Conditional Expectation and Martingales*

T. Bucci†, G. Corbo‡

January 8, 2022

1 Conditional Expectation

$(\Omega, \mathcal{F}, \mathbb{P})$ probability space.
$\mathcal{G} \subseteq \mathcal{F}$ $\mathcal{G}$ sub-σ-algebra of $\mathcal{F}$.
$X : \Omega \to \mathbb{R}$ random variable with $\mathbb{E}[|X|] < +\infty$.

Definition 1.1. We call conditional expectatiom of X w.r.t. $\mathcal{G}$ any $\mathcal{G}$-measurable random variable V s.t.:

$$\int_G X \mathrm{d}\mathbb{P} = \int_G V \mathrm{d}\mathbb{P} \qquad \forall G \in \mathcal{G} \tag{1}$$

Notation. $V = \mathbb{E}[X|\mathcal{G}]$.

Example 1.1. $X_1, \dots, X_n$ $B(1, p)$ independent $S_n = X_1 + \cdots + X_n$
$\mathcal{G} = \sigma(S_n)$ σ-algebra generated by the events $\{S_n = k\}$ $k = 0, \dots, n$

$$\forall i \in 1, \dots, n \quad \mathbb{E}[X_i|\sigma(S_n)] = \frac{S_n}{n} \left(\text{another notation } \mathbb{E}[X_i|S_n = k] = \frac{k}{n}\right)$$

Proof. Any set in $\mathcal{G}$ is union of $S_n = k$ for some k's. Therefore it suffices to show that:

$$\int_{S_n=k} X_i \mathrm{d}\mathbb{P} = \int_{S_n=k} \frac{S_n}{n} \mathrm{d}\mathbb{P} \qquad \forall k = 0, \dots, n$$

X_i has values in $\{0, 1\}$ so

$$\int_{\{S_n=k\}} \mathbb{1}_{\{X_i=1\}} \mathrm{d}\mathbb{P} = \frac{k}{n} \cdot \mathbb{P}\{S_n = k\} \qquad \forall k = 0, \dots, n$$

We call $\hat{S}_n = X_1 + \cdots + X_{i-1} + X_{i+1} + \cdots + X_n$, then:

$$\begin{aligned}
\int_{\{S_n=k\}} \mathbb{1}_{X_i} \mathrm{d}\mathbb{P} &= \mathbb{P}\{S_n = k, X_i = 1\} \\
&= \mathbb{P}\{\hat{S}_n = k - 1, X_i = 1\} \\
&= \mathbb{P}\{\hat{S}_n = k - 1\}\mathbb{P}\{X_i = 1\} \\
&= \binom{n-1}{k-1} p^{k-1}(1-p)^{n-k} p \qquad \forall k = 0, \dots, n
\end{aligned}$$

*Taken during the course of Stochastic Dynamical Models by Professor Franco Fagnola, Politecnico di Milano, A.Y. 2021-22.

†teo.bucci@mail.polimi.it
‡gabriele.corbo@mail.polimi.it

We have to check that

$$\binom{n-1}{k-1}p^k(1-p)^{n-k} = \frac{k}{n}\binom{n}{k}p^k(1-p)^{n-k}\left(= \frac{k}{n}\mathbb{P}\{S_n = k\}\right)$$

indeed

$$\frac{k}{n}\binom{n}{k} = \frac{k}{n}\frac{n!}{k!(n-k)!} = \binom{n-1}{k-1}.$$

Then

$$\int_{S_n=k} X_i \mathrm{d}\mathbb{P} = \int_{S_n=k} \frac{S_n}{n}\mathrm{d}\mathbb{P} \qquad \forall k = 0, \ldots, n$$

So

$$\mathbb{E}[X_i|\sigma(S_n)] = \frac{S_n}{n}$$

□

Example 1.2. X, Y real random variable with joint distribution $f_{X,Y}$ and conditional distribution:

$$f_{Y|X}(y|x) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_X(x)} & \text{if } f_X(x) > 0 \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\mathbb{E}[Y|\sigma(X)] = \int y \cdot f_{Y|X}(y|X)\mathrm{d}y$$
$$\text{r.v.} \quad \omega \to \int y \cdot f_{Y|X}(y|X(\omega))\mathrm{d}y$$

Theorem 1.1. Given X with $\mathbb{E}[|X|] < +\infty$ and $\mathcal{G}$ sub-σ-algebra of $\mathcal{F}$, one can always find $\mathbb{E}[X|\mathcal{G}]$.

Proof. Suppose $X \geq 0$ (if not we write $X = X^+ - X^-$ and work on the two parts separately).
Consider the following measures on $\mathcal{G}$:

- $\mathbb{P}_{|\mathcal{G}}$ ($\mathbb{P}$ is defined on $\mathcal{F}$).
- $\mathbb{Q}(G) = \int_G X\mathrm{d}\mathbb{P}$.

Note that $\mathbb{Q}$ is absolutely continuous w.r.t $\mathbb{P}_{|\mathcal{G}}$ ($\mathbb{Q} \ll \mathbb{P}_{|\mathcal{G}}$).
Then, by Radon-Nykodim theorem, there exists the density V of $\mathbb{Q}$ w.r.t. $\mathbb{P}_{|\mathcal{G}}$ $\left(V = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}_{|\mathcal{G}}}\right)$.
V is $\mathcal{G}$-measurable because $\mathbb{Q}$ and $\mathbb{P}_{|\mathcal{G}}$ are measures on $\mathcal{G}$.
Then:

$$\mathbb{Q}(G) = \int_G X\mathrm{d}\mathbb{P} = \int_G V\mathrm{d}\mathbb{P}$$

□

Example 1.3. $(X_n)_{n>0}$ Markov chain with state space I and transition matrix P.
$f : I \to \mathbb{R} \quad \text{s.t.} \quad \mathbb{E}[|f(X_n)|] < +\infty \qquad \forall n \in \mathbb{N}$.
Then:

$$\mathbb{E}[f(X_{n+1})|\sigma(X_1, \dots, X_n)] = (Pf)(X_n)$$

Where $(Pf)(j) = \sum_{j' \in I} p_{jj'} f(j')$.

Proof. We want to show that $\forall G \in \sigma(X_1, \dots, X_n)$

$$\int_G f(X_{n+1}) \mathrm{d}\mathbb{P} = \int_G (Pf)(X_n) \mathrm{d}\mathbb{P}$$

Each G is the countable union of a set $X_1 = i_1, \dots, X_n = i_n$, therefore we must show that:

$$\int_{\{X_1=i_1,\dots,X_n=i_n\}} f(X_{n+1}) \mathrm{d}\mathbb{P} = \int_{\{X_1=i_1,\dots,X_n=i_n\}} (Pf)(X_n) \mathrm{d}\mathbb{P}$$

The LHS can be written as follows:

$$\begin{aligned}
\int_{\{X_1=i_1,\dots,X_n=i_n\}} f(X_{n+1}) \mathrm{d}\mathbb{P} &= \sum_{j \in I} \int_{\{X_1=i_1,\dots,X_n=i_n\}} f(j) \mathbb{1}_{\{X_{n+1}=j\}} \mathrm{d}\mathbb{P} \\
&= \sum_{j \in I} f(j) \mathbb{P}\{X_{n+1} = j, X_n = i_n, \dots, X_1 = i_1\} \\
&= \sum_{j \in I} f(j) \mathbb{P}\{X_{n+1} = j | X_n = i_n, \cancel{\dots}\} \mathbb{P}\{X_n = i_n, \dots, X_1 = i_1\} \\
&= \sum_{j \in I} f(j) p_{i_n j} \int_{\{X_1=i_1,\dots,X_n=i_n\}} \mathrm{d}\mathbb{P} \\
&= \int_{\{X_1=i_1,\dots,X_n=i_n\}} \sum_{j \in I} f(j) p_{X_n j} \mathrm{d}\mathbb{P} \\
&= \int_{\{X_1=i_1,\dots,X_n=i_n\}} (Pf)(X_n) \mathrm{d}\mathbb{P}
\end{aligned}$$

$\square$

Remark (for CTMC)**.**

$$s < t \qquad \mathbb{E}[f(X_t)|\sigma(X_r | r \le s)] = (P_{t-s} f)(X_s)$$

1.1 Properties of $\mathbb{E}[\,\cdot\,|\mathcal{G}]$

1. **(linear)** $\mathbb{E}[\alpha X + \beta Y|\mathcal{G}] = \alpha \cdot \mathbb{E}[X|\mathcal{G}] + \beta \cdot \mathbb{E}[Y|\mathcal{G}]$
2. **(positive)** $X \ge 0 \implies \mathbb{E}[X|\mathcal{G}] \ge 0$
3. **(normalized)** $\mathbb{E}[1|\mathcal{G}] = 1$
4. **Projective Property**

   Theorem 1.2. If $\mathcal{H}$ is a sub-σ-algebra of $\mathcal{G}$ then:

   $$\mathbb{E}[\,\mathbb{E}[X|\mathcal{G}]\,|\,\mathcal{H}\,] = \mathbb{E}[X|\mathcal{H}]$$

Proof. We must check that $\forall H \in \mathcal{H}$

$$\int_H \mathbb{E}[\,\mathbb{E}[X|\mathcal{G}]\,|\,\mathcal{H}\,]\mathrm{d}\mathbb{P} = \int_H \mathbb{E}[X|\mathcal{H}]\mathrm{d}\mathbb{P}$$

Since $H \in \mathcal{H}$

$$\int_H \mathbb{E}[X|\mathcal{H}]\mathrm{d}\mathbb{P} = \int_H X\mathrm{d}\mathbb{P}$$

and since $\mathcal{H} \subseteq \mathcal{G}$, $H \in \mathcal{G}$, so:

$$\int_H \mathbb{E}[\,\mathbb{E}[X|\mathcal{G}]\,|\,\mathcal{H}\,]\mathrm{d}\mathbb{P} = \int_H \mathbb{E}[X|\mathcal{G}]\mathrm{d}\mathbb{P} = \int_H X\mathrm{d}\mathbb{P}$$

□

Definition 1.2. X real r.v. with $\mathbb{E}[|X|] < +\infty$, $\mathcal{G}$ sub-σ-algebra of $\mathcal{F}$

$$X \perp\!\!\!\perp \mathcal{G} \iff \mathbb{P}\left(\{X \in B\} \cap G\right) = \mathbb{P}\{X \in B\} \cdot \mathbb{P}\{G\} \qquad \forall B \in \mathcal{B}(\mathbb{R}) \quad \forall G \in \mathcal{G}$$

Remark. If $\mathcal{G} = \sigma(Y)$ the definition matches with $X \perp\!\!\!\perp Y$ becaus every $G \in \mathcal{G}$ is in the form $\{Y \in B'\}$ with $B' \in \mathcal{B}(\mathbb{R})$.

Theorem 1.3. $X \perp\!\!\!\perp \mathcal{G} \implies \mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$

Proof. The constant r.v. $\mathbb{E}[X]$ is $\mathcal{G}$-measurable $\forall G \in \mathcal{G}$
Note that $\forall B \in \mathcal{B}(\mathbb{R})$

$$\int_G \mathbb{1}_B(X)\mathrm{d}\mathbb{P} = \mathbb{P}\left(\{X \in B\} \cap G\right) = \mathbb{P}\{X \in B\} \cdot \mathbb{P}\{G\} =$$
$$= \int_G \mathbb{P}\{X \in B\}\mathrm{d}\mathbb{P} = \int_G \mathbb{E}[\mathbb{1}_B(X)]\mathrm{d}\mathbb{P}$$

Then $\forall f : \mathbb{R} \to \mathbb{R}$ measurable and bounded

$$\int_G f(X)\mathrm{d}\mathbb{P} = \int_G \mathbb{E}[f(X)]\mathrm{d}\mathbb{P}$$

Hence we can construct a sequence of boundend functions to approximate the identity $x \to x$, defined as:

$$f_n(x) = \max\{-n, \min\{x, n\}\}$$

s.t.:

$$\int_G f_n(X)\mathrm{d}\mathbb{P} = \int_G \mathbb{E}[f_n(X)]\mathrm{d}\mathbb{P}$$

Finally, applying the monotone convergence theorem as $n \to \infty$

$$\int_G X\mathrm{d}\mathbb{P} = \int_G \mathbb{E}[X]\mathrm{d}\mathbb{P}$$

□

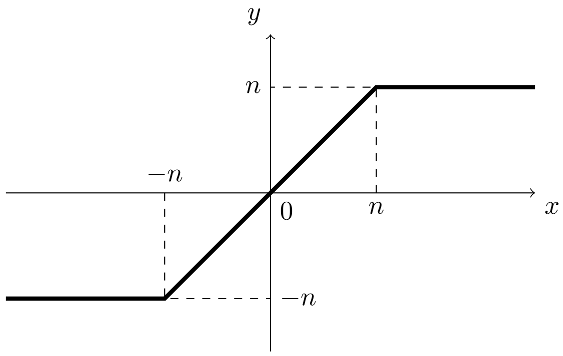

Figure 1: Approximation of the identity map.

5. **Contractivity in** $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$, $p \geq 1$

$$X \in \mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P}) \implies \mathbb{E}[X|\mathcal{G}] \in \mathcal{L}^p(\Omega, \mathcal{G}, \mathbb{P})$$

Indeed, is a consequence of the Jensen's inequality: If $\Phi : \mathbb{R} \to \mathbb{R}$ is convex and $\mathbb{E}[|\Phi(X)|] < +\infty$, then:

$$\Phi(\mathbb{E}[X]) \leq \mathbb{E}[\Phi(X)]$$

(that also holds for conditional expectations $\Phi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\Phi(X)|\mathcal{G}]$). In our case $\Phi(X) = |X|^p$.

Theorem 1.4. X r.v. with $\mathbb{E}[|X|^p] < +\infty$
V $\mathcal{G}$-measurable random variable with $\mathbb{E}[|V|^q] < +\infty$ with $\frac{1}{p} + \frac{1}{q} = 1$ conjugate index. Then

$$\mathbb{E}[V \cdot X|\mathcal{G}] = V \cdot \mathbb{E}[X|\mathcal{G}]$$

Proof. V indicator function $V = \mathbb{1}_G \quad G \in \mathcal{G}$ (V $\mathcal{G}$-measurable), then, by definition of conditional expectation, $\forall G' \in \mathcal{G}$:

$$\begin{aligned}\int_{G'} \mathbb{E}[\mathbb{1}_G \cdot X|\mathcal{G}]\mathrm{d}\mathbb{P} &= \int_{G'} \mathbb{1}_G \cdot X \mathrm{d}\mathbb{P} = \int_{G' \cap G} X \mathrm{d}\mathbb{P} = \\ &= \int_{G' \cap G} \mathbb{E}[X|\mathcal{G}]\mathrm{d}\mathbb{P} = \int_{G'} \mathbb{1}_G \cdot \mathbb{E}[X|\mathcal{G}]\mathrm{d}\mathbb{P}\end{aligned}$$

by the arbitrarity of G':

$$\mathbb{E}[\mathbb{1}_G \cdot X|\mathcal{G}] = \mathbb{1}_G \cdot \mathbb{E}[X|\mathcal{G}]$$

by linearity:

$$\mathbb{E}[V \cdot X|\mathcal{G}] = V \cdot \mathbb{E}[X|\mathcal{G}] \qquad \forall V \text{simple and } \mathcal{G}\text{-measurable}$$

If $V > 0$, $V = \sup_{n \geq 1} V_n$ with V_n simple $\mathcal{G}$-measurable, so:

$$\mathbb{E}[V \cdot X|\mathcal{G}] = \sup_{n \geq 1} \mathbb{E}[V_n \cdot X|\mathcal{G}] = \sup_{n \geq 1} V_n \cdot \mathbb{E}[X|\mathcal{G}] = V \cdot \mathbb{E}[X|\mathcal{G}]$$

If V is real we can write $V = V^+ - V^-$ and the thesis follows. $\square$

Theorem 1.5. If X r.r.v. with $\mathbb{E}[|X|^2] < +\infty$, then:
Z $\mathcal{G}$-measurable random variable with $\mathbb{E}[|Z|^2] < +\infty$

$$\mathbb{E}[|X - Z|^2] = \mathbb{E}[\ |X - \mathbb{E}[X|\mathcal{G}]|^2\ ] + \mathbb{E}[\ |\mathbb{E}[X|\mathcal{G}] - Z|^2\ ]$$

In particular:

$$\mathbb{E}[|X - \mathbb{E}[X|\mathcal{G}]|^2] = \min_{Z \in \mathcal{L}^2(\Omega,\mathcal{F},\mathbb{P})} \mathbb{E}[|X - Z|^2]$$

Proof.

$$\begin{aligned}\mathbb{E}[|X - Z|^2] &= \mathbb{E}[\ [(X - \mathbb{E}[X|\mathcal{G}]) + (\mathbb{E}[X|\mathcal{G}] - Z)]^2\ ] \\ &= \mathbb{E}[\ |X - \mathbb{E}[X|\mathcal{G}]|^2\ ] + \mathbb{E}[\ |\mathbb{E}[X|\mathcal{G}] - Z|^2\ ] \\ &\quad + 2 \cdot \mathbb{E}[\ [(X - \mathbb{E}[X|\mathcal{G}])(\mathbb{E}[X|\mathcal{G}] - Z)]\ ]\end{aligned}$$

We only need to prove that the last term is zero. By Projective property:

$\mathbb{E}[\ [(X - \mathbb{E}[X|\mathcal{G}])(\mathbb{E}[X|\mathcal{G}] - Z)]\ ] = \mathbb{E}[\mathbb{E}[\ [(X - \mathbb{E}[X|\mathcal{G}])(\mathbb{E}[X|\mathcal{G}] - Z)]\ ]|\mathcal{G}]$

$\mathbb{E}[X|\mathcal{G}] - Z$ is $\mathcal{G}$-measurable, so:

$\mathbb{E}[\mathbb{E}[\ [(X-\mathbb{E}[X|\mathcal{G}])(\mathbb{E}[X|\mathcal{G}]-Z)]\ ]|\mathcal{G}] = \mathbb{E}[(\mathbb{E}[X|\mathcal{G}]-Z)\cdot\mathbb{E}[\ [(X-\mathbb{E}[X|\mathcal{G}])]\ ]|\mathcal{G}]$

Finally:

$$\begin{aligned}\mathbb{E}[\ [(X - \mathbb{E}[X|\mathcal{G}])]\ |\mathcal{G}] &= \mathbb{E}[X|\mathcal{G}] - \mathbb{E}[\ [(\mathbb{E}[X|\mathcal{G}])]\ |\mathcal{G}] \\ &= \mathbb{E}[X|\mathcal{G}] - \mathbb{E}[X|\mathcal{G}] = 0\end{aligned}$$

□

2 Martingales

2.1 Introduction

$(\Omega, \mathcal{F}, \mathbb{P})$ T set of times either $\mathbb{N}$ or $[0, \infty)$.

Definition 2.1. A **filtration** on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a sequence $\mathcal{F}_t : t = 0, 1, 2, \ldots$ of sub-sigma algebras of $\mathcal{F}$, such that for all t, $\mathcal{F}_t \subseteq \mathcal{F}_{t+1}$.

The filtration represent our knowledge at successive times. This increases with time (i.e. we don't forget things).

Definition 2.2. $(M_t)_{t\in T}$ is a Martingale w.r.t. a filtration $(\mathcal{F}_t)_{t\in T}$ of sub-σ-algebras of $\mathcal{F}$ if:

1. **(adaptedness)** M_t is $\mathcal{F}_t$-measurable $\forall t \in T$
2. **(integrability)** $\mathbb{E}[|M_t|] < +\infty \qquad \forall t$
3. **(martinagle property)** $\forall s < t \qquad \mathbb{E}[M_t|\mathcal{F}_s] = M_s$

Example 2.1. $(X_n)_{n\geq 0}$ symmetric RW on $\mathbb{Z}$ starting from 0.
$X_n = \sum_{k=1}^n Y_k \qquad Y_k$ iid with $\mathbb{P}\{Y_k = 1\} = \frac{1}{2} = \mathbb{P}\{Y_k = -1\}$
$(X_n)_{n\geq 0}$ is a martingale w.r.t. $\{\mathcal{F}_n\}_{n\geq 0}$ with $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$.

Proof. We check the definition.

1. X_n is $\mathcal{F}_n$-measurable.
2. X_n is integrable ($|X_n| \le n$).
3. If $m < n$,
$$\mathbb{E}[X_n|\mathcal{F}_m] = \mathbb{E}\left[\sum_{k=m+1}^{n} Y_k + \sum_{k=1}^{m} Y_k|\mathcal{F}_m\right]$$
The first sum is independent of $\mathcal{F}_m$, the second one is $\mathcal{F}_n$-measurable, so:
$$\mathbb{E}[X_n|\mathcal{F}_m] = \mathbb{E}\left[\sum_{k=m+1}^{n} Y_k\right] + \sum_{k=1}^{m} Y_k = \sum_{k=m+1}^{n} \mathbb{E}[Y_k] + X_m = X_m$$
Since $\mathbb{E}[Y_k] = 0$.

$\square$

Example 2.2 (Another martingale)**.** $M_n = X_n^2 - n \qquad (M_n)_{n\ge 0}$

1. M_n is $\mathcal{F}_n$-measurable.
2. $|M_n| \le |X_n^2 - n| \le n^2 + n < +\infty \implies \mathbb{E}[|M_n|] < +\infty$.
3.

$$\begin{aligned}
\mathbb{E}[X_n|\mathcal{F}_m] &= \mathbb{E}[X_n^2 - n|\mathcal{F}_m] \\
&= \mathbb{E}[X_n^2|\mathcal{F}_m] - \mathbb{E}[n|\mathcal{F}_m] = \\
&= \mathbb{E}\left[\left(\sum_{k=1}^{m} Y_k + \sum_{k=m+1}^{n} Y_k\right)^2 |\mathcal{F}_m\right] - n = \\
&= \mathbb{E}\left[\left(\sum_{k=1}^{m} Y_k\right)^2 |\mathcal{F}_m\right] + 2\cdot\mathbb{E}\left[\left(\sum_{k=1}^{m} Y_k\right)\cdot\left(\sum_{k=m+1}^{n} Y_k\right)|\mathcal{F}_m\right] \\
&\quad + \mathbb{E}\left[\left(\sum_{k=m+1}^{n} Y_k\right)^2 |\mathcal{F}_m\right] - n = \\
&= X_m^2 + 2\cdot X_m\cdot\mathbb{E}\left[\sum_{k=m+1}^{n} Y_k|\mathcal{F}_m\right] + \mathbb{E}\left[\left(\sum_{k=m+1}^{n} Y_k\right)^2\right] - n = \\
&= X_m^2 + 2\cdot X_m\cdot 0 + \sum_{k=m+1}^{n} \mathbb{E}[Y_k^2] - n = \\
&= X_m^2 + (n-m) - n = X_m^2 - m = M_m
\end{aligned}$$

Since $\mathbb{E}[Y_k] = 0$ and $\mathbb{E}[Y_k^2] = 1$.

2.2 Stopping time

T stopping time of the filtration $(\mathcal{F}_t)_{t\in T}$ (i.e. $\{T \leq t\} \in \mathcal{F}_t \qquad \forall t \geq 0$)
We call **stopped martingale at time** T: $M^{IT} = (M_{\min(T,t)})_{t\geq 0}$.

Theorem 2.1 (Stopping Theorem). $(M_{\min(T,t)})_{t\geq 0}$ is a martingale w.r.t. the same filtration $(\mathcal{F}_t)_{t\in T}$ and in particular:

$$\mathbb{E}[M_{\min(T,t)}] = \mathbb{E}[M_0] \tag{2}$$

Remark. A martingale $(M_t)_{t\in T}$ necesseraly has constant means by:

$$\mathbb{E}[M_T] = \mathbb{E}[\mathbb{E}[M_T|\mathcal{F}_s]] = \mathbb{E}[M_0] \qquad \forall s < t$$

Example 2.3. $(X_n)_{n\geq 0}$ symmetric RW on $\mathbb{Z}$ starting from 0 and $a, b \in \mathbb{N}$ or $a, b > 0$.
$X_n = \sum_{k=1}^n Y_k \qquad Y_k$ iid with $\mathbb{P}\{Y_k = 1\} = \frac{1}{2} = \mathbb{P}\{Y_k = -1\}$
We want to compute the exit probability from $(-a, b)$ either from $-a$ or b and the mean exit time $T = \min(n \geq 1 | X_n = -a \text{ or } X_n = b)$.
Consider the martingales $(X_n)_{n\geq 0}$ and $(X_n^2 - n)_{n\geq 0}$ and apply the stopping theorem:

$$\mathbb{E}[X_{T\wedge n}^2 - T \wedge n] = \mathbb{E}[0] = 0$$
$$\mathbb{E}[X_{T\wedge n}^2] = \mathbb{E}[T \wedge n] \tag{3}$$

Thus $\mathbb{E}[T \wedge n] \leq \max(a^2, b^2)$ because $|X_{T\wedge n}^2| \leq \max(a^2, b^2)$, thus we have: $\mathbb{E}[T] = \sup_{n\geq 1} \mathbb{E}[T \wedge n]$
Applying the stopping theorem to $(X_n)_{n\geq 0}$:

$$\mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_0] = 0$$

$$\mathbb{E}[X_{T\wedge n}] \leq \max(a, b)$$

Using dominated convergence:

$$\mathbb{E}[X_T] = \lim_{n\to\infty} \mathbb{E}[X_{T\wedge n}] = 0$$

$$0 = -a \cdot \mathbb{P}\{X_T = -a\} + b \cdot \mathbb{P}\{X_T = b\}$$

Recalling that:

$$\mathbb{P}\{X_T = -a\} + \mathbb{P}\{X_T = b\} = 1$$

Solving the system:

$$\mathbb{P}\{X_T = -a\} = \frac{b}{a+b} \qquad \mathbb{P}\{X_T = b\} = \frac{a}{a+b}$$

Then using monotone convergence on the RHS on (3) and dominated convergence on the LHS as $n \to \infty$ we obtain:

$$\begin{aligned} \mathbb{E}[T] = \mathbb{E}[X_T^2] &= a^2 \cdot \mathbb{P}\{X_T = -a\} + b^2 \cdot \mathbb{P}\{X_T = b\} = \\ &= \frac{a^2 b}{a+b} + \frac{ab^2}{a+b} = ab \end{aligned}$$

Remark. Note the resemblance of this result with the mean ending time starting from i in the gambler's ruin problem:

$$\mathbb{E}_i[T] = i(N-i)$$

Remark (Warning)**.** It is not always possible to take the limit $n \to \infty$ to find $\mathbb{E}[M_{\min(T,t)}] = \mathbb{E}[M_t]$

Example 2.4. $(X_n)_{n\geq 0}$ symmetric RW on $\mathbb{Z}$ starting from 0.
$S = \min\{n \geq 1 : X_n = 1\}$
Stopping theorem $\mathbb{E}[X_{\min(S,n)}] = 0$ but $\mathbb{E}[X_S] = 1$, therefore:

$$\lim_{n\to\infty} \mathbb{E}[X_{\min(S,n)}] \neq \mathbb{E}[X_S]$$

How to find martingales of a discrete time MC?

Theorem 2.2. $(X_n)_{n\geq 0}$ MC with values in a discrete set I
$f : I \to \mathbb{R} \quad$ s.t. $\mathbb{E}[|f(X_n)|] < +\infty$, then:

$$M_n = f(X_n) - \sum_{k=0}^{\mathbf{n\text{-}1}} (Pf(X_k) - f(X_k)) \tag{4}$$

$(M_n)_{n\geq 0}$ is a martingale w.r.t. the natural filtration of $(X_n)_{n\geq 0}$. Where $(Pf)(j) = \sum_{j'\in I} p_{jj'} f(j')$.

Proof. We check the definition:

1. M_n is $\mathcal{F}_n$-measurable.

2. M_n is integrable.
   $\mathbb{E}[|f(X_n)|] < +\infty$ by assumption $\forall n$.
   $\mathbb{E}[|Pf(X_k)|] = \mathbb{E}[|\mathbb{E}[f(X_{k+1})|\mathcal{F}_k]|] < +\infty$ The equality holds since:

   $$\mathbb{E}[f(X_{k+1})|\mathcal{F}_k] = Pf(X_k)$$

   indeed for $m < n$:

   $$\mathbb{E}[f(X_n)|\mathcal{F}_m] = P^{n-m} f(X_m)$$

   because:

   $$\begin{aligned}\mathbb{E}[f(X_n)|\mathcal{F}_m] &= \mathbb{E}[\mathbb{E}[f(X_n)|\mathcal{F}_{n-1}]|\mathcal{F}_m] \\ &= \mathbb{E}[Pf(X_{n-1})|\mathcal{F}_m] \\ &= \mathbb{E}[P^2 f(X_{n-2})|\mathcal{F}_m] = \ldots\end{aligned}$$

3. For $n > m$:

   $$\begin{aligned}\mathbb{E}[M_n|\mathcal{F}_m] &= \mathbb{E}\left[f(X_n) - \sum_{k=0}^{n-1}(Pf(X_k) - f(X_k))\middle|\mathcal{F}_m\right] \\ &= \mathbb{E}\left[f(X_n) - f(X_m) - \sum_{k=m}^{n-1}(Pf(X_k) - f(X_k))\middle|\mathcal{F}_m\right] \\ &\quad \underbrace{+\, f(X_m) - \sum_{k=0}^{m-1}(Pf(X_k) - f(X_k))}_{M_m} \\ &= (*) + M_m\end{aligned}$$

The last term is M_m so it suffices to show that the first term is zero. It is equal to:

$$
\begin{aligned}
(*) &= \mathbb{E}\left[\mathbb{E}\left[f(X_n) - f(X_m) - \sum_{k=m}^{n-1}(Pf(X_k) - f(X_k))\middle|\mathcal{F}_{n-1}\right]\middle|\mathcal{F}_m\right] = \\
&= \mathbb{E}\left[\cancel{Pf(X_{n-1})} - f(X_m) - \cancel{Pf(X_{n-1})} + f(X_{n-1}) - \sum_{k=m}^{n-2}(\dots)\middle|\mathcal{F}_m\right] = \\
&= \mathbb{E}\left[f(X_{n-1}) - f(X_m) - \sum_{k=m}^{n-2}(Pf(X_k) - f(X_k))\middle|\mathcal{F}_m\right] = \\
&= \text{iterating...} = \mathbb{E}[f(X_m) - f(X_m)] = 0
\end{aligned}
$$

□

Example 2.5. $(X_n)_{n\geq 0}$ symmetric RW on $\mathbb{Z}$ starting from 0.
$f(x) = x^2$.
$\left(M_n = f(X_n) - \sum_{k=1}^{n-1}(Pf(X_k) - f(X_k))\right)$

$$Pf(i) = p_{i,i+1}f(i+1) + p_{i,i-1}f(i-1) = \frac{1}{2}(i+1)^2 + \frac{1}{2}(i-1)^2 = i^2 + 1$$

$$Pf(i) - f(i) = 1$$

$$\sum_{k=0}^{n-1} Pf(i) - f(i) = n \implies M_n = X_n^2 - n$$

Remark. If you add a constant to a martingale it remains a martingale.

Theorem 2.3 (Martingale stopping theorem)**.** Let $\mathbb{T}$ be a set of times, $(M_t)_{t\in\mathbb{T}}$ a Martingale w.r.t. a filtration $(\mathcal{F}_t)_{t\in\mathbb{T}}$ and T a stopping time w.r.t. $(\mathcal{F}_t)_{t\in\mathbb{T}}$. Then $(M_{T\wedge t})_{t\in\mathbb{T}}$ is still a Martingale w.r.t. the same filtration $(\mathcal{F}_t)_{t\in\mathbb{T}}$. In particular:

$$\mathbb{E}[M_{T\wedge t}] = \mathbb{E}[M_0]$$

Proof. (case with T *discrete* stopping time[1]) We have to check the three properties.

1. (adaptedness) i.e. $\forall t$, $M_{T\wedge t}$ is $\mathcal{F}_t$-measurable.

$$T \text{ discrete} \quad \Rightarrow \quad T = \sum_{r\in D} r\mathbb{1}_{\{T\leq r\}}$$

$\{T = r\} \in \mathcal{F}_r$ because T is a stopping time.[2] We can decompose the RV, one up to time t and one after t:

$$M_{T\wedge t} = \sum_{r\in D\cap[0,t]} \underbrace{M_r\mathbb{1}_{\{T=r\}}}_{\substack{\mathcal{F}_r\text{-measurable} \\ r\leq t \\ \Rightarrow \mathcal{F}_t\text{-measurable}}} + \underbrace{M_t}_{\mathcal{F}_t\text{-measurable}} \underbrace{\mathbb{1}_{\{T>t\}}}_{\{T\leq t\}^C\in\mathcal{F}_t}$$

[1] Otherwise approximation stuff...
[2] $\{T = r\} = \{T \leqslant r\} \setminus \bigcup_{n\geqslant 1}\left\{t \leqslant r - \frac{1}{n}\right\}$. We have to use inequalities.

as a result, $M_{T\wedge t}$ is $\mathcal{F}_t$-measurable.

2. (integrability) i.e. $\mathbb{E}[|M_{T\wedge t}|] < +\infty$. We can use again the decomposition. Moreover, since D is discrete, this implies that $D \cap [0,t]$ is *finite* for any t. Thus:

$$|M_{T\wedge t}| \leq \sum_{r \in D\cap[0,t]} |M_r| \mathbb{1}_{\{T=r\}} + |M_t| \mathbb{1}_{\{T<t\}}$$

is a finite sum of integrable RVs.

3. $\mathbb{E}[M_{T\wedge t}|\mathcal{F}_s] = M_{T\wedge s}$ for any $s < t$, i.e. we have to check that

$$\forall A \in \mathcal{F}_s : \quad \int_A M_{T\wedge s} d\mathbb{P} = \int_A M_{T\wedge t} d\mathbb{P}$$

Notice that both $A \cap \{T \leq s\}$ and $A \cap \{T > s\}$ belong to $\mathcal{F}_s$.

Computing:

$$\begin{aligned}
\int_A M_{T\wedge s} d\mathbb{P} &= \int_{A\cap\{T\leq s\}} M_T d\mathbb{P} + \int_{A\cap\{T>s\}} M_s d\mathbb{P} \\
&= \int_{A\cap\{T\leq s\}} M_T d\mathbb{P} + \int_{A\cap\{T>s\}} M_t d\mathbb{P} \\
&= \int_{A\cap\{T\leq s\}} M_T d\mathbb{P} + \underbrace{\int_{A\cap\{s<T\leq t\}} M_t d\mathbb{P}}_{(*)} + \int_{A\cap\{T>t\}} M_t d\mathbb{P}
\end{aligned}$$

where in the second step we used the Martingale property of $(M_t)_{t\in\mathbb{T}}$ and in the third we split the set. Notice that

$$(*) = \int_{A\cap\{s<T\leq t\}} M_t d\mathbb{P} = \sum_{r\in D\cap(s,t]} \int_{A\cap\{T=r\}} M_t d\mathbb{P}$$

recall that $A \in \mathcal{F}_s$, $\{T = r\} \in \mathcal{F}_r$, $s < r$ then $A \cap \{T = r\} \in \mathcal{F}_r$ (the biggest). So using again the Martingale property for the given Martingale the integral in the sum can be written as

$$\int_{A\cap\{T=r\}} M_t d\mathbb{P} = \int_{A\cap\{T=r\}} M_r d\mathbb{P}$$

Summing up

$$\begin{aligned}
\int_A M_{T\wedge s} d\mathbb{P} &= \int_{A\cap\{T\le s\}} M_T d\mathbb{P} + \sum_{r\in D\cap(s,t]} \int_{A\cap\{T=r\}} M_r d\mathbb{P} + \int_{A\cap\{T>t\}} M_t d\mathbb{P} \\
&= \int_{A\cap\{T\le s\}} M_T d\mathbb{P} + \sum_{r\in D\cap(s,t]} \int_{A\cap\{T=r\}} M_T d\mathbb{P} + \int_{A\cap\{T>t\}} M_{T\wedge t} d\mathbb{P} \\
&= \int_{A\cap\{T\le s\}} M_{T\wedge t} d\mathbb{P} + \int_{A\cap\{s<T\le t\}} M_{T\wedge t} d\mathbb{P} + \int_{A\cap\{T>t\}} M_{T\wedge t} d\mathbb{P} \\
&= \int_A M_{T\wedge t} d\mathbb{P}
\end{aligned}$$

□

Martingales of a continuous time Markov Chain

Theorem 2.4. $(X_t)_{t\ge 0}$ continuous time MC (regular..., FKE/BKE have a unique solution...). Consider $f : I \to \mathbb{R}$ such that $\mathbb{E}[|f(X_t)| < +\infty$, Q transition rate matrix, and $\mathbb{E}[|Qf(X_s)|] < +\infty$ $\forall s$, then defining

$$M_t := f(X_t) - \int_0^t (Qf)(X_s) ds \tag{5}$$

$(M_t)_{t\ge 0}$ is a Martingale w.r.t. the natural filtration of the given MC $(X_t)_{t\ge 0}$, namely the collection $(\mathcal{F}_t)_{t\ge 0}$ of $\mathcal{F}_t := \sigma(X_s | s \le t)$.

Remark. Compare (5) with the discrete-time analog:

$$f(X_n) - \sum_{k=0}^{n-1} (Pf(X_k) - f(X_k))$$

Proof. Recall that $\mathbb{E}[f(X_t)|\mathcal{F}_s] = (P_{t-s}f)(X_s)$ where $P_{t-s} = (p_{ij}(t-s))_{i,j\in I}$. We check the Martingale property:

$$\mathbb{E}[M_t|\mathcal{F}_s] \stackrel{?}{=} M_s$$

We begin the computation

$$\mathbb{E}[M_t|\mathcal{F}_s] = \mathbb{E}\left[f(X_t) - \int_0^t Qf(X_r)dr|\mathcal{F}_s\right] \stackrel{?}{=} f(X_s) - \int_0^s Qf(X_r)dr = M_s$$

$\int_0^t = \int_0^s + \int_s^t$

$$\mathbb{E}\left[f(X_t) - f(X_s) - \int_s^t Qf(X_r)dr|\mathcal{F}_s\right] \stackrel{?}{=} 0$$

$$\mathbb{E}[f(X_t) - f(X_s)|\mathcal{F}_s] \stackrel{?}{=} \mathbb{E}\left[\int_s^t Qf(X_r)dr|\mathcal{F}_s\right] = (*)$$

then since the conditional expectation is essentially an integral, by Fubini's theorem and remembering that $(P_{r-s}Qf)(j) = \sum_{jk} p_{ij}(r-s)q_{jk}f(k)$, we have the FKE, $\ldots = \frac{d}{dr}\sum_k p_{ik}(r-s)f(k)$

$$\begin{aligned}
(*) &= \int_s^t \mathbb{E}[Qf(X_r)|\mathcal{F}_s]dr \\
&= \int_s^t (P_{r-s}Qf)(X_s)dr \\
&= \int_s^t \frac{d}{dr}(P_{r-s}f)(X_s)dr \\
&= (P_{r-s}f)(X_s)|_s^t \\
&= P_{t-s}f(X_s) - f(X_s) \\
&= \mathbb{E}[f(X_t)|\mathcal{F}_s] - f(X_s) \\
&= \mathbb{E}[f(X_t) - f(X_s)|\mathcal{F}_s]
\end{aligned}$$

□

Application/example.

Computation of the mean depletion time of an $M/M/1$ queue (i.e. mean extinction time of a population).

$(X_t)_{t\geq 0}$ where X_t is the number of individuals in the system, with 1 counter. Suppose the system can host at most N individuals[3] (it's not exactly a $M/M/1$ queue). The set of states is then $\{0, 1, \ldots, N\}$.

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & \cdots & 0 \\ \mu & -(\lambda+\mu) & \lambda & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \mu & -(\lambda+\mu) & \lambda \\ 0 & \cdots & 0 & 0 & \mu & -\mu \end{bmatrix}, \quad \lambda < \mu$$

We consider as initial state n, where $n \ll N$, and define

$$T = \min\{t \geq 0 | X_t = 0\},$$

we want to compute $\mathbb{E}_n[T]$ using the Martingales theorem.

Consider $f(n) = n$:

$$\begin{aligned}
\text{if } n > 0, \quad Qf(n) &= \mu f(n-1) - (\lambda+\mu)f(n) + \lambda f(n+1) \\
&= \mu(n-1) + (\lambda+\mu)n + \lambda(n+1) \\
&= \lambda - \mu \\
\text{if } n = 0, \quad Qf(0) &= \lambda(f(1) - f(0)) = \lambda
\end{aligned}$$

The martingale is

$$X_t - \int_0^t (\lambda \mathbb{1}_{\{X_s=0\}} + (\lambda-\mu)\mathbb{1}_{\{X_s>0\}})ds$$

Martingale stopping theorem:

$$\mathbb{E}_n\left[X_{T\wedge t} - \int_0^{T\wedge t} \left(\cancel{\lambda \mathbb{1}_{\{X_s=0\}}} + (\lambda-\mu)\mathbb{1}_{\{X_s>0\}}\right) ds\right] = n$$

[3] to avoid technical problems of integrability.

when $s < T$ you have not yet reached 0, so the first $\mathbb{1}$ is zero and the second is one.

$$\mathbb{E}_n[\underbrace{X_{T\wedge t}}_{\geq 0} + (\mu-\lambda)(T\wedge t)] = n \tag{6}$$

thus

$$(\mu-\lambda)\mathbb{E}_n[T\wedge t] \leq n \quad \Rightarrow \quad \text{when } t\to\infty: \ \mathbb{E}_n[T] \leq \frac{n}{\mu-\lambda}$$

so T is integrable. Back to (6)

$$X_{T\wedge t} \xrightarrow[\text{a.s.}]{t\to\infty} X_T = 0 \quad \Rightarrow \quad (\mu-\lambda)\mathbb{E}_n[T] = n \quad \Rightarrow \quad \mathbb{E}_n[T] = \frac{n}{\mu-\lambda}$$

Theorem 2.5 (Martingale Convergence Theorem)**.** Given $(M_t)_{t\geq 0}$ (with either $t\in\mathbb{N}$ or $t\in\mathbb{R}^+$) Martingale w.r.t. some filtration $(\mathcal{F}_t)_{t\geq 0}$. If

$$\sup_{t\geq 0}\mathbb{E}[|M_t|] < +\infty$$

then

$$\lim_{t\to+\infty} M_t$$

exists almost surely (may not be in L^1).

Example 2.6. Consider $(X_n)_{n\geq 0}$ symmetric RW on $\mathbb{Z}$.

$$X_0 = 0 \qquad X_n = \sum_{k=1}^{n} Y_k \qquad Y_k \text{ i.i.d. } \mathbb{P}(Y_k = \pm 1) = \frac{1}{2}$$

Let

$$M_n = e^{\vartheta X_n - cn}, \quad c, \vartheta \in \mathbb{R}$$

1. Find c such that $(M_n)_{n\geq 0}$ is a Martingale.
2. Show that for such c and all ϑ, M_n converges a.s., but not in L^1.

For $(M_n)_{n\geq 0}$ to be a Martingale, beside checking the two basic properties, we have to check that if $\mathbb{E}[M_{n+1}|\mathcal{F}_n] \overset{?}{=} M_n$. The first quantity is equal to:

$$\begin{aligned}
\mathbb{E}[M_{n+1}|\mathcal{F}_n] &= \mathbb{E}\left[e^{\vartheta X_{n+1} - c(n+1)}|\mathcal{F}_n\right] \\
&= \mathbb{E}\left[e^{\vartheta Y_{n+1} - c} e^{\vartheta X_n - cn}|\mathcal{F}_n\right] \\
&= \mathbb{E}\left[e^{\vartheta Y_{n+1}}|\mathcal{F}_n\right] e^{\vartheta X_n - c(n+1)}
\end{aligned}$$

thus

$$\begin{aligned}
\mathbb{E}\left[e^{\vartheta Y_{n+1}}|\mathcal{F}_n\right] e^{\vartheta X_n - c(n+1)} &\overset{?}{=} e^{\vartheta X_n - cn} \\
\mathbb{E}\left[e^{\vartheta Y_{n+1}}\right] e^{-c} &\overset{?}{=} 1 \\
e^{c} &\overset{?}{=} \mathbb{E}\left[e^{\vartheta Y_{n+1}}\right] = \frac{1}{2}\left(e^{\vartheta} + e^{-\vartheta}\right) = \cosh\vartheta
\end{aligned}$$

so

$$c = \log\cosh\vartheta \quad (= c(\vartheta))$$

With such c, $(M_n)_{n\geq 0}$ is a Martingale

$$|M_n| = M_n \quad \Rightarrow \quad \mathbb{E}[|M_n|] = \mathbb{E}[M_n] = \mathbb{E}[M_0] = 1$$

By Martingale convergence theorem, $M_n \xrightarrow[n\to\infty]{\text{a.s.}} k$ and by the Law of Large Numbers we can compute k.

$$X_n = \sum_{k=1}^{n} Y_k \quad \Rightarrow \quad \frac{1}{n}\sum_{k=1}^{n} Y_k \xrightarrow[n\to\infty]{\text{a.s.}} \mathbb{E}[Y_1] = 0$$

then

$$e^{\vartheta X_n - c(\vartheta)n} = \left(e^{\overbrace{\vartheta \frac{1}{n}\sum_{k=1}^{n} Y_k}^{\to 0} - c(\vartheta)}\right)^n \sim \left(e^{-c(\vartheta)}\right)^n \to 0 = k$$

since $c(\vartheta) > 0$, $e^{-c(\vartheta)} < 1$.

Summing up:

$$M_n \xrightarrow[n\to\infty]{\text{a.s.}} 0 \quad \text{but} \quad \mathbb{E}[M_n] = 1 \quad \Rightarrow \quad M_n \nrightarrow 0 \text{ in } L^1$$

2.3 Supermartingales and Submartingales

Consider $(X_t)_{t\geq 0}$ with either $t \in \mathbb{N}$ or $t \in \mathbb{R}^+$ and a filtration $(\mathcal{F}_t)_{t\geq 0}$. We say that $(X_t)_{t\geq 0}$ is a supermartingale (resp. submartingale) if

- (adaptedness) X_t is $\mathcal{F}_t$-measurable for all t.
- (integrability) $\mathbb{E}[|X_t|] < +\infty$
- $\forall s < t$,

$$\mathbb{E}[X_t|\mathcal{F}_s] \leq X_s, \text{ for supermartingales}$$
$$\mathbb{E}[X_t|\mathcal{F}_s] \geq X_s, \text{ for submartingales}$$

Remark. A process is a Martingale if and only if it's both a sub- and super-Martingale.

Theorem 2.6. If $(M_t)_{t\geq 0}$ is a Martingale and $\varphi : \mathbb{R} \to \mathbb{R}$ is a **convex** (resp. **concave**) function such that $\mathbb{E}[|\varphi(M_t)|] < +\infty$ for any t, then

$$(\varphi(M_t))_{t\geq 0}$$

is a **submartingale** (resp. **supermartingale**) w.r.t. the same filtration of $(M_t)_{t\geq 0}$.

Proof. Consider φ convex, using the Jensen inequality

$$s < t \quad \varphi(M_s) = \varphi(\mathbb{E}[M_t|\mathcal{F}_s]) \leq \mathbb{E}[\varphi(M_t)|\mathcal{F}_s]$$

□

Examples.

$\varphi(x) = |x|$, $\varphi(x) = |x|^2$, $\varphi(x) = \max\{0, x\}$.

2.4 Doob-Meyer decomposition.

Consider a submartingale $(X_n)_{n\geq 0}$ with $n \in \mathbb{N}$ w.r.t. a filtration $(\mathcal{F}_n)_{n\geq 0}$ then it can be written as

$$X_n = M_n + A_n$$

where $(M_n)_{n\geq 0}$ is a martingale and A_n is $\mathcal{F}_{n-1}$-measurable and increasing $A_{n+1} \geq A_n$ a.s.[4]

[4] A_n is a *predictable* process.