# Lecture 21: Martingales

Martingales are models for a system of wagers coming from a sequence of fair games.

submartingales model advantageous games

supermartingales model disadvantageous games

Another view point is that sub & super martingales model the stochastic equivalent to monotonic sequences of numbers. Indeed we will prove that under "bdd" type conditions sub & super martingales are guarenteed to converge.

Indeed, a basic theme in the next lecture is that "Martingales like to converge" so "looking for a martingale" is a technique for establishing convergence.

**Assumption:** For the rest of this lecture fix some probability space $(\Omega, \mathcal{F}, P)$.

**Definition:** A **filtration** is an increasing sequence of sub $\sigma$-fields

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

**Definition:** A sequence of r.v.s $X_1, X_2, \ldots$ defined on $(\Omega, \mathcal{F}, P)$ is **adapted** to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ if $X_n$ is $\mathcal{F}_n$-measurable for each $n\in\mathbb{N}$.



**Definition:**

Suppose $(X_n)_{n\in\mathbb{N}}$ is adapted to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ & $X_n \in L_1(\Omega, \mathcal{F}, P)$ for all $n\in\mathbb{N}$. Then

- $(X_n)_{n\in\mathbb{N}}$ is a **martingale** if
$$E(X_{n+1}|\mathcal{F}_n) \overset{P\text{-a.e.}}{=} X_n \quad \forall n\in\mathbb{N}$$
- $(X_n)_{n\in\mathbb{N}}$ is a **submartingale** if
$$E(X_{n+1}|\mathcal{F}_n) \overset{P\text{-a.e.}}{\geq} X_n \quad \forall n\in\mathbb{N}$$
- $(X_n)_{n\in\mathbb{N}}$ is a **supermartingale** if
$$\underbrace{E(X_{n+1}|\mathcal{F}_n)}_{\substack{\text{when } \mathcal{F}_n \text{ represents the current state of information \& } X_n \\ \text{represents a gambler fortune, this says future fortune} \\ \text{decreases in expected value given } \mathcal{F}_n.}} \overset{P\text{-a.e.}}{\leq} X_n \quad \forall n\in\mathbb{N}$$

**Definition:** If $(X_n)_{n\in\mathbb{N}}$ are r.v.s on $(\Omega, \mathcal{F}, P)$ then the **natural filtration** is

$$\mathcal{F}_n := \sigma\langle X_1, \ldots, X_n\rangle$$

**Notation:** As a shorthand we let

"subM" $\equiv$ submartingale

"supM" $\equiv$ super martingale

"M" $\equiv$ martingale

Notice that if $(X_n)_{n\geq 1}$ is a subM then clearly $(-X_n)_{n\geq 1}$ is a supM by (e1).

Also if $(X_n)_{n\geq 1}$ is both a subM & supM it is then a M.

Here is another immediate consequence of the definitions

**Proposition**

$(X_n)_{n\geq 1}$ is a subM w.r.t. filtration $(\mathcal{F}_n)_{n\geq 1}$

$$\Updownarrow$$

$$E(X_m \mid \mathcal{F}_n) \overset{a.e.}{\geq} X_n \quad \forall m > n$$

**Proof:**

($\Leftarrow$): Trivial

($\Rightarrow$): If $m > n$ then $\mathcal{F}_n \subset \mathcal{F}_m$ ← a more dramatic smoother

$$E(X_{m+1} \mid \mathcal{F}_n) \underset{(s2)}{\overset{a.e.}{=}} E(\underbrace{E(X_{m+1} \mid \mathcal{F}_m)}_{\text{an intermediate smoothing}} \mid \mathcal{F}_n)$$

$$\underset{(e3)}{\overset{a.e.}{\geq}} E(X_m \mid \mathcal{F}_n).$$

The result now follows by induction.

QED

**Corollary:**

$(X_n)_{n\geq}$ is a subM $\Longrightarrow E(X_m) \geq E(X_n) \quad \forall m > n$

$(X_n)_{n\geq}$ is a supM $\Longrightarrow E(X_m) \leq E(X_n) \quad \forall m > n$

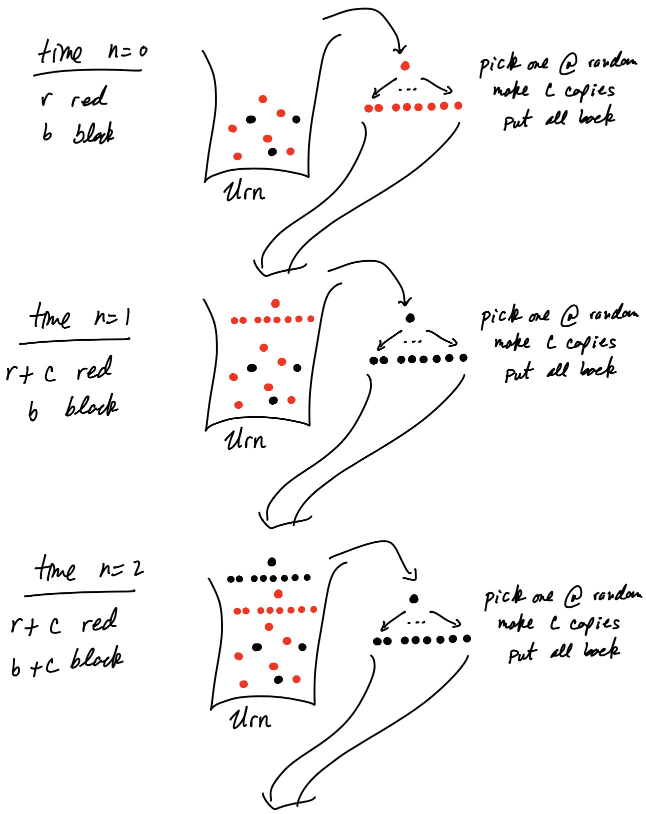
$(X_n)_{n\geq}$ is a M $\Longrightarrow E(X_m) = E(X_n) \quad \forall m > n$

Proof: Simple application of (s1)



## Pólya's Urn

This is a classic example of a martingale (and exchangable sequences of r.v.s) and a good canonical example of a martingale to have in your mind.




Repeat

⋮

Let $Y_n := \begin{cases} 1 & \text{if the } n^{th} \text{ draw is red.} \\ 0 & \text{if} \end{cases}$

$R_n :=$ proportion of reds in the urn at time $n$

$\mathcal{F}_n := \sigma\langle R_1, R_2, \ldots, R_n \rangle$

Question: Does $R_n$ have a limiting distribution?

First thing to notice is that

$$P(Y_n = 1 | R_1, \dots, R_n) = R_n$$

so that

$$E(Y_n | \mathcal{F}_n) = E(Y_n | R_1, \dots, R_n)$$

$$= \sum_{y \in 0,1} y\, P(Y_n = y | R_1, \dots, R_n)$$

$$= R_n$$

Moreover $R_n = \dfrac{r + c(Y_0 + \dots + Y_{n-1})}{r+b+nc}$ ← each time you add c more balls

which implies

$$E(R_{n+1} | \mathcal{F}_n)$$

$$= E\left( \frac{r + c(Y_0 + \dots + Y_n)}{r+b+(n+1)c} \middle| \mathcal{F}_n \right)$$

$$= E\left( \frac{r+b+nc}{r+b+(n+1)c} \underbrace{\frac{r + c(Y_0 + \dots + Y_{n-1})}{r+b+nc}}_{= R_n \text{ (m) } \mathcal{F}_n} \middle| \mathcal{F}_n \right) + E\left( \frac{c Y_n}{r+b+(n+1)c} \middle| \mathcal{F}_n \right)$$

$$= \frac{r+b+nc}{r+b+(n+1)c} R_n + \frac{c}{r+b+(n+1)c} R_n$$

$$= R_n$$

Since $R_n$ is trivially (m) $\mathcal{F}_n$ we therefore have that $R_n$ is a martingale. We will see that under mild conditions martingales converge.



## Transformation of Martingales

**Theorem: (Transformations)**

(i) If $(X_n)_{n \geq 1}$ & $(Y_n)_{n \geq 1}$ are subMs w.r.t. $(\mathcal{F}_n)_{n \geq 1}$ then so are $(X_n + Y_n)_{n \geq 1}$ and $(X_n \vee Y_n)_{n \geq 1}$.

(ii) If $(X_n)_{n \geq 1}$ is a M w.r.t. $(\mathcal{F}_n)_{n \geq 1}$, $f: \mathbb{R} \to \mathbb{R}$ is convex and $f(X_n) \in L_1(\Omega, \mathcal{F}, P)$ then $(f(X_n))_{n \geq 1}$ is a subM w.r.t. $(\mathcal{F}_n)_{n \geq 1}$.

(iii) If $(X_n)_{n \geq 1}$ is a subM w.r.t. $(\mathcal{F}_n)_{n \geq 1}$, $f: \mathbb{R} \to \mathbb{R}$ is convex, $f$ is non-decreasing and $f(X_n) \in L_1(\Omega, \mathcal{F}, P)$ then $(f(X_n))_{n \geq 1}$ is a subM w.r.t. $(\mathcal{F}_n)_{n \geq 1}$.

(iv) If $(X_{1n})_{n \geq 1}, (X_{2n})_{n \geq 1}, \dots, (X_{mn})_{n \geq 1}$ are m subM wrt $(\mathcal{F}_n)_{n \geq 1}$ and $w_1, w_2, \dots, w_m \geq 0$ then

$$Y_n := \sum_{i=1}^{m} w_i X_{in}$$

is a subM wrt $(\mathcal{F}_n)_{n \geq 1}$.

*Proof:*

For (i):

Remember that the def of subM, supM & M require $X_n, Y_n \in L_1(\Omega, \mathcal{F}, P)$ so the fact that $(X_n + Y_n)_{n \geq 1}$ is a subM is a trivial consequence of (e2) from lecture 20.

To show $(X_n \vee Y_n)_{n\geq 1}$ is a subM notice that

$$X_{n+1} \leq X_{n+1} \vee Y_{n+1}$$
$$Y_{n+1} \leq X_{n+1} \vee Y_{n+1}$$

$$\therefore\ X_n \overset{p\text{-a.e.}}{\leq} E(X_{n+1} | \mathcal{F}_n) \overset{p\text{-a.e.}}{\leq} E(X_{n+1} \vee Y_{n+1} | \mathcal{F}_n) \qquad (*)$$
$$Y_n \overset{p\text{-a.e.}}{\leq} E(Y_{n+1} | \mathcal{F}_n) \overset{p\text{-a.e.}}{\leq} E(X_{n+1} \vee Y_{n+1} | \mathcal{F}_n)$$

by monotonicity (e3).

$$\therefore\ X_n \vee Y_n \overset{p\text{-a.e.}}{\leq} E(X_{n+1}|\mathcal{F}_n) \vee E(Y_{n+1}|\mathcal{F}_n)$$

by subM prop.

$$\overset{p\text{-a.e.}}{\leq} E(X_{n+1} \vee Y_{n+1} | \mathcal{F}_n)$$

by $(*)$

Also $|X_n \vee Y_n| \leq |X_n| + |Y_n| \in L_1(\Omega, \mathcal{F}, P)$ so that $(X_n \vee Y_n)_{n\geq 1}$ satisfies the subM properties.

For (ii) & (iii)

$$E(f(X_{n+1}) | \mathcal{F}_n) \overset{p\text{-a.e.}}{\geq} f(E(X_{n+1}|\mathcal{F}_n)) \overset{p\text{-a.e.}}{=} f(X_n)$$

by Jensen for conditional expected value (proved in the exact same way as was done in Lecture 11

The prores (ii) & (iii) is similar.

For (iv)

Follows easily by linear properties of $E(\cdot|\mathcal{F}_n)$.

QED



Notice a few simple consequences of the previous thm:

$(X_n)_{n\geq 1}$ is a M $\Rightarrow$ $(|X_n|)_{n\geq 1}$ is a subM

$\left.\begin{array}{l}(X_n)_{n\geq 1} \text{ is a M} \\ \& \ X_n \in L_2\end{array}\right\} \Rightarrow (X_n^2)_{n\geq 1}$ is a subM

$(X_n)_{n\geq}$ is a subM $\Longrightarrow$ $(X_n^+)_{n\geq 1}$ is a subM

## A collection of examples

**Example: (strategy against a bad game)**

$X_0$ = a gamblers initial fortune

$X_n$ = fortune after n-plays of a game when always betting #1 on each play,

i.e. $X_n - X_{n-1}$ = winnings obtained on the $n^{th}$ play

Suppose $(X_n)_{n\geq 1}$ is a supM wrt $(\mathcal{F}_n)_{n\geq 1}$.

Since $X_n$ is a supM your expected winnings are monotonically decreasing with n

Question:

Can I wager differently (bet more on hot streaks) to convert $X_n$ into a subM?

Let $W_n$ be the wager on n-th play s.t.

$W_n \geq 0$

$W_n$ is $\mathcal{F}_{n-1}$-measurable — I can change bet based on the future

$W_n$ is bounded — i.e. there is a house max bet

Winnings on the $n^{th}$ play is

$$W_n \underbrace{(X_n - X_{n-1})}_{=: \Delta_n \text{ = winnings on } n^{th} \text{ play wagering \$1}}$$

The total fortune after n plays with wagers $W_n$ is given by

$$X_n^* = X_0 + W_1 \Delta_1 + W_2 \Delta_2 + \cdots + W_n \Delta_n$$

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

claim: $(X_n^*)_{n \geq 1}$ is a supM w.r.t $\mathcal{F}_n$ ← so no matter the strategy still a losing game

Prof:

clearly $X_n^*$ is adapted to $\mathcal{F}_n$ an $L_1(\Omega, \mathcal{F}, P)$ by bddness.

Also

@ $\mathcal{F}_n$    @ $\mathcal{F}_{n+1}$

$$E(X_{n+1}^* | \mathcal{F}_n) \overset{P\text{-a.e}}{=} E(X_n^* + W_{n+1} \Delta_{n+1} | \mathcal{F}_n)$$

$$\overset{P\text{-a.e.}}{=} X_n^* + W_{n+1} E(\Delta_{n+1} | \mathcal{F}_n)$$

$$\overset{P\text{-a.e.}}{=} \underbrace{X_n^* + W_{n+1}}_{\text{positive}} \underbrace{\left[\underbrace{E(X_{n+1}|\mathcal{F}_n)}_{\leq X_n} - \underbrace{E(X_n|\mathcal{F}_n)}_{= X_n}\right]}_{\text{negative}}$$

$$\leq X_n^*$$

∴ indeed $(X_n^*)_{n \geq 1}$ is a supM.



Now suppose $(X_n)_{n \geq 1}$ is a subM ← a favorable game

By similar reasoning $(X_n^*)_{n \geq 1}$ is also a subM.

Suppose $W_n \in [0,1]$ ← timid play

Now

$$E(\underbrace{X_{n+1}^* - X_n^*}_{n+1^{st} \text{ gain}} | \mathcal{F}_n) \overset{a.e.}{=} E(W_{n+1}(X_{n+1} - X_n) | \mathcal{F}_n)$$

@ $\mathcal{F}_n$

$$\overset{a.e.}{=} \underbrace{W_{n+1}}_{\in [0,1]} \underbrace{E(X_{n+1} - X_n | \mathcal{F}_n)}_{\geq 0 \text{ by subM}}$$

$$\overset{a.e.}{\leq} E(X_{n+1} - X_n | \mathcal{F}_n)$$

Take E of both sides gives

$$E(X_{n+1}^* - X_n^*) \leq E(X_{n+1} - X_n)$$

$$\therefore \sum_{n=0}^{N-1} E(X_{n+1}^* - X_n^*) \leq \sum_{n=0}^{N-1} E(X_{n+1} - X_n)$$

$$\| \qquad\qquad \|$$

$$E(X_N^*) - \underbrace{E(\cancel{X_0^*})} \leq E(X_N) - \underbrace{E(\cancel{X_0})}$$

same inital fortune

$$\therefore E(X_N^*) \leq E(X_N)$$

i.e. timid play reduces expected fortune.

**Example:** Smoothing Martingale

$(\mathcal{F}_n)_{n\geq 1}$ is a filtration.

Let $X \in L_1(\Omega, \mathcal{F}, P)$ &

$$X_n := E(X|\mathcal{F}_n)$$

Clearly $X_n$ is $\mathcal{F}_n$-measurable and in $L_1(\Omega,\mathcal{F},P)$ by smoothing property (S3) from Lecture 20. Also

$$E(X_{n+1}|\mathcal{F}_n) \underset{def}{\overset{a.e.}{=}} E(E(X|\mathcal{F}_{n+1})|\mathcal{F}_n)$$

$$\underset{(S2)}{\overset{a.e.}{=}} E(X|\mathcal{F}_n)$$

$$\overset{a.e.}{=} X_n$$

$\therefore$ $(X_n)_{n\geq 1}$ is a $M$ w.r.t. $(\mathcal{F}_n)_{n\geq 1}$.

Next lecture we will investigate if

$$X_n \overset{?}{\longrightarrow} X$$

**Example:** Martingale difference sequences

If $(X_n)_{n\geq 1}$ is a $M$ w.r.t. $(\mathcal{F}_n)_{n\geq 1}$ we an always write

$$X_n = \underbrace{X_1}_{ⓜ\mathcal{F}_1} + \underbrace{(X_2 - X_1)}_{ⓜ\mathcal{F}_2} + \cdots + \underbrace{(X_n - X_{n-1})}_{ⓜ\mathcal{F}_n}$$

where the terms in the sum satisfy

$$E\big((X_{i+1} - X_i)\,|\,\mathcal{F}_i\big) \overset{a.e.}{=} E(X_{i+1}|\mathcal{F}_i) - X_i$$

$$\overset{a.e.}{=} 0$$



**Definition:** If $(\Delta_n)_{n\geq 1}$ are real valued r.v.s on $(\Omega, \mathcal{F}, P)$ then $(\Delta_n)_{n\geq 1}$ is a subM, M or supM difference sequence w.r.t. $(\mathcal{F}_n)_{n\geq 1}$ iff

(i) $\Delta_n \in L_1(\Omega, \mathcal{F}_n, P|_{\mathcal{F}_n})$ and

(ii) $$E(\Delta_{n+1}|\mathcal{F}_n) \begin{cases} \leq 0 & \text{for supM} \\ = 0 & \text{for M} \\ \geq 0 & \text{for subM} \end{cases}$$

**Example:** Lebesgue decomposition supM

Suppose you have two models $P$ & $Q$ for a random $\omega \in \Omega$ generating a infinite sequence of r.v.s $X = (X_1, X_2, \dots)$

$$\underset{\mathcal{F}\quad P \text{ or } Q}{\Omega} \xrightarrow{(X_1, X_2, \dots)ⓜ} \underset{\mathcal{B}(\mathbb{R}^\infty)}{\mathbb{R}^\infty}$$

Question 1: How an one check if

$$\underbrace{PX^{-1} \perp QX^{-1}}_{} \quad \text{or} \quad \underbrace{PX^{-1} \equiv QX^{-1}}_{}$$

if this holds then one sample from $X$ allows you to determine which method generated it

$\omega \sim P \Rightarrow X(\omega)$

or

$\omega \sim Q \Rightarrow X(\omega)$

if this hold it will be impossible to distinguish $P$ or $Q$ from one sample of $X = (X_1, \dots)$

Question 2: If $PX^{-1} \perp QX^{-1}$ can one distinguish $P$ from $Q$ as $n \to \infty$ when observing $X$ at the first $n$-coordinates:

$$X_{1:n} := (X_1, X_2, \dots, X_n) \in \mathbb{R}^n$$

Let

$$\mathcal{F}_n := \sigma\langle X_1, \dots, X_n \rangle$$

$$\mathcal{F}_\infty := \sigma\langle X_i : i \geq 1 \rangle$$

$$Q_n := Q\big|_{\mathcal{F}_n}$$

$$Q_\infty := Q\big|_{\mathcal{F}_\infty}$$

} & similarly for $P$

$$Q_n = \underbrace{Q_n^a + Q_n^s}$$

The Lebesgue decomposition w.r.t $P_n$ where $Q_n^a \ll P_n$ & $Q_n^s \perp P_n$

$$Q_\infty = \underbrace{Q_\infty^a + Q_\infty^s}$$

The Lebesgue decomposition w.r.t $P_\infty$

Note: $Q_\infty \perp P_\infty \iff Q_\infty^a \equiv 0$

Suppose $q_n$ & $p_n$ are densities of $Q_n(X_1,\dots,X_n)^{-1}$ and $P_n(X_1,\dots,X_n)^{-1}$ w.r.t. some dominating measure $\mu_n$ (could be $\mathcal{L}^n$ or $Q_n(X_1,\dots,X_n)^{-1} + P_n(X_1,\dots,X_n)^{-1}$ e.g.)

Define $q_\infty$, $p_\infty$ & $\mu_\infty$ similarly.

Now

$$\frac{dQ_n^a}{dP_n} \overset{P_n\text{-a.e.}}{=} \frac{q_n(X_1,\dots,X_n)}{p_n(X_1,\dots,X_n)} I_{\{p_n(X_1,\dots,X_n) > 0\}}$$

and

$$\frac{dQ_\infty^a}{dP_\infty} \overset{P_\infty\text{-a.e.}}{=} \frac{q_\infty(X_1,X_2,\dots)}{p_\infty(X_1,X_2,\dots)} I_{\{p_\infty(X_1,X_2,\dots) > 0\}}$$

which follow by the proof of the Lebesgue Decomposition Thm & change of variables.



Now

$$PX^{-1} \perp QX^{-1}$$

since events in $\mathcal{F}_\infty$ are of the form $X^{-1}(B)$ for $B \in \mathcal{B}(\mathbb{R}^\infty)$

$$\iff Q_\infty \perp P_\infty$$

$$\iff Q_\infty^a \equiv 0$$

$$\iff \frac{dQ_\infty^a}{dP_\infty} = 0 \quad P_\infty\text{-a.e.}$$

$$\overset{?}{\iff} \frac{dQ_n^a}{dP_n} \overset{?}{\longrightarrow} 0$$

see next lecture..

$$\overset{?}{\iff} H(Q_n, P_n) \longrightarrow \sqrt{2}$$

see exercise.

where $H(\cdot,\cdot)$ is the Hellinger distance on probability measures:

$$H^2(Q_n,P_n) := \int_{\mathbb{R}^n} (\sqrt{q_n} - \sqrt{p_n})^2 \, d\mu_n = 2\Big(1 - \int_{\mathbb{R}^n} \sqrt{q_n}\sqrt{p_n}\, d\mu_n\Big)$$

$$\overset{?}{\iff} KL(Q_n \| P_n) \longrightarrow \infty$$

see this later for Gaussian measures

where $KL(\cdot\|\cdot)$ is the symmetrized Kullback-Liebler dist btwn $P_n$ & $Q$:

$$KL(Q_n \| P_n) := \int_\Omega \log \frac{dQ_n^a}{dP_n} \, dP_n + \int_\Omega \log \frac{dP_n^a}{dQ_n} \, dQ_n$$

We will resolve the question marks later but notice that we can get partway to resolving some by noticing

$\frac{dQ_n^a}{dP_n}$ is a supM w.r.t. $\mathcal{F}_n$ on $(\Omega, \mathcal{F}, P)$

Indeed $\frac{dQ_n^a}{dP_n} \in L_1(\Omega, \mathcal{F}_n, P_n)$ and $\forall A \in \mathcal{F}_n$

$$\int_A E\Big(\frac{dQ_{n+1}^a}{dP_{n+1}} \Big| \mathcal{F}_n\Big) dP = \int_A \frac{dQ_{n+1}^a}{dP_{n+1}} \, dP$$

can replace with $dP_{n+1}$ by C.V.

$$= Q_{n+1}^a(A)$$

$$\leq Q_{n+1}(A), \quad Q_{n+1} = Q_{n+1}^a + Q_{n+1}^s$$

$$= Q_n(A) \quad \text{since } A \in \mathcal{F}_n$$

$\therefore \int_A E\left(\frac{dQ^a_{n+1}}{dP_{n+1}} \Big| \mathcal{F}_n\right) dP_n \leq Q_n(A), \quad \forall A \in \mathcal{F}_n$

but we also have that

$$\int_A \underbrace{\frac{dQ^a_n}{dP_n}} dP_n \leq Q_n(A), \quad \forall A \in \mathcal{F}_n$$

where this is the $P_n$-largest such, by Lebesgue decomp thm

$$\therefore \quad E\left(\frac{dQ^a_{n+1}}{dP_{n+1}} \Big| \mathcal{F}_n\right) \overset{P_n\text{-a.e.}}{\leq} \frac{dQ^a_n}{dP_n}$$

i.e. $\left(\frac{dQ^a_n}{dP_n}\right)_{n\geq 1}$ is a supM w.r.t. $(\mathcal{F}_n)_{n\geq 1}$ on $(\Omega, \mathcal{F}, P)$.

Also notice that

$$Q_n \ll P_n \implies \Bigg(\underbrace{\overset{\text{likelihood ratio}}{\overset{\frac{dQ^a_n}{dP_n}}{\overset{\shortparallel}{\frac{dQ_n}{dP_n}}}}}_{\text{density}}\Bigg)_{n\geq 1} \text{ is a } M \text{ w.r.t. } (\mathcal{F}_n)_{n\geq 1} \text{ on } (\Omega, \mathcal{F}, P).$$

Since $\forall A \in \mathcal{F}_n$

$$\begin{aligned} \int_A E\left(\frac{dQ_{n+1}}{dP_{n+1}} \Big| \mathcal{F}_n\right) dP &= \int_A \frac{dQ_{n+1}}{dP_{n+1}} dP \\ &= Q_{n+1}(A) \\ &= Q_n(A) \\ &= \int_A \frac{dQ_n}{dP_n} dP \end{aligned}$$

so that $E\left(\frac{dQ_{n+1}}{dP_{n+1}} \Big| \mathcal{F}_n\right) \overset{P\text{-a.e.}}{=} \frac{dQ_n}{dP_n}$ by uniqueness of densities.



# Stopping times

Stopping times are a dynamic rule for when to stop gambling ... but needs to be computable before the next game is about to occur.

**Definition:** A function $\tau : \Omega \to \overline{\mathbb{N}}$ (where $\overline{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$) is a **stopping time** for a filtration $(\mathcal{F}_n)_{n\geq 1}$ if $\underbrace{\{\tau = n\}}_{\substack{\text{Determine to} \\ \text{stop or not after} \\ n^{th} \text{ game is played} \\ \text{but before } n+1^{st} \text{ is played}}} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}$.

When $X_n$ represents gamblers total fortune think of $X_\tau$ as "final fortune".

Note: $\tau$ is $\mathcal{F}$-measurable since $\mathcal{B}(\overline{\mathbb{N}}) = \sigma\langle \{n\} : n \in \overline{\mathbb{N}} \rangle$ and by generators are enough we just check that

$$\{\tau = n\} \in \mathcal{F}_n \in \mathcal{F} \checkmark$$

$$\{\tau = \infty\} = \{\tau < \infty\}^c = \bigcap_{n=1}^{\infty} \{\tau = n\}^c \in \mathcal{F} \checkmark$$

For continuous time martingales it will be convienent to use the condition $\{\tau \leq n\} \in \mathcal{F}_n$ rather than $\{\tau = n\} \in \mathcal{F}_n$. In our case they are equivalent:

$$\{\tau = n\} \in \mathcal{F}_n, \forall n \in \mathbb{N} \Rightarrow \{\tau \leq n\} = \bigcup_{k=1}^{n} \underbrace{\{\tau = k\}}_{\in \mathcal{F}_k \subset \mathcal{F}_n} \in \mathcal{F}_n$$

$$\{\tau \leq n\} \in \mathcal{F}_n, \forall n \in \mathbb{N} \Rightarrow \{\tau = n\} = \underbrace{\underbrace{\{\tau \leq n\}}_{\in \mathcal{F}_n} - \underbrace{\{\tau \leq n-1\}}_{\in \mathcal{F}_{n-1} \subset \mathcal{F}_n}}_{\in \mathcal{F}_n}$$

**Proposition:** (The $X_n$ values determine when to stop for the natural filtration)

If $(\mathcal{F}_n)_{n\geq 1}$ is the natural filtration for $(X_n)_{n\geq 1}$ and $\tau$ is a stopping time w.r.t. $(\mathcal{F}_n)_{n\geq 1}$ then $\forall\, w_1, w_2 \in \Omega$

$$\left.\begin{array}{r} \tau(w_1) = n < \infty \ \& \\ X_k(w_1) = X_k(w_2),\ \forall\, k \leq n \end{array}\right\} \Longrightarrow \tau(w_2) = n$$

*Proof:*

Since $\{\tau = n\} \in \mathcal{F}_n = \sigma\langle X_1, \dots, X_n \rangle$ we have

$$\{\tau = n\} = \{w \in \Omega : (X_1(w), \dots, X_n(w)) \in B\}$$

for some $B \in \mathcal{B}(\mathbb{R}^n)$.

$$\begin{aligned} \tau(w_1) = n < \infty &\iff (X_1(w_1), \dots, X_n(w_1)) \in B \\ &\iff (X_1(w_2), \dots, X_n(w_2)) \in B \\ &\iff \tau(w_2) = n < \infty \end{aligned}$$

QED

**Example:** Defining $\tau = 5$ gives a stopping time since $\forall\, n \in \mathbb{N}$

$$\{\tau = n\} = \begin{cases} \Omega & \text{if } n = 5 \\ \emptyset & \text{o.w.} \end{cases} \in \mathcal{F}_n$$

**Example:**

Let $c \in \mathbb{R}$ & $(X_n)_{n\geq 1}$ be adapted to $(\mathcal{F}_n)_{n\geq 1}$

& $\tau :=$ the first time $X_n$ enters $[c, \infty)$

$$= \inf\left(\{n : X_n \geq c\} \cup \underbrace{\{\infty\}}_{\text{so } \tau = \infty \text{ if no } n \in \mathbb{N} \text{ s.t. } X_n \geq c}\right)$$

$\tau$ is a stopping time since

$$\{\tau \leq n\} = \underbrace{\{X_1 \geq c\}}_{\in \mathcal{F}_1} \cup \dots \cup \underbrace{\{X_n \geq c\}}_{\in \mathcal{F}_n} \in \mathcal{F}_n$$



**Proposition:** ($\sigma \wedge \tau$ is a ST)

If $\sigma$ & $\tau$ are stopping times for a filtration $(\mathcal{F}_n)_{n\geq 1}$ then so is

$$\sigma \wedge \tau := \min(\sigma, \tau).$$

*Proof:* $\{\sigma \wedge \tau \leq n\} = \underbrace{\{\sigma \leq n\}}_{\in \mathcal{F}_n} \cup \underbrace{\{\tau \leq n\}}_{\in \mathcal{F}_n}.$

QED

The above proposition is particularly useful for truncating a ST $\tau$ @ $k$: $\underbrace{\tau \wedge k}_{\text{bdd stopping}}$ is a ST for any fixed $k < \infty$

Recall that $\mathcal{F}_n$ represents information available to you at time $n$.

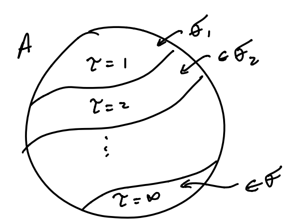
What represents information at time $\tau$?

**Definition:** If $\tau$ is a stopping time w.r.t filtration $(\mathcal{F}_n)_{n\geq 1}$ define $\mathcal{F}_\tau$ to be the collection of all $A \subset \Omega$ s.t.

$$A \cap \{\tau = n\} \in \mathcal{F}_n, \quad \forall\, n \in \mathbb{N} \ \&$$

$$A \cap \{\tau = \infty\} \in \mathcal{F} \quad \hookrightarrow \text{so that } \mathcal{F}_\tau \subset \mathcal{F}$$

So $A \in \mathcal{F}_\tau$ has the form



**Proposition: (equiv def of $\mathcal{F}_\tau$)**

If $\tau$ is a stopping time w.r.t filtration $(\mathcal{F}_n)_{n\geq 1}$ then

$$A \in \mathcal{F}_\tau \overset{(a)}{\iff} A \in \mathcal{F} \ \& \ A \cap \{\tau = n\} \in \mathcal{F}_n, \ \forall n \in \mathbb{N}$$

$$\overset{(b)}{\iff} A \in \mathcal{F} \ \& \ A \cap \{\tau \leq n\} \in \mathcal{F}_n, \ \forall n \in \mathbb{N}$$

Proof:

$(\overset{(a)}{\Longrightarrow})$: Easy since

$$A = \underbrace{(A \cap \{\tau = \infty\})}_{\in \mathcal{F}} \cup \underbrace{(A \cap \{\tau = 1\})}_{\in \mathcal{F}_1} \cup \underbrace{(A \cap \{\tau = 2\})}_{\in \mathcal{F}_2} \cup \cdots \in \mathcal{F}$$

$(\overset{(a)}{\Longleftarrow})$: Easy since

$$A \cap \{\tau = \infty\} = \underbrace{A}_{\in \mathcal{F}} \cap \underbrace{\{\tau < \infty\}^c}_{\in \mathcal{F}} \in \mathcal{F}.$$

$(\overset{(b)}{\Longrightarrow})$: $A \cap \{\tau \leq n\} = A \cap \bigcup_{k=1}^{n} \{\tau = k\} \in \mathcal{F}_n$

$(\overset{(b)}{\Longleftarrow})$: $A \cap \{\tau = n\} = \underbrace{A \cap \{\tau \leq n\}}_{\in \mathcal{F}_n} - \underbrace{A \cap \{\tau \leq n-1\}}_{\in \mathcal{F}_{n-1} \subset \mathcal{F}_n} \in \mathcal{F}_n$

QED

**Proposition: ($X_\tau$ is $\mathcal{F}_\tau$-measurable & $\mathcal{F}_\tau$ is a $\sigma$-field)**

If $\tau$ is a stopping time w.r.t filtration $(\mathcal{F}_n)_{n\geq 1}$ and $(X_n)_{n\geq 1}$ is adapted to $(\mathcal{F}_n)_{n\geq 1}$ then $\mathcal{F}_\tau$ is a sub $\sigma$-field of $\mathcal{F}$.

If, in addition, $\tau < \infty$ then

$$X_\tau = X_{\tau(\omega)}(\omega) \text{ is } \mathcal{F}_\tau\text{-measurable.}$$

Proof:

Clearly $\mathcal{F}_\tau \subset \mathcal{F}$ by the previous proposition. To show $\mathcal{F}_\tau$ is a $\sigma$-field notice ...



- $\Omega \in \mathcal{F}_\tau$ since $\Omega \in \mathcal{F}$ & $\Omega \cap \{\tau = n\} = \{\tau = n\} \in \mathcal{F}_n$
- $A \in \mathcal{F}_\tau \Longrightarrow A^c \in \mathcal{F}_\tau$ since $A^c \in \mathcal{F}$ & $A^c \cap \{\tau = n\} = \{\tau = n\} - A \cap \{\tau = n\} \in \mathcal{F}_n$
- $A_1, \ldots \in \mathcal{F}_\tau \Longrightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{F}_\tau$ since $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$ & $\left(\bigcup_{k=1}^{\infty} A_k\right) \cap \{\tau = n\} = \bigcup_{k=1}^{\infty} \underbrace{A_k \cap \{\tau = n\}}_{\in \mathcal{F}_n} \in \mathcal{F}_n$

$\therefore$ $\mathcal{F}_\tau$ is a sub $\sigma$-field of $\mathcal{F}$.

Now suppose $\tau(\omega) < \infty$, $\forall \omega \in \Omega$.

Let $B \in \mathcal{B}(\mathbb{R})$ and show $\{X_\tau \in B\} \in \mathcal{F}_\tau$

i.e. show

(i) $\{X_\tau \in B\} \cap \{\tau = n\} \in \mathcal{F}_n$, $\forall n \in \mathbb{N}$ &

(ii) $\{X_\tau \in B\} \cap \{\tau = \infty\} \in \mathcal{F}$

Condition (i) clearly holds since the LHS is

$$\underbrace{\{X_n \in B\}}_{\in \mathcal{F}_n} \cap \underbrace{\{\tau = n\}}_{\in \mathcal{F}_n} \in \mathcal{F}_n$$

Condition (ii) holds since $\{\tau = \infty\} = \emptyset$ by assumption.

QED

## Proposition: ($\sigma \le \tau \Rightarrow \mathcal{F}_\sigma \subset \mathcal{F}_\tau$)

If $\sigma$ & $\tau$ are ST wrt filtration $(\mathcal{F}_n)_{n\ge1}$

then (i) $\{\sigma \le \tau\} \in \mathcal{F}_\tau$

(ii) $\sigma(\omega) \le \tau(\omega)\ \forall \omega \in \Omega \Rightarrow \mathcal{F}_\sigma \subset \mathcal{F}_\tau$.

**Proof:**

For (i)

$$\underbrace{\{\sigma \le \tau\}}_{\substack{\in \mathcal{F} \text{ since} \\ \text{both are (m)}}} \cap \{\tau = n\} = \{\sigma \le n\} \in \mathcal{F}_n$$

For (ii)

$$A \in \mathcal{F}_\sigma \Rightarrow A \in \mathcal{F} \ \& \ A \cap \{\sigma = n\} \in \mathcal{F}_n$$

$$\Rightarrow A \in \mathcal{F} \ \&$$

$$A \cap \{\tau = n\} = A \cap \{\tau = n\} \cap \overbrace{\{\sigma \le \tau\}}^{\Omega}$$

$$= A \cap \underbrace{\{\tau = n\}}_{\in \mathcal{F}_n} \cap \{\sigma \le n\}$$

the intersection of these $\in \mathcal{F}_n$

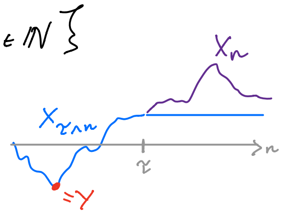
$$\Rightarrow A \in \mathcal{F}_\tau$$

QED

**Example:** Let $(X_n)_{n\ge1}$ be adapted to the filtration $(\mathcal{F}_n)_{n\ge1}$ & $\tau$ be a ST.

Then $Y := \inf\{X_{\tau\wedge n} : n \in \mathbb{N}\}$

is $\mathcal{F}_\tau$-measurable since




$\tau \wedge n$ is a ST $\Rightarrow X_{\tau\wedge n}$ (m) $\mathcal{F}_{\tau\wedge n}$

$\Rightarrow X_{\tau\wedge n}$ (m) $\mathcal{F}_\tau$ since $\mathcal{F}_{\tau\wedge n} \subset \mathcal{F}_\tau$

$\Rightarrow \inf_n X_{\tau\wedge n}$ (m) $\mathcal{F}_\tau$ by closure.



## Example: Strong Markov property of Brownian Motion

## Theorem: (Finite optimal sampling)

Let $(X_1, X_2, \dots, X_n)$ be a subM w.r.t. filtration $(\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_n)$ and $\sigma, \tau$ be stopping times s.t. $\sigma \le \tau \le n$. Then

$(X_\sigma, X_\tau)$ is a subM wrt $(\mathcal{F}_\sigma, \mathcal{F}_\tau)$

Proof:

Previous propositions establish that $(\mathcal{F}_\sigma, \mathcal{F}_\tau)$ is a filtration and $(X_\sigma, X_\tau)$ is adapted to $(\mathcal{F}_\sigma, \mathcal{F}_\tau)$.

Integrability follows since

$$|X_\sigma|, |X_\tau| \le |X_1| + \dots + |X_n| \in L_1(\Omega, \mathcal{F}, P).$$

Now

$$E(X_\tau | \mathcal{F}_\sigma) \overset{a.e.}{\geq} X_\sigma.$$

$$\iff \int_A E(X_\tau | \mathcal{F}_\sigma)\, dP \geq \int_A X_\sigma\, dP \quad \forall A \in \mathcal{F}_\sigma$$

"

$$\iff \int_A X_\tau\, dP \geq \int_A X_\sigma\, dP, \quad \forall A \in \mathcal{F}_\sigma$$

$$\iff E((X_\tau - X_\sigma) I_A) \geq 0 \quad \forall A \in \mathcal{F}_\sigma$$

Note that

$$(X_\tau - X_\sigma) I_A = \sum_{k=\sigma}^{\tau-1} (X_{k+1} - X_k) I_A$$

$$= \sum_{k=1}^{n-1} (X_{k+1} - X_k) I_{A_k}$$

where $A_k = A \cap \{\sigma \le k \le \tau - 1\}$

$$= \underbrace{A \cap \{\sigma \le k\}}_{\in \mathcal{F}_k} \cap \underbrace{\{\tau \le k\}^c}_{\in \mathcal{F}_k} \in \mathcal{F}_k$$

since $A \in \mathcal{F}_\sigma$



$$\therefore E((X_{k+1} - X_k) I_{A_k} | \mathcal{F}_k)$$

$$\overset{a.e.}{=} I_{A_k} \underbrace{E(X_{k+1} - X_k | \mathcal{F}_k)}_{\overset{\geq 0}{a.e.}} \overset{a.e.}{\geq} 0$$

$$\therefore E((X_{k+1} - X_k) I_{A_k}) \geq 0$$

$\therefore E((X_\tau - X_\sigma) I_A) \geq 0$ as was to be shown

QED

The Finite optional sampling thm is a tool for generating some useful inequalities. Here is an example which generalizes Kolmogorov's Maximal inequality.

## Theorem: (Kolmogorov's ineq for subM)

Let $(X_1, X_2, \dots, X_n)$ be a subM w.r.t a filtration $(\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_n)$.

If $M_n := \max(X_1, X_2, \dots, X_n)$ then

$$P(M_n \geq c) \le \frac{E(X_n I_{\{M_n \geq c\}})}{c} \le \frac{E(X_n^+)}{c}$$

for all $c > 0$.

Proof:

Let $\tau :=$ first time $X_k$ enters $[c, \infty)$ (or $n$ if it never does)

$$= \begin{cases} \min\{k : X_k \geq c\}, & \text{if } M_n \geq c \\ n & \text{o.w.} \end{cases}$$

A previous example showed $\tau$ is a ST.

$\therefore$ $(X_\tau, X_n)$ is a subM wrt $(\mathcal{F}_\tau, \mathcal{F}_n)$

$\therefore$ $E(X_n \mid \mathcal{F}_\tau) \overset{P\text{-a.e.}}{\geq} X_\tau$

$\therefore$ $\int_A X_n \, dP \geq \int_A X_\tau \, dP \quad \forall A \in \mathcal{F}_\tau$

Since $\{M_n \geq c\} = \{X_\tau \geq c\} \in \mathcal{F}_\tau$ we have

$$\int_{M_n \geq c} X_n \, dP \geq \int_{X_\tau \geq c} X_\tau \, dP \geq c P(\underbrace{X_\tau \geq c}_{= M_n \geq c})$$

$\therefore$ $P(M_n \geq c) \leq \frac{1}{c} \int_{M_n \geq c} X_n \, dP \leq \frac{1}{c} \int_\Omega X_n^+ \, dP$

QED.

**Example:** This example shows that Kolmogorov's Max-ineq for subM is a generalization of the earier version.

Let $X_1, X_2, \ldots$ be indep r.v.s s.t.

$$E(X_i) = 0$$
$$\text{var}(X_i) < \infty$$

Set $S_n = X_1 + \cdots + X_n$.

Since the $X_i$'s form a M diff seq w.r.t. the filtration $\mathcal{F}_n := \sigma\langle X_1, \ldots, X_n \rangle$ we have

$$(S_k)_{k=1}^n \text{ is a M w.r.t. } (\mathcal{F}_k)_{k=1}^n$$

$\therefore$ $(S_k^2)_{k=1}^n$ is a subM w.r.t. $(\mathcal{F}_n)_{k=1}^n$

used $\text{var}(X_i) < \infty$ for integrability condition.

$\therefore$ Kolmogorov's Max-ineq for subM gives

$$P\left(\max_{1 \leq k \leq n} |S_k| \geq c\right) = P\left(\max_{1 \leq k \leq n} S_k^2 \geq c^2\right)$$
$$\leq \frac{E(S_n^2)^+}{c^2} = \frac{\text{var}(S_n)}{c^2}$$

A similar method generalizes Hoeffding's inequality.

**Theorem:** (Azuma)

**Theorem:** (McDiarmid's concentration ineq)