mathcal{F}_n) \geq$$

$$\geq \varphi\Big(\underbrace{\mathbb{E}(X_{n+1} | \mathcal{F}_n)}_{= X_n}\Big) = \varphi(X_n) = Y_n$$

Jensen

Ex 6 Pólya Urn

In an urn we have RED and BLUE balls

- Start with $P_0 = k$, $P = l$ $\quad (k, l \geq 1)$

- at discrete times, $n = 1, 2, \ldots$
  - draw a randomly (uniformly) chosen ball from the urn
  - put it back along with another one with the same colour

(B_n, R_n) Markov chain on $\mathbb{Z}_+ \times \mathbb{Z}_+$

$$\mathbb{E}\left(f(B_{n+1}, R_{n+1}) \mid \mathcal{F}_n\right) =$$

$$f(B_n + 1, R_n) \frac{B_n}{B_n + R_n} + f(B_n, R_{n+1}) \frac{R_n}{B_n + R_n}$$

$$X_n := \frac{B_n}{B_n + R_n} =$$ the proportion of blue balls in the urn at stage n

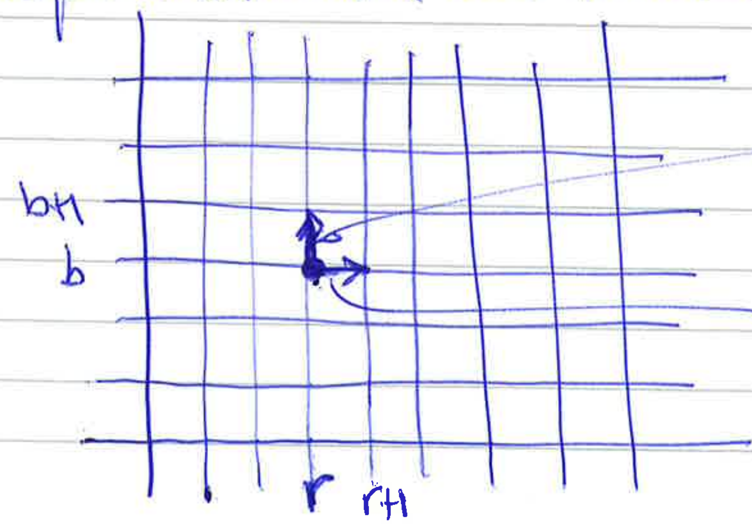
$(X_n)_{n \geq 1}$ is a martingale

$$\mathbb{E}(X_{n+1} | \mathcal{F}_n) = \mathbb{E}\left(\frac{B_{n+1}}{B_{n+1} + R_{n+1}} \Big| \mathcal{F}_n \right) =$$

$$\frac{B_n + 1}{B_n + R_n + 1} \cdot \frac{B_n}{B_n + R_n} + \frac{B_n}{B_n + R_n + 1} \cdot \frac{R_n}{B_n + R_n}$$

$$= \cdots = \frac{B_n}{B_n + R_n} = X_n$$

Representation as Random Walk on $\mathbb{Z}_+ \times \mathbb{Z}_+$



probab $\frac{b}{r+b}$

probab $\frac{r}{r+b}$

HW: Write computer simulation

Polya Urn Continued

$B_0 = k, \; R_0 = l$

$\xi_m := B_m - B_{m-1} \in \{0,1\}$

$m \geq 1$

$1 - \xi_m = R_m - R_{m-1} \in \{0,1\}$

$$S_n := \sum_{m=1}^{n} \xi_m = B_n - B_0 \quad , n \geq 0$$

Compute the probability of a particular sequence of drawings

$\xi_m = x_m \quad 1 \leq m \leq n$, where

$x_m \in \{0,1\}$ are given

Notation: $s_n := \sum_{m=1}^{n} x_m$

$b_n := b_0 + s_n = k + s_n$

$r_n := r_0 + (n - s_n) = l + n - s_n$

$$\mathbb{P}\left(\xi_m = x_m, \quad m = 1, 2, \ldots, n\right) = \cdots =$$

$$\frac{\prod_{j=0}^{s_n-1} b_j \cdot \prod_{j=0}^{n-s_n-1} r_j}{\prod_{j=0}^{n-1} (b_j + r_j)} =$$

$$\frac{\frac{(b_n-1)!}{(b_0-1)!} \cdot \frac{(r_n-1)!}{(r_0-1)!}}{\frac{(b_n+r_n-1)!}{(b_0+r_0-1)!}} =$$

$$\frac{\Gamma(b_n)}{\Gamma(b_0)} \cdot \frac{\Gamma(r_n)}{\Gamma(r_0)} \cdot \frac{\Gamma(b_0+r_0)}{\Gamma(b_n+r_n)}$$

The probability depends only on the number of blue/red balls drawn up to n, and not on their order

$$\mathbb{P}(brrbrb) = \mathbb{P}(rrrbbb) \text{ etc}$$

$\boxed{\text{exchangeability}}$

Now a different problem:

I have a biased coin

$$\mathbb{P}(\text{Head}) = \Theta = 1 - \mathbb{P}(\text{Tail})$$

The bias $\Theta \in [0,1]$ itself is random with distribution

$$\Theta \sim \text{BETA}(k,l)$$

$$\mathbb{P}(\Theta \in (\theta, \theta + d\theta)) = \frac{\Gamma(k+l)}{\Gamma(k)\Gamma(l)} \theta^{k-1} (1-\theta)^{l-1} d\theta$$

Toss this coin

$$\xi_m = \begin{cases} 1 & \text{if Head} \\ 0 & \text{if Tail} \end{cases}$$

$$P\left(\xi_m = x_m,\ m = 1, 2, \dots, n\right) =$$

$$E\left(P\left(\xi_m = x_m;\ m = 1, 2, \dots, n \,\middle|\, \sigma(\Theta)\right)\right) =$$

$$\int_0^1 \theta^{s_n} (1-\theta)^{n-s_n} \frac{\Gamma(k+l)}{\Gamma(k)\Gamma(l)} \theta^{k-1} (1-\theta)^{l-1} \, d\theta =$$

$$\frac{\Gamma(k+l)}{\Gamma(\underbrace{k}_{b_0})\Gamma(\underbrace{l}_{r_0})} \int_0^1 \theta^{\overbrace{k+s_n-1}^{b_n}} (1-\theta)^{\overbrace{l+n-s_n-1}^{r_n}} \, d\theta =$$

$$\frac{\Gamma(b_0+r_0)}{\Gamma(b_0)\Gamma(r_0)} \cdot \frac{\Gamma(b_n)\Gamma(r_n)}{\Gamma(b_n+r_n)}$$

SAME!

Now let $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$ (Θ not known!) and compute

$\boxed{\text{Recall example } 3}$

$$M_n := \mathbb{E}(\Theta \mid \mathcal{F}_n)$$

Note that $\mathcal{F}_n$ is atomic (discrete) σ-algebra. So we can compute "naïvely". Use Bayes' Theorem:

$$\mathbb{E}(\Theta \mid \xi_m = x_m,\ 1 \leq m \leq n) =$$

$$\frac{\mathbb{E}(\Theta \cdot \mathbb{P}(\xi_m = x_m,\ 1 \leq m \leq n \mid \sigma(\Theta)))}{\mathbb{E}(\mathbb{P}(\xi_m = x_m,\ 1 \leq m \leq n \mid \sigma(\Theta)))} =$$

$$\dots = \frac{\Gamma(b_n + 1) \cdot \Gamma(b_n + r_n)}{\Gamma(b_n) \cdot \Gamma(b_n + r_n + 1)} = \frac{b_n}{b_n + r_n}$$

So,

$$M_n := \mathbb{E}(\Theta \mid \mathcal{F}_n) = \frac{B_n}{B_n + R_n}$$

The same martingale as in Pólya Urn!

Ex 7 Galton-Watson Branching process

$\xi_{n,k}$: $n \geq 1$, $k \geq 1$
i.i.d. random variables

$\xi_{n,k} \in \mathbb{N} = \{0, 1, 2, \dots\}$

$E(\xi_{n,k}) = m < \infty$
$Var(\xi_{n,k}) = \sigma^2 < \infty$ (assumed)

$Z_0 \in \mathbb{N}$ given, assume $Z_0 > 0$

$$Z_{n+1} = \sum_{k=1}^{Z_n} \xi_{n+1,k}$$

Interpretation:

Z_n = population size in n-th generation

$\xi_{n+1,k}$ = # of children of k-th member of n-th generation

$$\mathcal{F}_0 = \{\emptyset, \Omega\}, \quad \mathcal{F}_n := \sigma\left(\xi_{m,k} : \begin{array}{l} 1 \le m \le n \\ k \ge 1 \end{array}\right)$$

Then

$$X_n := m^{-n} \cdot Z_n, \quad n \ge 0$$

is an $(\mathcal{F}_n)_{n \ge 0}$ - martingale

$$\mathbb{E}(X_{n+1} | \mathcal{F}_n) = m^{-(n+1)} \mathbb{E}(Z_{n+1} | \mathcal{F}_n)$$

$$= m^{-(n+1)} \mathbb{E}\left(\sum_{k=1}^{Z_n} \xi_{n+1,k}\right) = m^{-(n+1)} \cdot Z_n \cdot m =$$

$$= m^{-n} Z_n = X_n \quad \checkmark$$

8 Gambling = discrete time stochastic integration

Math model of a casino / stock market

$(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$ filtered probab. sp.

$(X_n)_{n\geq 0}$ adapted process, $X_0 = 0$

$V_n := X_n - X_{n-1}$ the difference process

$V_n =$ the amount earned or lost on unit bet in the n-th round

The bet is

fair if $\mathbb{E}(V_n \mid \mathcal{F}_{n-1}) = 0$

unfavourable if $\mathbb{E}(V_n \mid \mathcal{F}_{n-1}) \leq 0$

favourable if $\mathbb{E}(V_n \mid \mathcal{F}_{n-1}) \geq 0$

How do we gamble? In the n-th round knowing $(V_1, V_2, \dots V_{n-1})$ (but not V_n!) we decide [according to some strategy/reasoning] that in the n-th round we put $C_n \ (\geq 0)$ units on bet and we win/loose $C_n \cdot V_n$.

Definition $\boxed{\text{predictible process}}$

The process $(C_n)_{n \geq 1}$ is predictible if $(\forall n)$ C_n is $\mathcal{F}_{n-1}$ - measurable

The total earnings (or loss) in the first n rounds

$$Y_n = \sum_{m=1}^{n} C_m V_m = \sum_{m=1}^{n} C_m \underbrace{(X_m - X_{m-1})}_{\text{"}dX_m\text{"}}$$

a discrete baby-version of stochastic integration

<u>Proposition</u>: $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$ –

$(X_n)_{n\geq 0}$ adapted; $(C_n)_{n\geq 1}$ predictible bounded

$$Y_0 = 0; \quad Y_n := \sum_{m=1}^{n} C_m (X_m - X_{m-1}) \quad \boxed{C_m \geq 0}$$

Then, if $(X_m)_{m\geq 0}$ is mart./supermart/submart

then so is $(Y_m)_{m\geq 0}$.

<u>Proof</u> Straightforward.