Chapter 1

A first study in stochastic processes: Martingales

Okay, I'll admit that this part is really hard for me to follow. I am already broken down by Professor Polito's lectures and right now... I only want to be happy again. Please note that the astonishing drop in the quality of the notes is not my fault. This shit should be easier to explain but they manage to mess this up anyway.

1.1 Revising stochastic processes and filtrations

> **Note the notation!**
>
> 1. we will work in a fixed probability space $(\Omega, \mathcal{H}, \mathbb{P})$;
>
> 2. we will denote an index set $\mathbb{T}$ (which is usually interpreted as time). It could be $\mathbb{T} = \mathbb{N}$ or $\mathbb{T} = \mathbb{R}$, for example;
>
> 3. $\{X_t : t \in \mathbb{T}\}$ is a random variable taking values in the product space
>
> $$(F, \mathcal{F}) = (E^{\mathbb{T}}, \mathcal{E}^{\mathbb{T}}).$$
>
> We can think of this as an evolving stochastic process.

Let $\mathbb{T}$ be a subset of $\mathbb{R}$. For each $t \in \mathbb{T}$ let $\mathcal{F}_t$ be a sub-σ-algebra of $\mathcal{H}$. The family $\mathcal{F} = \{\mathcal{F}_t : t \in \mathbb{T}\}$ is called a **filtration** provided that

$$\mathcal{F}_s \in \mathcal{F}_t \qquad \forall \, s < t.$$

> **Definition 1.1.1**
>
> A **stochastic process** $X = (X_t)_{t \in \mathbb{T}}$ generates a filtration that is named **natural filtration**.

As the time passes, I discover more and more information about $\mathcal{H}$. Each process defines a filtration by generating a "flow of information".

> **Example 1.1.1**
>
> Imagine a car competition composed by two different races. Consider the events
>
> $$\omega = \{\text{car } A \text{ wins}\}$$
> $$\ell = \{\text{car } A \text{ loses}\}.$$
>
> Consider 2 races: our Ω will be
>
> $$\Omega = \{\omega\omega, \ell\omega, \omega\ell, \ell\ell\}.$$
>
> Before the starting of the race, at time 0, we have no information about what is going to happen so our filtration will be
>
> $$\mathcal{F}_0 = \left\{ \underbrace{\{0\}}_{\varnothing}, \underbrace{\{\omega\omega, \ell\omega, \omega\ell, \ell\ell\}}_{\Omega} \right\}$$

which is the trivial σ-algebra. Remember that in this case the empty set $\varnothing$ represents impossible events (e.g. all cars win all races). So the random variable X_0 measurable with respect to $\mathcal{F}_0$ maps all the outcomes of the race to the same value. The idea is that we cannot ask ourselves "what is the probability of the car winning the first and the second race?" because we aren't able to measure these events in $\mathcal{H}$. Now consider the situation after the first race: I am now able to "discern" the events regarding the first race:

$$\mathcal{F}_1 = \{\{0\}, \{\omega\ell, \omega\omega\}, \{\ell\ell, \ell\omega\}, \{\omega\omega, \omega\ell, \ell\omega, \ell\ell\}\}.$$

Now I am able to know what happened after the first race and this gives me a richer sub-σ-algebra, with $\mathcal{F}_0 \subset \mathcal{F}_1$. I am now able to have a random variable X_1 that is measurable with respect to $\mathcal{F}_1$. For instance, imagine defining the following random variable:

$$X_1 = \begin{cases} 1 & \text{if the car wins the first race} \\ 0 & \text{otherwise.} \end{cases}$$

But, for instance, the random variable "the car wins the second race" is still not measurable with respect to $\mathcal{F}_1$. Consider now the filtration after the second race:

$$\mathcal{F}_1 = \{\{0\}, \{\omega\omega\}, \{\ell\ell\}, \{\omega\ell\}, \{\ell\omega\}, \{\omega\ell, \omega\omega\}, \{\ell\ell, \ell\omega\}, \{\omega\omega, \omega\ell, \ell\omega, \ell\ell\}\}.$$

Of course $\mathcal{F}_2 \supset \mathcal{F}_1 \supset \mathcal{F}_0$. We could define a random variable X_3 of the following type:

$$X_3 = \begin{cases} 0 \\ 1 \\ 2 \end{cases}$$

which counts the numbers of victories of our car. X_3 is $\mathcal{F}_2$-measurable but not $\mathcal{F}_1$-measurable. Now $\mathcal{F}$ is a filtration:

$$\mathcal{F} = \{\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2\}.$$

Example 1.1.2

What if the same experiment is observed by different people? In other ways, what if we have different learning from the same experiment? Imagine that we have two filtrations $\mathcal{F}$ and $\mathcal{G}$. We say that $\mathcal{F}$ is finer than $\mathcal{G}$ if

$$\mathcal{F}_t \supset \mathcal{G}_t \qquad \forall\, t \in \mathbb{T}.$$

Definition 1.1.2

Given $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}$, a stochastic process $\{X_t\}_{t \in \mathbb{T}}$ is **adapted** to $\mathcal{F}$ if X_t is $\mathcal{F}_t$-measurable for each $t \in \mathbb{T}$ and $\mathcal{E}$-measurable.

We already mentioned that every process generates its *natural filtration* $\sigma(X_t, t \in \mathbb{T})$, for which the process is clearly adapted. But we may want to consider another filtration $\mathcal{F} = (F_t)$ such that $\{X_t\}$ is adapted to $\mathcal{F}$ because all the information on X that I care about are contained in this filtration. What does this mean and what is the relationship between the two? Which one of the two is finer considering that our process is adapted to both? The answer is

$$\mathcal{F} \supset \sigma(X_t) \qquad (\mathcal{F} \text{ is finer than the natural filtration})$$

because $\sigma(X_t)$ is always the smallest possible σ-algebra that a process can generate (the natural filtration is the *minimum* filtration).

Another way we could see the question of adaptation is:

"Can X be determined using only information available in $\mathcal{F}$?"

If yes, then X is adapted to $\mathcal{F}$.

Example 1.1.3

This is called "the secretary problem": in this case we must start from the filtration and then understand the problem. Here i have N candidates for a position; a candidate disregarded after the interview is lost. The

interviewer wants to hire exactly 1 candidate and each candidate has different abilities and the interviewer knows only the relative ability of those already interviewed so far. Our goal is to maximizing the probability of hiring the best one. We have three questions:

1. what is Ω?

2. what is the filtration $\mathcal{F}$ for this experiment?

3. what process should we use?

In this case $\Omega = N!$ permutations of the ranking of the candidates (the order in which they show up) and the filtration is the information earned from interview up to time t (that is the ranking of the candidates up to time t). But what is the process that I should use? Consider the sequence

$$V_1, V_2, \ldots \qquad \{V_i\}_{i \geqslant 1}$$

with $V_i = 1$ if and only if the best candidate is the i-th candidate and $V_i = 0$ otherwise. Could this process $\{V_i\}_{i \geqslant 1}$ be used? No, because V is not adapted to $\mathcal{F}$... because to understand if i-th candidate is te best we need to compare it to the other candidates, including the ones that didn't show up yet! But then how can we get an adapted process? Let us consider the expectation

$$U_n = \mathbb{E}\left[V_n | \mathcal{F}_n\right]$$

What do we know about the measurability of U_n? We know that it is for sure $\mathcal{F}_n$-measurable. This trick gives us a simple way to build an adapted process. So now we will have: $U_n = 0$ if the candidate is not the best up to n and $U_n = 1$ otherwise. More specifically, we will have

$$U_n = 1 \cdot \underset{\substack{\text{is among the first } n}}{\text{proability that the best candidate}} + 0 \cdot \underset{\substack{\text{is not among the first } n}}{\text{proability that the best candidate}}$$
$$= 1 \cdot \frac{n}{N} + 0 \cdot \frac{N-n}{N}.$$
$$= \frac{n}{N}$$

This is a quantity that I can measure and it is therefore adapted.

Figure 1.1: You conditioned in the wrong hood, buddy.

1.1.1 Stopping times

Consider now the same problem under a different point of view: we want to establish the "best time" when to stop discarding candidates. This is a random time and it is called a **stopping time**. This is nothing else than a time that we choose basing ourselves on elements or events that may or may not happen: this is why they're a random variable. The stopping times are a random variable that can cause problems, since stopping times can be infinite: imagine that our stopping time is "whenever a volcano erupts" but the stupid fucking volcano decides to go into quiescence forever. This is why beside our index set $\mathbb{T}$ we need to introduce $\overline{\mathbb{T}} = \mathbb{T} \cup \{\infty\}$.

Definition 1.1.3

$$T : \Omega \mapsto \overline{\mathbb{T}}$$

is called **stopping time** of $\mathcal{F}$ if

$$\{T \leqslant t\} \in \mathcal{F}_t \qquad \forall t \in \mathbb{T}.$$

Remember that stopping times have to be $\mathcal{F}$-measurable: for example "when will the fucking volcano erupt for the last time" is *not* a stopping time because it requires information that we don't have in t!

Stopping time have an alternative definition:

Definition 1.1.4

T is a stopping time if the process

$$Z_t = \mathbb{1}_{\{T \leqslant t\}} \qquad t \in \mathbb{T}$$

is adapted to F.

From this definition it follows that when $\mathbb{T} = \mathbb{N}$ then this is equivalent to having

$$\underbrace{\hat{Z}_n}_{Z_n - Z_{n-1}} = \mathbb{1}_{\{T=1\}} \qquad \forall n \in \mathbb{N}$$

that needs being adapted to $\mathcal{F}_n$.

Not all random times are stopping time: think about an experiment in which we produce a signal when a physical event happens. Let us call Z the process indicator and T our stopping time:

$$Z_t \omega = \begin{cases} 0 & t < T(\omega) \\ 1 & t \geqslant T(\omega). \end{cases}$$

Then if the information flow allows us to recognize the occurrence of the event then T is a stopping time.

Example 1.1.4

Some stopping times:

① The first time that $X(\omega) \in H \in \Omega$;

$$T(\omega) = \begin{cases} \inf \{n \in \mathbb{N} : X_n(\omega) \in H\} \\ +\infty & \text{if } X_n(\omega) \notin H \ \forall n \end{cases}$$

So

$$\{T \leqslant n\} = \bigcup_{k=0}^{n} \{X_k \in H\}.$$

② Consider i.i.d. random variables $X_1, X_2, \ldots$. Consider the probabilities

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

and the random walk

$$S_n = \sum_{1}^{n} X_i.$$

Let's define

$$T_1 = \begin{cases} \min \{n < 50 : S_n = 3\} \\ 50 & \text{otherwise.} \end{cases}$$

This is a stopping time because I can write $\{T_1 \leqslant n\}$ as

$$\underbrace{\bigcup_{k=1}^{n} \underbrace{\{S_k = 3\}}_{\mathcal{F}_k\text{-measurable}}}_{\mathcal{F}_n\text{-measurbale}} \qquad n < 50.$$

Moreover for $n = 50$ we have $T_1 \in \mathcal{F}_{50}$.

③ Starting from the previously deifned random walk, consider the quantity

$$M_n = \min(S_1, \ldots, S_n)$$

And the random time

$$T_2 = \min\{n : S_n \geqslant M_m + 2\}$$

is a stoppimg time. On the contrary,

$$T_3 = \begin{cases} \max\{n < 50 : S_n = 7\} & \text{if not empty} \\ 50 & \text{otherwise} \end{cases}$$

is not a stopping time. Why? Because I have to wait until $n = 50$ to answer the question.

Consider the random times on $\mathbb{R}_+$

$$0 < T_1 < T_2 < \ldots$$

With $\lim_{n \to \infty} T_n = +\infty$. Define the process $\{N_t\}$ as

$$N_t := \sum \mathbb{1}_{[0,t]}(T_n).$$

This is called **counting process**. It is a basic count of the number of events happened up to time n.
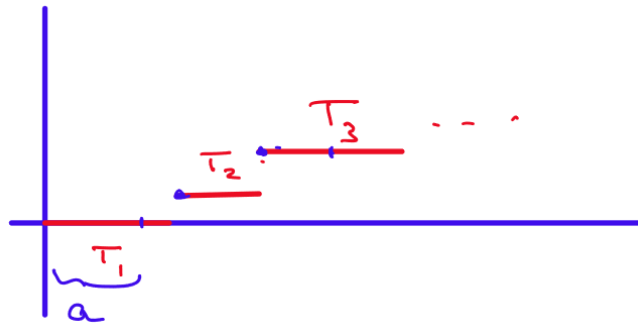

Figure 1.2: Yeah I'm counting on ending my life very soon.

N_t is increasing, right continuous and incresases by unitary jumps. Moreover, $N_0 = 0$, $N_t < \infty$ for $t \in \mathbb{R}_+$. Of course $\lim_{t \to \infty} N_t = \infty$. Counting processes generate their natural filtration $\mathcal{F}$.

Some problems require to stop the observation at a stopping time (because we don't care anymore[1])... So we don't actually need the whole knowledge of the complete filtration $\mathcal{F} = (\mathcal{F}_t)_{t \geqslant 0}$. The problem is that as we said before the stopping time is a random time. So what we need is the *information known up to time T* $\mathcal{F}_T$, basically the σ-field that is the filtration at time T.

> **Remark**
>
> We said that $\mathcal{F}_t$ does not include $t = \infty$... but T may diverge to ∞! We need to enlarge $\mathcal{F}$ to include $t = \infty$.

Think, for instance the stopping time "first time of overflowing of a dam". It is something that may very well never happen and is therefore called "Vajont stopping time[2]".

---

[1] Assuming we ever did.

[2] Okay I'm sorry I'm sorry I am a terrible person yes sir yes yes I concur it was an extremely inappropriate joke yes yes sir yes you have my word yes no no never again no you don't have to worry no no.

Definition 1.1.5

Let $\overline{\mathbb{T}} = \mathbb{T} \cup \{+\infty\}$. We define
$$\mathcal{F}_\infty = \lim_{t \to \infty} \mathcal{F}_t = \underbrace{\bigvee_{t \in \mathbb{T}} \mathcal{F}_+}_{\sigma(\bigcup \mathcal{F}_t)} .$$

Definition 1.1.6

$\mathcal{F}_{\overline{\mathbb{T}}}$ is a filtration on $\overline{\mathbb{T}}$ and T is a stopping time on it if and only if it is a stopping time on $\mathcal{F}_{\mathbb{T}}$.

Remark

Every stochastic process on $(\mathcal{F}_t)_{t \in \mathbb{T}}$ can be extended to $\overline{\mathbb{T}}$ by appending any random variable X_∞ to it (with $X_\infty \in \mathcal{F}_\infty$).

So adding this random variable doesn't really add any information but only ensures that nerds don't whine about completeness. It is *technically* information that we will never user, since $\mathcal{F}_\infty$ does not contain any further information with respect to $\mathcal{F}$. We denote with $\mathcal{F}$ also the extended filtration, with the ease of mind given by the fact that a divergent stopping time won't make us commit harakiri.

Definition 1.1.7

Let $\mathcal{F}$ be a filtration on $\overline{\mathbb{T}}$ and let T be a stopping time on $\mathcal{F}$. We call **past until T** the σ-algebra $\mathcal{F}_t \subset \mathcal{F}_\infty \subset \mathcal{H}$ such that:
$$\mathcal{F}_T = \Big\{ H \in \mathcal{H} : H \cap \{T < t\} \in \mathcal{F}_t, \forall t \in \overline{\mathbb{T}} \Big\} .$$

This means that $\mathcal{F}_T$, which represents the evolution within T, contains all the events H such that they are before the stopping time in all the filtrations. This is selecting the events H that happen within the time T.

Remark

If we fix a $T \equiv t$ then $\mathcal{F}_T \equiv \mathcal{F}_t$ (it's the normal filtration at time t).

If T is a stopping time of $\mathcal{F}$ then $\{T \leqslant r\}$ belongs to $\mathcal{F}_T$, $\forall r \geqslant 0$:
$$\{T \leqslant r\} \cap \{T \leqslant t\} = \{T \leqslant \min(T,r)\} \in \mathcal{F}_t$$

So T is $\mathcal{F}_T$-measurable.

Remark

$\mathcal{F}_t$ can be read as the collection of all $\mathcal{F}_t$-measurable random variables V: the value of $V(\omega)$ can be told by the time $T(\omega)$: we can read the value of $V(\omega)$ before the ringing of an alarm.

Theorem 1.1.1

A random variable V belongs to $\mathcal{F}_T$ if and only if
$$V\mathbb{1}_{T \leqslant t} \in \mathcal{F}_t, \qquad \forall t \in \overline{\mathbb{T}}. \tag{i}$$

If $\overline{\mathbb{T}} \equiv \overline{\mathbb{N}}$ the condition becomes
$$V\mathbb{1}_{T=n} \in \mathcal{F}_n \qquad \forall n \in \overline{\mathbb{N}}.$$

Proof

Note that
$$X_t = V\mathbb{1}_{T \leqslant t}$$

is a stochastic process. For $\forall r \in \mathbb{R}_+$ and $t \in \mathbb{T}$

$$\{V > r\} \cap \{T \leqslant t\} = \{X_t > r\}.$$

But now we know that (i) is just a consequence of the definition of $\mathcal{F}_t$:

$$\{V > r\} \in \mathcal{F}_T \quad \forall r \underset{\text{def. of } \mathcal{F}_T}{\iff} \{X_t > r\} \in \mathcal{F}_t \qquad \forall r, \forall t \in \overline{\mathbb{T}}$$

The special case in which $\overline{\mathbb{T}} \equiv \overline{\mathbb{N}}$ means that

$$V\mathbb{1}_{\{T=n\}} = \begin{cases} X_n - X_{n-1} & \text{if } n \in \mathbb{N} \\ X_\infty - \sum_{n \in \mathbb{N}} (X_n - X_{n-1}) & \text{if } n = \infty \end{cases}$$

so that

$$V \in \mathcal{F}_T \iff V\mathbb{1}_{T=n} \in \mathcal{F}_n$$

$\square$

But what if there is a problem that involves more than 1 stopping time? We cannot use two different "pasts" because that would cause problems with the measurability of random variables.

Example 1.1.5

Imagine a factory. The production is blocked when

    a) the temperature of the room is above a certain threshold;

    b) the machinery has not been cleaned for more than 12 hours.

Alternatively, imagine that the sales of shares happens:

    a) when the price is above a fixed value;

    b) when the increase of the price is smaller than a fixed value.

Both of this cases require the use of 2 different stopping times S and T. We may think to evaluate two typical functions of S and T of interest:

$$\underbrace{\min\{S,T\}}_{S \wedge T} \qquad \underbrace{\max\{S,T\}}_{S \vee T}$$

Example 1.1.6

Truncated stopping time: let T be a stopping time (for example, the time at which we sell certain shares) and that we want a finite horizon for this decision. In this case the quantity of interest is

$$S = T \wedge n = \min\{T, n\}$$

where n could be some sort of time horizon.

Imagine that two cyclists participate to a race. Their children will have their snack when both the parents will arrive to the finish line. How long will the children wait for their snack? We can think about the following stopping times:

$$\begin{aligned} T &: \quad \text{time employed by the first cyclist} \\ S &: \quad \text{time employed by the second cyclist} \\ U &: \quad \max\{S,T\}. \end{aligned}$$

The waiting time for the children will be U.

This is cool and all, but are we sure that $S \wedge T$ and $S \vee T$ are stopping times?

Why?

Aw hell nah. Now even the color boxes are against me.

Well this is not a surprise.

Proof

We know that $\{S \leqslant t\}$ and $\{T \leqslant t\}$ belong to $\mathcal{F}_t$ for every t, since they are stopping times. Consider now $\{\min\{(S, T) \leqslant t\}\}$ and $\{\max\{(S, T) \leqslant t\}\}$. We have

$$\{\min\{(S, T) \leqslant t\}\} = \underbrace{\underbrace{\{S \leqslant t\}}_{\in \mathcal{F}_t} \cup \underbrace{\{T \leqslant t\}}_{\in \mathcal{F}_t}}_{\in \mathcal{F}_t}$$

and

$$\{\max\{(S, T) \leqslant t\}\} = \underbrace{\underbrace{\{S \leqslant t\}}_{\in \mathcal{F}_t} \cap \underbrace{\{T \leqslant t\}}_{\in \mathcal{F}_t}}_{\in \mathcal{F}_t}.$$

□

There is a further theorem:

Theorem 1.1.2

① if $S \leqslant T \implies \mathcal{F}_S \subset \mathcal{F}_T$;

② $\mathcal{F}_{\min\{S,T\}} = \mathcal{F}_S \cap \mathcal{F}_T$;

③ if $V \in \mathcal{F}$, the following processes are in $\mathcal{F}_{\min\{S,T\}}$:

$$\begin{array}{ccc} (a) & (b) & (c) \\ V\mathbb{1}_{S \leqslant T}, & V\mathbb{1}_{S=T}, & V\mathbb{1}_{S<T}. \end{array}$$

Proof

1. Imagine $\{S \leqslant T\}$. Now consider the event $\{S \leqslant n\}$. Since S happens before T then we have that $\{T \leqslant n\} \subset \{S \leqslant n\}$. We know that if $A \in \mathcal{F}_S$ then

$$A \cap \{T \leqslant n\} = A \cap \underbrace{\{S \leqslant n\}}_{\in \mathcal{F}_n} \cap \underbrace{\{T \leqslant n\}}_{\in \mathcal{F}_n}$$

which means that $A \cap \{T \leqslant n\} \in \mathcal{F}_n$ and by definition of $\mathcal{F}_T$ this correponds to saying

$$\mathcal{F}_S \subset \mathcal{F}_T.$$

2.   (a) Let us first prove that $\mathcal{F}_{\min\{S,T\}} \subset \mathcal{F}_S \cup \mathcal{F}_T$. Note that $\min\{S,T\}$ is dominated by $\mathcal{F}_s$ and $\mathcal{F}_T$, which means that $\mathcal{F}_{\min\{S,T\}}$ is included both in $\mathcal{F}_S$ and $\mathcal{F}_T$. By point 1 we have

$$\mathcal{F}_{\min\{S,T\}} \subset \mathcal{F}_S \cap \mathcal{F}_T.$$

  (b) It remains to prove that $\mathcal{F}_{\min\{S,T\}} \supset \mathcal{F}_S \cup \mathcal{F}_T$. Let $H \in \mathcal{F}_S \cap \mathcal{F}_T$. This means that

$$H \cap \{S \leqslant T\} \in \mathcal{F}_{\min\{S,T\}}$$

by part the of this theorem, since $H \in \mathcal{F}_S$. But we also know that

$$H \cap \{T \leqslant S\} \in \mathcal{F}_{\min\{S,T\}}$$

by part 3 of this theorem since also $H \in \mathcal{F}$. But this means that

$$H = \{H \cap \{S \leqslant T\}\} \cup \{H \cap \{S \leqslant T\}\}$$

belongs to $\mathcal{F}_{\min\{S,T\}}$ and so

$$\mathcal{F}_{\{S \cap T\}} \subset \mathcal{F}_{\min\{S,T\}}.$$

3.   (a) $V\mathbb{1}_{\{S \leqslant T\}} \in \mathcal{F}_{\min\{S,T\}}$. To prove this we use the following theorem:

> **Theorem 1.1.3**
>
> Let T be a stopping time of $\mathcal{F}$. Then
>
> $$\mathcal{F}_T = \{X_T : X \in \mathcal{F}\}$$
>
> with X in $\mathcal{F}$.

This baiscally means that $\mathcal{F}_t$ becomes the values X_T of X in $\mathcal{F}$ at time T. Let us consider $X_t = V\mathbb{1}_{\{S \leqslant t\}}$ with $t = \min\{S,T\}$. We know that $\min\{S,T\}$ is a stopping time of $\mathcal{F}_{\min\{S,T\}}$ and by the theorem above

$$X_{\min\{S.T\}} \in \mathcal{F}_{\min\{S,T\}}$$

or, alternatively,

$$\mathbb{1}_{\{S \leqslant T\}} \in \mathcal{F}_{\min\{S,T\}}.$$

  (b) Let $V = 2 \implies \{S \leqslant T\} \in \mathcal{F}_{\min\{S,T\}}$ and by symmetry $\{T \leqslant S\} \in \mathcal{F}_{\min\{S,T\}}$. This means that

$$\{S = T\} = \{S \leqslant T\} \cap \{T \leqslant S\} \in \mathcal{F}_{\min\{S,T\}}.$$

Furthermore

$$\{S < T\} = \{S \leqslant T\} \setminus \{S = T\}$$

Which implies memebership of $\mathcal{F}_{\min\{S,T\}}$.

  (c) Consider the fact already proven that $V\mathbb{1}_{\{S \leqslant T\}} \in \mathcal{F}_{\min\{S,T\}}$. Multiplying by $\mathbb{1}_{\{S = T\}}$ or $\mathbb{1}_{\{S < T\}}$ does not alter the membership in $\mathcal{F}_{\min\{S,T\}}$.

$\square$

Consider now a machinery what when it becomes too hot (above a certain threshold) an alarm rings and after 3 seconds the machinery stops. Here we deal with 2 times:

$$
\begin{aligned}
T : &\quad \text{time of the alarm (stopping time)} \\
t = 3 : &\quad \text{deterministic time (stopping time as well)}
\end{aligned}
$$

and the time of our interest is the time S when the machinery stops:

$$S = T + t$$

> **Definition 1.1.8**
>
> Let
>
> | | |
> |---|---|
> | S | be a stopping time |
> | T | be a random time $T > S$ but whose value can be told by time S. |
>
> We say that T is **foretold by** S (and it is a stopping time).

> **Example 1.1.7**
>
> Consider S and $T = 2S$. Surprise surprise, T is a foretold time.

Did we need an example for this?

> **Example 1.1.8**
>
> Yes.

Yeah this meta-joke has been running long enough.

1.1.2 Discrete and continuous stopping times

> **Definition 1.1.9**
>
> Let
> $$d_n(t) = \begin{cases} \frac{k+1}{2^n} & \text{if } \frac{k}{2^n} \leqslant t < \frac{k+1}{2^n} \\ +\infty & \text{if } t = \infty. \end{cases}$$
>
> We got a function
> $$d_n : \overline{\mathbb{R}}_+ \mapsto \overline{\mathbb{R}}_+.$$

Consider the case $n = 1$ (remember that $k = 0, 1, 2, \ldots$):

Figure 1.3: These are screenshots taken from the video feed. I couldn't bother. Sorry.

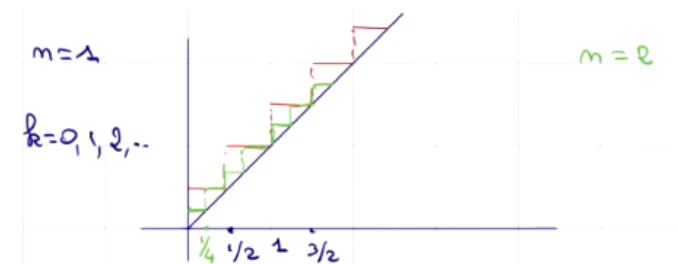But after $n = 2$ we get the same function with smaller "steps":

Figure 1.4: The aesthetic problem of musical execution is reduced by Enrico Fubini to a problem of musical improvisation: I honestly think that this perspective is limiting both in its conceptions of "execution" and "improvisation" even if I agree with the intent of blurring the line between the two concepts unjustly separated in our Western modern musical. That is, a modern perspective: the reader will surely remember that Bach was famous first and foremost for his improvisational abilities and that the musical form of the fugue was often used to display pyrotechnical capabilities in that regard.

Observe that $d_1 \geqslant d_2 \geqslant d_3 \geqslant \ldots$. What are the other properties of this function?

1. it is a step function;

2. it is right-continuous;

3. $d_n(t) > t$ (the diagonal line);

4. $\lim_{n\to\infty} d_n(t) = t$.

Now we apply this simple function to our stopping times so that we get the following proposition.

> **Proposition 1.1.1**
>
> Let $\mathcal{F}$ be a filtration on $\overline{\mathbb{R}}_+$ and let T be a stopping time. Let
> $$T_n = d_n \circ T.$$
> Then $\{T_n\}_n$ is a sequence of **discrete stopping times** of $\mathcal{F}$ decreasing to T.

> **Remark**
>
> Of course T_n is a foretold time by T, because if I know T I also know the value of T_n.

Proof

Fix n:

- T_n is a measurable function of $T \implies T_n \in \mathcal{F}_n$;
- $d_n(t) \geqslant t \quad \forall t < \infty$ and $d_n(\infty) = \infty \implies d_n(T) = T_n \geqslant T$;
- T_n is a foretold time by T;
- T_n is a stopping time;
- T_n is discrete.

But $d_n(t) \searrow t$ as $n \to \infty \implies d_n(T) = T_n \searrow T$. We have thus switched from a continuous stopping time to a sequence of discrete stopping times. □

An important thing to note is that during our process of discretization we never break the property of the stopping times... and this is possible because we are using foretold times!

1.1.3 Conditioning at stopping times

Imagine that we have our information about X that stops at the stopping time T: this means knowing $\mathcal{F}_T$. Consider the classic conditioning situation
$$\mathbb{E}(X\mathcal{F}_T).$$

We already said how this is our best estimate of X using our information available (that is $\mathcal{F}_T$).

> **Note the notation!**
>
> We will write
> $$\mathbb{E}_T \equiv \mathbb{E}_{\mathcal{F}_T} \equiv \mathbb{E}(\cdot|\mathcal{F}_T).$$

Can I use T fixed as a deterministic time t and thus get $\mathbb{E}(X|\mathcal{F}_t)$? Is this a special case or a different think altogether? Well yes, because t *is* a stopping time being $\mathcal{F}_t$-measurable! $\mathbb{E}(X|\mathcal{F}_t)$ is a special case of $\mathbb{E}(X|\mathcal{F}_T)$.

This becomes clear when we see the properties of stopped expectation. For $\forall\, X, Y, Z$ being positive random variables and for $\forall\, S, T$ stopping times of $\mathcal{F}$ the following properties hold:

- **defining property**:
$$\mathbb{E}_T X = Y \iff V \in \mathcal{F}_T \text{ and } \mathbb{E}VX = \mathbb{E}VY;$$

- **unconditioning**:
$$\mathbb{E}\mathbb{E}_T X = \mathbb{E}X;$$

- **repeated conditioning**:
$$\mathbb{E}_S \mathbb{E}_T X = \mathbb{E}_{\min\{S,T\}} X = \mathbb{E}_{S \wedge T} X;$$

- **conditional determinism**:
$$\mathbb{E}_T(X + YZ) = X + Y\mathbb{E}_T(Z)$$

if $X, Y \in \mathcal{F}_T$.

We need to prove the third property.

Proof

Notice that if $S \leqslant T \implies \mathcal{F}_S \subset \mathcal{F}_T$. This means that

$$\mathbb{E}_S \mathbb{E}_T = \mathbb{E}_S \; (= \mathbb{E}_{S \wedge T})$$

(the "poorer" σ-algebra wins!). Moreover notice that if S and T are arbitrary we can apply the same result to $S \wedge T$ since $S \wedge T \leqslant T$ and

$$\mathbb{E}_{S \wedge T} \mathbb{E}_T = \mathbb{E}_{S \wedge T} \qquad\qquad (\bullet)$$

Remember that $Y = \mathbb{E}_T X$. Let's now try to write the statement of our proof in a different way:

$$\mathbb{E}_S Y = \mathbb{E}_{S \wedge T} X.$$

But using our property $(\bullet)$ we get

$$\mathbb{E}_S Y = \mathbb{E}_{S \wedge T} \mathbb{E}_T X$$
$$= \mathbb{E}_{S \wedge T} Y.$$

Let's prove this, by steps.

1. note that $\mathcal{F}_{S \wedge T} \subset \mathcal{F}_S$. This means that $\mathbb{E}_{S \wedge T} Y$ is in $\mathcal{F}_S$ and therefore $\mathbb{E}_{S \wedge T} Y$ has the measure property and it is a candidate to be $\mathbb{E}_S Y$. So we should check whether this candidate satisfies the defining property of stopped expectation.

2. Check the defining property
$$\mathbb{E}VY = \mathbb{E}V\mathbb{E}_{S \wedge T} Y$$

for each positive V in $\mathcal{F}_S$. We start by fixing such V, positive and in $\mathcal{F}_S$. We proved that

$$V\mathbb{1}_{\{S \leqslant T\}} \in \mathcal{F}_{S \wedge T}$$

and the defining property for $\mathbb{E}_{S \wedge T}$ gives

$$\mathbb{E}V\mathbb{1}_{\{S \leqslant T\}} Y = \mathbb{E}V\mathbb{1}_{\{S \leqslant T\}} \mathbb{E}_{S \wedge T} Y \qquad\qquad (\bullet\bullet)$$

Now we observe that $Y \in \mathcal{F}_T$ by definition. Also notice that $Y\mathbb{1}_{\{T < S\}} \in \mathcal{F}_{S \wedge T}$. Apply now this fact to the defining property and obtain

$$\mathbb{E}VY\mathbb{1}_{\{T < S\}} = \mathbb{E}V\underbrace{\mathbb{E}_{S \wedge T} Y \, \mathbb{1}_{\{T < s\}}}_{\mathcal{F}_{S \wedge T}\text{-meas.}}$$

and by conditional determinism we get

$$\mathbb{E}VY\mathbb{1}_{\{T<S\}} = \mathbb{E}V\mathbb{1}_{\{T<S\}}$$
$$= \mathbb{E}V\mathbb{1}_{\{T<S\}}\mathbb{E}_{S\wedge T}Y.$$

By putting together this result with (••) we get

$$\mathbb{E}VY = \mathbb{E}V\mathbb{E}_{S\wedge T}Y.$$

□

1.2 Martingales

When we study stochastic processes, the typical way of introducing classes of stochastic processes for which we have methods to study is determining the finite dimensional distributions. If we have them, thanks to the Kolmogorov Consistency theorem, we can give a complete description of the process but this is boring and for fucking losers. So we introduce classes of stochastic processes for which we can write the distribution.

Imagine, from a different view point, another approach: instead of starting from the distribution of a stochastic process we start from a property; then as a consequence of that property we infer the behavior towards infinity. This is the philosophy behind the class of **martingales**.

> **Example 1.2.1**
>
> Consider the process $\{X_n\}$ that describes the fortune of a player. This player wins with probability p and loses with probability $1 - p = q$. Consider the fortune at time n $\{S_n\}_{n\geqslant 0}$, which is
>
> $$S_n = \sum_{i=0}^{n} X_i.$$

Is the game fair? Imagine also that the X_i's are the value that I get by losing or selling options or stock. We know that the game is fair if $p = q = \frac{1}{2}$. But this could also be written as a property in the following way:

$$\mathbb{E}\left[S_{n+1}\Big|\underbrace{S_n}_{\mathcal{F}_n}\right] = S_n.$$

So I don't have any "advantages" or "bias".

Proof

$$\mathbb{E}\left[S_{n+1}|S_n\right] = \mathbb{E}\left[S_n + X_{n+1}|S_n\right]$$
$$= S_n + \mathbb{E}\left[X_{n+1}|S_n\right]$$
$$= S_n + \underbrace{\mathbb{E}\left[X_{n+1}\right]}_{=0 \text{ if } p=q=\frac{1}{2}}$$
$$= S_n$$

□

We will not define martingales through the event/filtration that allows the process Y to remain constant:

$$\mathbb{E}[Y_{n+1}|\mathcal{F}_n] = Y_n.$$

Then we should ask some questions about this σ-algebra. We know that $\{\mathcal{F}_n\}$ can be increasing, constant or decreasing but Y must remain constant in mean. Look at it through the tower property.

$$\mathbb{E}Y_n = \mathbb{E}\left[\mathbb{E}\left[Y_{n+1}|\mathcal{F}_n\right]\right] = \mathbb{E}Y_{n+1}.$$

> **Definition 1.2.1**
>
> A real-valued process
>
> $$X = (X_t)_{t \in \mathbb{T}}$$
>
> is called a $\mathcal{F}$-**martingale** if:
>
> 1. it is adapted to $\mathcal{F}$;
>
> 2. it is integrable for each $t \in \mathbb{T}$;
>
> 3. $\mathbb{E}(X_t - X_s | \mathcal{F}_s) = 0 \quad \forall s < t$.
>
> If $\mathbb{E}(X_t - X_s | \mathcal{F}_s) \geqslant 0 \quad \forall s < t$ then the process is called $\mathcal{F}$-**submartingale** and if $\mathbb{E}(X_t - X_s | \mathcal{F}_s) \leqslant 0 \quad \forall s < t$ it is called $\mathcal{F}$-**supermaringlae**.

These three properties are what we need to check to establish whether a process has martingale properties. The first two are really just regularity conditions, while the third is the "real" defining property of a martingale. There is is another definition for the third property:

> **Definition 1.2.2**
>
> 3'. $\mathbb{E}(X_t | \mathcal{F}_s) = X_s$.

In the first definition we were looking at the increment, now we are looking at the mean of a further point t in time for the process X and discovering that, conditioning it to information available in time s, it doesn't change from the process in time s.
If $\mathbb{E}X_t^+ < \infty$ then the sequence $\{X_t\}$ (that can be a sub-martingale) converges to a limit.

> **Remark**
>
> A martingale is also a sub-martingale and a super-martingale.

We will use sub-martingales in our demonstrations but it really does not change anything because if $\{X_t\}$ is a $\mathcal{F}$-sub-martingale then $\{-X_t\}$ is a $\mathcal{F}$-super-martingale.

> **Remark**
>
> The number of terms in $\{X_t\}$ can be bounded or unbounded: even a process with 4 terms can be a martingale.

There are consequence to the way we defined martingales. Consider the times $s < t < u$ and the $\mathcal{F}$-martingale $\{X_t\}$. Then

$$\mathbb{E}_s(X_u - X_t)$$

means knowing the evolution of the process at time s and looking for the increment of the process between t and u.
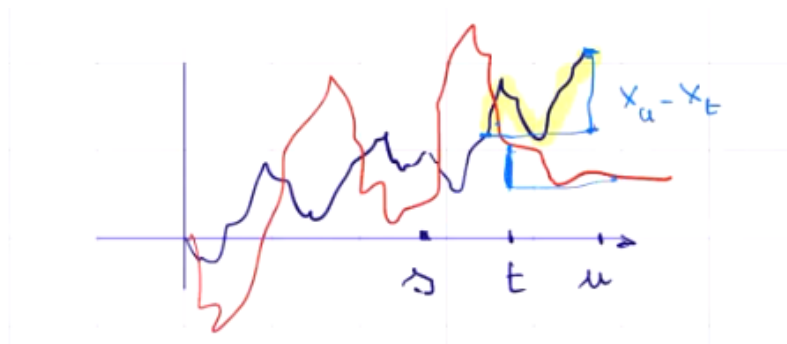

Figure 1.5: It should not

We can write

$$\mathbb{E}_s(X_u - X_t) = \mathbb{E}_s\left(\underbrace{\mathbb{E}_t(X_u - X_t)}_{=0 \text{ bc it is a martingale}}\right)$$

$$= \mathbb{E}_s 0 = 0.$$

14

This means that given the information up to time s the estimation of any future increment will be null! In the special case $\mathbb{T} = \mathbb{N}$ then this property becomes

$$\mathbb{E}_n(X_{n+1} - X_n) = 0.$$

> **Proposition 1.2.1**
>
> If $\{X_i\}$ is an $\mathcal{F}_n$-martingale, it holds
>
> $$\mathbb{E}X_i = \mathbb{E}X_{i+1} \quad \text{if it is a martingale}$$
> $$\mathbb{E}X_{i+1} > \mathbb{E}X_i \quad \text{if it is a sub-martingale.}$$

Proof

Consider the expectation

$$\mathbb{E}(X_{i+1}|\mathcal{F}_i) \geqslant X_i$$

being a sub-martingale. Take the expectation

$$\underbrace{\mathbb{E}\left[\mathbb{E}(X_{i+1}|\mathcal{F}_i)\right]}_{\mathbb{E}X_{i+1}} \geqslant \mathbb{E}X_i.$$

We can now define the process $\{Y_i\}_{i \in \mathbb{N}}$ as

$$Y_{i+1} = X_{i+1} - X_i.$$

We know that $\{X_i\}$ is a $\mathcal{F}$-martingale. Note that

$$\sum_{j=1}^{i} Y_j = X_1 - X_0 + X_2 - X_1 + \ldots + X_i - X_{i-1}$$

so that we can write

$$X_i = c + \sum_{j=1}^{i} Y_j.$$

We know that

$$\mathbb{E}\left[X_{i+1}|\mathcal{F}\right] = X_i$$

But we also know that

$$\mathbb{E}[Y_{i+1}|\mathcal{F}] = 0.$$

$\square$

The quantity $\{\mathbb{E}Y_{i+1}|\mathcal{F}_i\}$ is called **martingale difference**.

> **Example 1.2.2**
>
> Imagine we have a sequence $\{Y_i\}$ of independent random variables with $\mathbb{E}Y_i = 0$. Consider $\mathcal{F}_1$, the σ-algebra generated by $\{Y_j, 0 \leqslant j \leqslant i\}$. Call $S_i = c + \sum^i Y_j$. $\{S_i\}$ is an $\mathcal{F}$-martingale.

1.2.1 Properties of martingales

1. Consider X and Y being $\mathcal{F}$-sub-martingales and $a, b \in \mathbb{R}^+$. Then

$$aX + bY$$

   is a $\mathcal{F}$-sub-martingale.

2. Consider X, Y being two $\mathcal{F}$-sub-martingales. Then

$$\max\{X, Y\}$$

   is a $\mathcal{F}$-sub-martingale.

> ***Proof***
>
> Consider $\mathbb{E}[\max\{X_n, Y_n | \mathcal{F}_{n-1}\}]$. Remember that we have to prove:
>
> - measurability property;
> - integrability property;
> - martingale property.
>
> Often times measurability and integrability are implied in the form of the function. Since we are using the maximum between X and Y we know that
>
> $$\mathbb{E}[\max\{X_n, Y_n | \mathcal{F}_{n-1}\}] \geqslant \mathbb{E}[X_n | \mathcal{F}_{n-1}]$$
> $$\mathbb{E}[\max\{X_n, Y_n | \mathcal{F}_{n-1}\}] \geqslant \mathbb{E}[Y_n | \mathcal{F}_{n-1}].$$
>
> But since X and Y are martingales we know that $\mathbb{E}[X_n | \mathcal{F}_{n-1}] \geqslant X_{n-1}$ and $\mathbb{E}[Y_n | \mathcal{F}_{n-1}] \geqslant Y_{n-1}$; hence we get that
>
> $$\mathbb{E}[\max\{X_n, Y_n | \mathcal{F}_{n-1}\}] \geqslant \max\{X_{n-1}, Y_{n-1}\}$$
>
> so the martingale property is satisfied. □

3. Consider two $\mathcal{F}$-super-martingales X_n, Y_n. Then

$$\min\{X_n, Y_n\}$$

is a $\mathcal{F}$-super-martingale.

> ***Proof***
>
> Consider $\max\{-X_n, -Y_n\}$. We know that $-X_n$ is a sub-martingale and so is $-Y_n$. So as proved above $\max\{-X_n, Y_n\}$ is a sub-martingale. But $\max\{-X_n, -Y_n\} = -\min\{X_n, Y_n\}$ so
>
> $$-\min\{X_n, Y_n\} \text{ is a sub-martingale} \iff \min\{X_n, Y_n\} \text{ is a super-martingale.}$$
>
> □

4. Consider a function f that is <u>convex</u> on $\mathbb{R}$. If X is a $\mathcal{F}$-martingale and $f \circ X$ is integrable then $f \circ X$ is a $\mathcal{F}$-sub-martingale.

> ***Proof***
>
> We have $s < t$ and we want to study $\mathbb{E}\left[f(X_t) | \mathcal{F}_s\right]$. By Jensens's inequality we have
>
> $$\mathbb{E}\left[f(X_t) | \mathcal{F}_s\right] \geqslant f\left[\underbrace{\mathbb{E}(X_t | \mathcal{F}_s)}_{X_s}\right]$$
> $$= f(X_s).$$
>
> □

> **Remark**
>
> Some examples of positive functions of X are $X^+, X^-, |X|$. If X is martingale, these are sub-martingales. $|X|^p$ is a sub-martingale if X is a martingale and $\mathbb{E}|X|^p < \infty$.

5. if f is convex and increasing and X is a $\mathcal{F}$-sub-martingale with $f(X_t)$ integrable $\forall t$ then $f(X_t)$ is again a $\mathcal{F}$-sub-martingale.

1.2.2 Create your own martingale

1. Consider an increasing family of σ-algebras $\{\mathcal{F}_j\}$ and a random variable X on $(\Omega, \mathcal{F}, \mathbb{P})$ which is integrable and adapted to $\mathcal{F}$. Define

$$X_i = \mathbb{E}[X | \mathcal{F}_i].$$

Then the process $\{X_i\}_{i \in \mathbb{N}}$ is called **Doob martingale** and plays an important role in may processes. As i increases, so does the information about our random variable while retaining the underlying process X. We have now to prove that $\{X_i\}$ is a $\mathcal{F}$-martingale.

Proof

We need to check that

$$\mathbb{E}[X_{i+1}|\mathcal{F}_i] = \mathbb{E}\left[\mathbb{E}\left[X|\mathcal{F}_{i+1}\right]|\mathcal{F}_i\right]$$
$$= \mathbb{E}\left[X|\mathcal{F}_i\right]$$

and this is because of the properties of conditional expectation since $\mathcal{F}_{i+1} \supset \mathcal{F}_i$. But $\mathbb{E}\left[X|\mathcal{F}_i\right] = X_i$ by definition. $\square$

The same result holds when $\mathbb{T}$ is continuous.

2. Consider sums of independent random variable $\{X_i\}$ with $\mathbb{E}[X_i] = 0$. We have

$$S_0 = 0 \qquad S_n = S_0 + X_1 + X_2 + \ldots + X_n \qquad n \geqslant 1.$$

We define $\mathcal{F} = (\mathcal{F}_n)_{n \in \mathbb{N}}$ as the filtration generated by $S = (S_n)_{n \in \mathbb{N}}$. Consider the following facts:

   (a) $\mathcal{F}_n$ is adapted to S;
   (b) Each S_n is integrable and $\mathbb{E}S_n = 0$;
   (c) $\mathbb{E}_n(S_{n+1} - S_n) = \mathbb{E}_n(X_{n+1}) = \mathbb{E}X_{n+1} = 0$ (because of independence).

So S_n is a martingale. What happens, tough, if $\mathbb{E}X_i = 1$ for each i? We would get a sub-martingale (because we would have an increasing process).

3. Consider the product of random variables. Consider $R_1, R_2, \ldots$ being independent random variables with

$$\mathbb{E}R_i = 1 \qquad \text{Var } R_i < \infty$$

that don't necessarily need to be identically distributed. Let $M_0 = 1$ and $M_n = M_0 \cdot R_1 \cdot R_2 \cdot \ldots \cdot R_n$ and let $\mathcal{F}$ be the filtration generated by $M = \{M_n\}_{n \in \mathbb{N}}$. Observe that

   (a) M is adapted to $\mathcal{F}$ by definition;
   (b) $\forall M_n$ is integrable: consider $\|M_n\| \leqslant \|M_{n-1}\| \|R_n\|$ since $M_n = M_{n-1}R_n$. But since R_n has finite variance then $\|M_{n-1}\| \|R_n\| < \infty$. So since it is bounded it is integrable;
   (c) consider the expectation

$$
\begin{aligned}
\mathbb{E}[M_{n+1}|\mathcal{F}_n] &= \mathbb{E}[M_n R_{n+1}|\mathcal{F}_n] \\
&= M_n \mathbb{E}[R_{n+1}|\mathcal{F}_n] \qquad \text{(beacuse } M_n \text{ is } \mathcal{F}_n\text{-meas.)} \\
&= M_n \mathbb{E}[R_{n+1}] \\
&= M_n.
\end{aligned}
$$

Example 1.2.3

This is considered a good model for finance: M_n is the price of a share at time n and R_{n+1} is the return at time $(n+1)$ per euro invested at time n. $\mathcal{F}_n$ is the information that is available to the whole market.

4. Consider functions of sums. Consider $\{X_n\}$, a sequence of i.i.d. random variables with

$$\mathbb{E}X_i = 0 \qquad \text{Var } X_i = \sigma^2 < \infty \quad \forall i.$$

Consider

$$S_n = \sum_{i}^{n} X_i$$

and define

$$M_n = S_n^2 - n\sigma^2.$$

M_n is a $\mathcal{F}$-martingale. This is useful to study for example the occurrence of the time in which something happens for the first time. Consider:

(a) M_n is integrable ($\mathbb{V}\mathrm{ar}\, X_i = \sigma^2$);

(b) M_n is adapted by definition;

(c) consider the expectation

$$\mathbb{E}[M_{n+1}|\mathcal{F}_n] = \mathbb{E}\left[\underbrace{(S_n + X_{n+1})^2}_{S_{n+1}^2} - (n+1)\sigma^2|\mathcal{F}_n\right]$$

$$= \mathbb{E}\left[S_n^2 + N_{n+1}^2 + 2S_n X_{n+1} - (n+1)\sigma^2|\mathcal{F}_n\right]$$

$$= S_n^2 + 2S_n \underbrace{\mathbb{E}(X_{n+1}|\mathcal{F}_n)}_{=0} + \underbrace{\mathbb{E}[X_{n+1}^2|\mathcal{F}_n]}_{\mathbb{E}[X_{n+1}^2]=\sigma^2} - (n+1)\sigma^2$$

$$= S_n^2 + \cancel{\sigma^2} - n\sigma^2 - \cancel{\sigma^2} = M_n.$$

Example 1.2.4

Consider $(X_n)_{n\geqslant 1}$ being i.i.d. random variables with common probability density function $g(x)$. Let $f(x)$ be another probability density function with the property

$$f(x) = 0 \qquad \text{for } \frac{x}{g(x)} = 0.$$

Define

$$L_n = \underbrace{\prod_{i=1}^{n} \frac{f(X_i)}{g(X_i)}}_{\text{likelihood ratio}}$$

$$L_0 = 1.$$

Likelihood ration is a $\mathcal{F}$-martingale where $\mathcal{F}$ is the filtration generated by $\{X_n\}$. Contemplate:

(a) it is integrable by definition;

(b) it is adapted by definition;

(c) consider the expectation:

$$\mathbb{E}\left[L_{n+1}|\mathcal{F}_n\right] = \mathbb{E}\left[L_n \frac{f(X_{n+1})}{g(X_{n+1})|\mathcal{F}_n}\right]$$

$$= L_n \mathbb{E}\left[\frac{f(X_{n+1})}{g(X_{n+1})}|\mathcal{F}_n\right] \qquad \text{(beacuse } L_n \text{ is } \mathcal{F}_n\text{-meas.)}$$

$$= L_n \mathbb{E}\left[\frac{f(X_{n+1})}{g(X_{n+1})}\right]$$

$$= L_n \underbrace{\int_{\mathbb{R}} \frac{f(X)}{g(X)} g(x)\, \mathrm{d}x}_{=1} = L_n.$$

This really is exhausting. Like it literally drains the life from me

1.2.3 Doob's decomposition

We may want to understand whether there is any relationship between a stochastic process and a martingale. In this part our set $\mathbb{T}$ will be $\mathbb{N}$.

Up to now we said that a process is adapted to $\mathcal{F}$ if X_n is $\mathcal{F}_n$-measurable.

Definition 1.2.3

The process $F = (F_n)_{n\geqslant 1}$ is $\mathcal{F}$-predictable if $F_0 \in \mathcal{F}_0$ and $\mathcal{F}_{n+1} \in \mathcal{F}_n$, for $\forall n \in \mathbb{N}$.

This means that the available information up to n is "enough" to have a bet in the period $n+1$. Some predictable processes are:

- any deterministic processes;

- consider two stopping times S, T of $\mathcal{F}$ and let $S \leqslant T$. Consider the random variable V in $\mathcal{F}_n$. Then

$$
\begin{array}{cccc}
(1) & (2) & (3) & (4) \\
V\mathbb{1}_{(S,T]} & V\mathbb{1}_{(S,\infty]} & \mathbb{1}_{(S,T]} & \mathbb{1}_{(0,T]}
\end{array}
$$

are predictable processes.

Proof

(2) Consider $F = V\mathbb{1}_{(S,\infty]}$, so that $F_n = \mathbb{1}_{(S,\infty]}(n)$. Consider

$$
\begin{aligned}
F_{n+1} &= V\mathbb{1}_{(S,\infty]}(n+1) \\
&= V\mathbb{1}_{\{S<n+1\}} \\
&= V\mathbb{1}_{\{S\leqslant n\}} \in \mathcal{F}_n
\end{aligned}
$$

so the process is $\mathcal{F}$-measurable and predictable.

(1) Since we know by hypothesis that $S \leqslant T$ then $V \in \mathcal{F}_S \subset \mathcal{F}_T$. This means that $V \in \mathcal{F}_T$. Hence, of a consequence,

$$
V\mathbb{1}_{(T,\infty]} - V\mathbb{1}_{(S,\infty]} = V\mathbb{1}_{(S,T]}
$$

is predictable.

(3) Take $V = 1$.

(4) Take $T = \infty$, $V = 1$. $\mathbb{1}_{(S,\infty]}$ is predictable. But then

$$
\mathbb{1}_{[0,S]} = \mathbb{1} - \mathbb{1}_{(S,\infty]}
$$

so it is predictable.

$\square$

Theorem 1.2.1

Doob's decomposition.
X is a stochastic process which is adapted to $\mathcal{F}$ and integrable. Then

1. it can be decomposed as
$$
X_n = X_0 + M_n + A_n, \qquad n \in \mathbb{N}
$$

   where:

   - M_n is a $\mathcal{F}$-martingale with $M_0 = 0$;
   - A_n is a predictable process with $A_0 = 0$.

2. The decomposition is unique up to equivalence.

3. If X_n is a sub-martingale then $\{A_n\}_{n\geqslant 0}$ increasing, while if X_n is a super-martingale then $\{A_n\}_{n\geqslant 0}$ is decreasing.

Proof

Put $A_0 = M_0 = 0$. Define M and A through their increments:

$$
\begin{aligned}
A_{n+1} - A_n &= \underbrace{\mathbb{E}\left[X_{n+1} - X_n | \mathcal{F}_n\right]}_{\mathcal{F}_n-\text{meas.}} \\
M_{n+1} - M_n &= (X_{n+1} - X_n) - (A_{n+1} - A_n).
\end{aligned}
$$

If we look at these quantities we see that A is predictable and M is martingale. Imagine now there is another decomposition: let

$$
X = X_0 + M' + A'
$$

be another decomposition. We must have

$$\cancel{X_0} + M' + A' = \cancel{X_0} + M + A \iff A - A' = M - M' = B.$$

Now B is a process and it is predictable and martingale (because it is the difference between two martingales). Since B is predictable *and* a martingale we have

$$B_{n+1} - B_n = \underbrace{\mathbb{E}[B_{n+1} - B_n | \mathcal{F}_n]}_{\text{predictable}}$$

$$= \underbrace{0}_{\text{martingale}}$$

$$\implies B_{n+1} = B_n = B_0 \text{ a.s.}, A = A' \text{ a.s.}, M = M' \text{ a.s.}$$

If X is a sub-martingale then we have

$$\mathbb{E}[X_{n+1} - X_n | \mathcal{F}_n] \geq 0$$

and this means that we have $A_{n+1} \geq A_n$ is increasing. $\qquad\square$

Consider the increments on X:

$$X_{n+1} - X_n = \underbrace{A_{n+1} - A_n}_{\text{prediction term}} + \underbrace{M_{n+1} - M_n}_{\text{innovation}}.$$

Engineers, much better people than us, often use this terms.

1.2.4 Important stochastic processes and related martingales

A quick note before starting:

Markov chains

Consider a stochastic process $\{X_n\}_{n \geq 1}$. All trials are on the same probability space $(E_n, \mathcal{E}_n) = (E, \mathcal{E})$. Define a **Markov kernel** on $(E, \mathcal{E})$

$$(E, \mathcal{E}) \mapsto (E, \mathcal{E})$$

such that

$$K_n(x_0, x_1, \ldots, x_n; A) = \mathbb{P}(x_n, A) \qquad \forall n \in \mathbb{N}, \ \forall (x_0, x_1, \ldots, x_n) \in E, A \in \mathcal{E}.$$

So the probability that governs the evolution is dependent *only* on the last state and nothing else. If $x_{n+1} = j, x_n = i$ then $\mathbb{P}(i, j) = p_{ij}$. The process $X = \{X_n\}_{n \geq 0}$ is said to be a **Markov chain** over $(\Omega, \mathcal{H}, \mathbb{P})$ with state space $(E, \mathcal{E})$, initial distribution μ and transition kernel P.

$$P = (p_{ij})_{i,j}$$

Now consider a Markov chain with transition probability matrix P such that

$$P \cdot \underbrace{f}_{\text{eigenvector}} = \underbrace{\lambda}_{\text{eigenvalue}} \cdot f.$$

This means that we have the following system of equations:

$$\begin{cases} \sum_j p_{ij} f(j) = \lambda f(i) & i = 1 \\ \vdots & i = 2 \\ \vdots & \ddots \end{cases}$$

We can clearly write these expressions in form of expectation:

$$\mathbb{E}[f(X_{n+1}) | X_n = i] = \lambda f(i).$$

We can also write it without fixing i:

$$\mathbb{E}[f(X_{n+1}) | X_n] = \lambda f(X_n). \qquad (\bullet)$$

Look at the last equation: it is not too different from the martingale property. Consider the ratio $\frac{f(X_n)}{\lambda^n}$ and the filtration generated by X $\sigma(X_0, X_1, \ldots, X_n)$. Thanks to the relation given by the equation ($\bullet$) it is a martingale with respect to $\sigma(X_0, X_1, \ldots, X_n)$.

Proof

Consider the expectation

$$
\mathbb{E}\left[\frac{f(X_{n+1})}{\lambda^{n+1}}\Big| X_0, \ldots, X_n\right] = \frac{1}{\lambda^n}\frac{1}{\lambda}\underbrace{\mathbb{E}\left[f(X_{n+1}|X_n)\right]}_{\lambda f(X_n)}
$$

$$
= \frac{1}{\lambda^n}\frac{1}{\lambda}\lambda f(X_n)
$$

$$
= \frac{1}{\lambda^n}f(X_n)
$$

$$
= Y_n
$$

where the first equality is thanks to Markov property. This is a general property for Markov chains. $\square$

Example 1.2.5

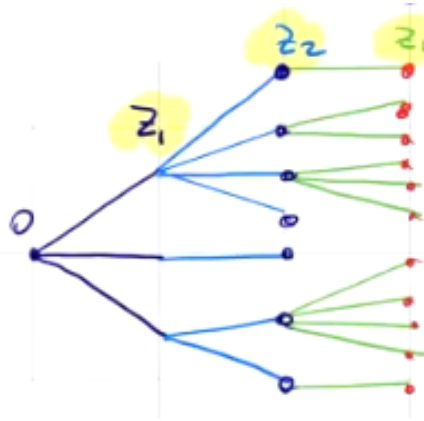**Branching processes**. We study the life of a generation of individual.

Figure 1.6: I don't even know anymore.

The offsprings have a distribution $\{p_k\}$. The mean number of offsprings per individual is $m = \sum k p_k$. Consider the i.i.id random variables $\{Z^{(n)}(i), n \geq 0\}$ describing the number of children of the i-th individual of generation n. Define

$$
Z_{n+1} = \begin{cases} Z^{(n)}(1) + Z^n(2) + \ldots + Z^{(n)}(Z_n) & \text{if } Z_n \neq 0 \\ 0 & \text{if } Z_n = 0 \end{cases}
$$

We immediately realize that this is a Markov process because every generation's number of children can be entirely derived from the previous generation's law or reproduction. I don't need to know whose child the individual is. So Z_n (number of individual of the n-th generation) gives rise to a Markov chain.

$$
\sum_{j=0}^{\infty} p_{ij} j = m \cdot i
$$

Is the equation that tells us the complete number of individuals. Now consider $Pf = mf$ with $f(j) = j$. So we have $Pj = mi$ where m plays the role of λ. The process $\{\frac{Z_n}{m^n}, \sigma(Z^{(0)}, \ldots, Z^{(n)})\}$ is a martingale.

Poisson martingale

Consider $\mathbb{R}_+$ as our index set. $\mathcal{F}$ is our filtration and we consider the counting process $N = (N_t)_{t \geq 0}$ (it counts the number of events up to time t). It has unit jumps and any path starts from 0 ($N_0(\omega) = 0$), is increasing and right

continuous.

Definition 1.2.4

The counting process N is said to be a **Poisson process** with rate λ with respect to $\mathcal{F}$ if it is adapted to $\mathcal{F}$ and

$$\mathbb{E}[f(N_{t+s} - N_s)|\mathcal{F}_s] = \sum_{k=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^k}{k!}f(k) \qquad \forall\, s, t \in \mathbb{R}_+, \forall \text{ positive } f \mapsto \mathbb{N}.$$

Theorem 1.2.2

Let N be a counting process. It is a Poisson process with rate λ with respect to $\mathcal{F}$ if and only if:

$$M_t = N_t - \lambda t$$

is a $\mathcal{F}$-martingale.

We only prove that M_t is a martingale if N_t is Poisson.

Proof

We know that

$$\begin{aligned}
\mathbb{E}[M_t|\mathcal{F}_s] &= \mathbb{E}[M_t - M_s + M_s|\mathcal{F}_s] \\
&= \mathbb{E}[M_t - M_s|\mathcal{F}_s] + M_s \\
&= \mathbb{E}[M_t - M_s] + M_s \\
&= \mathbb{E}[N_t - N_s + \lambda t + \lambda s] + M_s \\
&= \underbrace{\mathbb{E}[N_t - N_s]}_{\lambda(t-s)} - \lambda(t-s) + M_s \\
&= M_s
\end{aligned}$$

$\square$

We may use the Doob's decomposition theorem to get $N_t = M_t + \lambda t$ where λt is our predictable process. So the same result can be achieved from two different viewpoints.

1.2.5 Integration in discrete time

Let us consider two real-valued processes $M = (M_n)_n$ and $F = (F_n)_n$ and let us define

$$X_n = F_0 M_0 + (M_1 - M_0)F_1 + \ldots + (M_n - M_{n-1})F_n.$$

We say that $\{X_n\}$ is the integral of F with respect to M and we write

$$X_n = \int F\, \mathrm{d}M$$

where $\mathrm{d}M$ is a random signed measure. Remember the Lebesgue-Stieltjes integral? Me neither, but as long as M has bounded variation this is a Lebesgue-Stieltjes integral.

Yeah no wait hold your fucking horses.

What's up now?

First of all how is that shit an integral?? And what the fuck is a Lebesgue-Stieltjes integral?

I should have NEVER let you enroll in this master.

Professor Lods, I will exact my revenge on Università di Torino very soon.

Okay okay. So a little explanation is due since I actually never saw a Stieltjes integral. Too busy doing things, y'know. Things like, for example, sex. Heh. Yeah. Sex.

> **Revise with Kotatsu!**
>
> Oookay so very shortly: a Riemann-Stieltjes integral is a generalization of our dear old Riemann integral. It is notated as
>
> $$\int_a^b f(x)dg(x).$$
>
> Actually this is not too different from the concept of the Lebesgue integral. Basically is about partitioning the interval (a, b) and then summing for every partition (that is, for each i) each value of $f(x_i)$ multiplied for $[g(x_i) - g(x_{i-1})]$. Just take the limit for the partition to be infinitely small and voilà, le integràl c'est servie (I don't know French). El integral està servido (I know a little bit of Spanish). Omnibus integral paratum est ut facile intellegatur (I actually know Latin really well. Useless skills are my strength).
> Anyway what's important is: *we take the function $f(x)$ and use it to calculate how much the variation of $g(x)$ impacts on the total sum*. And if $g(x) = x$, well... We have our cute little Riemann integral!

Yeah after this smug declaration of sterile knowledge we are maybe ready to better understand how the discrete time integral is, well, an integral. Our function $g(x)$ is M so we consider the variation of M; since we are working in discrete time we will never reach a partition so fine it can be called continuous. So it's like... a *diet* Riemann-Stieltjes integral.

> **Theorem 1.2.3**
>
> Consider F bounded and predictable. Then if M is a martingale then X is a martingale; If M is a sub(super)-martingale then X is a sub(super)-martingale.

This means that... **we can't beat the system!**

I can beat something else though.

> **Example 1.2.6**
>
> Consider M_n as the price of a share at time n and F_n as the number of shares owned during $(n-1, n]$. Our profit will be
>
> $$(M_n - M_{n=1})F_n$$

and our total profit X_n gained during $(0, n]$ will be:

$$X_n = X_0 + \underbrace{\sum_{k=1}^{n}(M_k - M_{k-1})F_k}_{\text{discrete time integral}}$$

F_n is based on the knowledge in $n-1$ so it is predictable. The process M_n should be a martingale (otherwise if it is a sub/super-martingale everyone/no one will buy). So the total profit will also be a martingale! We can only choose our buying politics F_k, but there is no way to select a politics that will change a martingale in a super-martingale or sub-martingale.

Clearly this works in mean!

Proof

1. We have M being a martingale and $F_0, F_1, \ldots, F_n \in \mathcal{F}_n$ as well as $M_0, M_1, \ldots, M_n \in \mathcal{F}_n$. Therefore $X_n \in \mathcal{F}_n$ and X is adapted to $\mathcal{F}$.

2. We need to check whether the discrete time integral is a martingale. We know by hypothesis that F is bounded, so $F < b$ for some b. This implies

$$|X_n| < b(|M_0| + |M_1 + M_0| + \ldots + |M_n - M_{n-1}|)$$

Since M is a martingale and it is integrable, we get that X_n is bounded and integrable.

3. Consider
$$\mathbb{E}[X_{n+1} - X_n | \mathcal{F}] = \mathbb{E}[F_{n+1}(M_{n+1} - M_n) | \mathcal{F}_n]$$
since all the terms cancel out and only the last ones survive. But $F_{n+1} \in \mathcal{F}_n$ so we can take it out of the expectation:
$$F_n \underbrace{\mathbb{E}[M_{n+1} - M_n | \mathcal{F}_n]}_{=0} = 0.$$

$\square$

We know that using a policy that it is predictable it is impossible to beat the system. Finance bros try to overcome this possibility using stopping times. If my policy is not only based on a predictable process but I add the randomness of the time in which I decide to sell or buy can I break the curse of martingales and make shareholders want to suck my dick? Well, it depends.

1.2.6 Stopped martingales and Doob's stopping theorem

> Definition 1.2.5
>
> Define $M = (M_n)_{n \in \mathbb{N}}$ as a process and let T be a random time with values on $\overline{\mathbb{N}}$. The process
>
> $$X_n(\omega) = M_{n \wedge T}(\omega) = \begin{cases} M_n(\omega) & n < T(\omega) \\ M_T(\omega) & n > T(\omega) \end{cases}$$
>
> (where $n \wedge T$ is a truncated random time) is called M **stopped at** T.

As a consequence X is exactly the discrete time integral if $F = \mathbb{1}_{[0,T]}$:

$$X_n = \underbrace{M_0 F_0}_{0} + (M_1 + M_0)\mathbb{1}_{[0,T]}(1) + \ldots + (M_n - M_{n-1})\mathbb{1}_{[0,T]}(n).$$

The indicators only select the current time interval. If this is the case we can observe that $F_{[0,T]}$ is bounded, positive and predictable. Hence if M is a martingale the theorem applies with this special choice of M and we can write the result as a different theorem:

> **Theorem 1.2.4**
>
> Let T be a stopping time and let X be the process M stopped at T. If M is a martingale then so is X (the same holds for sub-martingales and super-martingales).

So we cannot determine a policy based on stopping times that can change the nature of our martingale. In the remote case in which you are interested in this you can read Williams - Introduction to martingales.

A further theorem about this is **Doob's stopping theorem**. For a martingale we know

$$\mathbb{E}[X_t - X_s|\mathcal{F}_s] = 0.$$

The question is: is this true also if s and t are substituted by stopping times $S, T, S \leqslant T$?

> **Theorem 1.2.5**
>
> Let M be adapted to $\mathcal{F}$. Then the following are equivalent:
>
> (1) M is a submartingale;
>
> (2) for every bounded stopping time $S \leqslant T$ the random variables M_S and M_T are integrable and
>
> $$\mathbb{E}[M_T - M_S|\mathcal{F}_S] \geqslant 0;$$
>
> (3) for each pair of bounded stopping times the random variables M_S and M_T are integrable and
>
> $$\mathbb{E}[M_T - M_S] \geqslant 0.$$

> **Remark**
>
> If M is a martingale the theorem can be read in a different way:
>
> $$\mathbb{E}M_T = \mathbb{E}M_S = \mathbb{E}M_0.$$
>
> Previously, $\mathbb{E}M_n = \mathbb{E}M_{n-1} = \mathbb{E}M_0$.

> *Proof*
>
> To prove the theorem we need to show that from condition 1 follows 2 from which follows 3 from which follows 1.
>
> $1 \to 2$ by hypothesis M is a sub-martingale and our thesis is that if $S(\omega) < T(\omega) < n$ (because we asked for bounded times) then:
>
> (a) M_S and M_T are integrable;
> (b) $\mathbb{E}[M_T - M_S|\mathcal{F}_S] \geqslant 0$.
>
> We know that S and T are bounded by n. Let V be a positive bounded random variable and define $F = V\mathbb{1}_{(S,T]}$ and use it in the discrete time integral:
>
> $$X_n = \underbrace{M_0 F_0}_{X_0} + (M_1 - M_0)\underbrace{F_1}_{V\mathbb{1}_{\{1\in(S,T]\}}} + \ldots + (M_n - M_{n-1})F_n$$
>
> So we have
>
> $$X_n - X_0 = V(M_T - M_S)$$
>
> So X_n is a sub-martingale. Take $V = 1$ and $S = 0$: now we have
>
> $$X_n \quad - \quad X_0 \quad = M_T$$
> $$\downarrow \qquad \qquad \downarrow$$
> $$\text{int.} \qquad \text{int.} \quad \implies M_T \text{ is integrable.}$$
>
> Now take $V = 1$ and $T = n$ so that we get
>
> $$X_n \quad - \quad X_0 \quad = M_n \quad - M_S$$
> $$\downarrow \qquad \quad \downarrow \qquad \downarrow$$
> $$\text{int.} \qquad \text{int.} \quad \text{int.} \quad \implies M_S \text{ is integrable.}$$

We recall that $V \in \mathcal{F}_S$ and we use the defining property for $\mathbb{E}(\cdot|\mathcal{F}_S)$. So we can write

$$\mathbb{E}V\mathbb{E}(M_T - M_S|\mathcal{F}_S) \underset{\text{def. prop.}}{=} \mathbb{E}V(M_T - M_S)$$

$$\underset{\text{discr. time int.}}{=} \mathbb{E}[X_n - X_0]$$

$$\underset{\text{proved above}}{\geqslant} 0$$

and this is true $\forall V > 0, V < b, V \in \mathcal{F}_s$. Hence

$$\mathbb{E}(M_T - M_S|\mathcal{F}_S) \geqslant 0$$

So $1 \to 2$.

$2 \to 3$ We can use the tower rule. Take the expectation of point 2:

$$\mathbb{E}[\mathbb{E}(M_T - M_S|\mathcal{F}_S)] = \mathbb{E}[M_T - M_S] \geqslant 0.$$

$3 \to 1$ Let 3 hold, so that $\mathbb{E}[M_T - M_S] \geqslant 0$. Choose $T = n$ and $S = 0$. Then M_n is integrable. Move to adaptness: this holds by hypothesis. Move to the martingale inequality:

$$\mathbb{E}[M_n - M_m|\mathcal{F}_m] \geqslant 0.$$

Note that this is equivalent to prove

$$\mathbb{E}\mathbb{1}_H\mathbb{E}[M_n - M_m|\mathcal{F}_m] \geqslant 0 \qquad H \in \mathcal{F}_m, 0 \leqslant m \leqslant n.$$

Fix H, m, n and define

$$S(\omega) = m \qquad T(\omega) = n\mathbb{1}_H(\omega) + m\mathbb{1}_{\{\Omega \setminus H\}}(\omega)$$

The indicators are non-zero in complementary instances. Notice that:

(a) S is a fixed time so it is a stopping time;

(b) $S \leqslant T \leqslant n$ by definition of S and T because the indicators are non-zero in complementary instances;

(c) $T \geqslant S$ is a foretold time by $S = m$;

(d) $H \in \mathcal{F}_S$ by definition.

So we can write $M_T - M_S = \mathbb{1}_H(M_n - M_m)$ where $M_T - M_S \geqslant 0$ by hypothesis. This means that we have

$$\underbrace{\mathbb{E}[\mathbb{1}_H\mathbb{E}(M_n - M_m|\mathcal{F}_m)]}_{\mathbb{E}[M_n - M_m|\mathcal{F}_m] \geqslant 0} \geqslant 0$$

$\square$

1.2.7 Oscillations

We will talk about oscillations in a strip $(a.b)$ during a time interval. Consider an adapted process M and two real numbers $a < b \in \mathbb{R}$. We want to set a level a and a level b. First, by convention, fix $T_0 = -1$. Then define:

$$S_k := \inf\{n > T_{k-1} : M_n \leqslant a\}$$
$$T_l := \inf\{n > S_k : M_n \geqslant b\}.$$

So it is the first time in which the process crosses the threshold a or b after having crossed the other threshold.
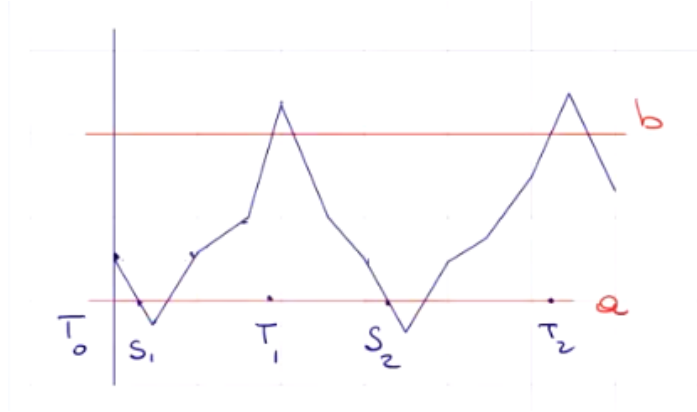
Figure 1.7: Up and downs. Just like my consumption of cocaine and my mood respectively on an average tuesday evening.

We call S_k **downcorssing time** while we call T_k **upcrossing time**.
Now consider the sequence $(T_0, S_1, T_1, S_2, T_2, \ldots)$. It is an increasing sequence of stopping times. If we want to focus on the number of oscillations we may count how many times we go from an upcrossing to a downcrossing time (or vice-versa). Say we want to count the upcrossing. We can define a quantity

$$U_n(a,b) = \sum_{k=1}^{\infty} \mathbb{1}_{(0,n]}(T_k)$$

which of course is a counting process.

Example 1.2.7

We are still with our financebros. Suppose that (M_n) is the price of an asset at time n. We want to buy when the price is below a at time S_i and sell when it is above b at time T_i. In $(0, n]$ we have $U_n(a, b)$ cycles of buying and selling so our strategy could consists in holding a number F_n of shares during period $(m - 1, m]$. This means introducing

$$F_m = \sum_{k=1}^{\infty} \mathbb{1}_{(S_k, T_k]}$$

with $F_0 = 0$. We can thus trace the evolution of our capital with the discrete-time integral:

$$X = \int F \, dM$$

And the profit during $(0, n]$ will be $X - X_0$.
The profit is at least

$$(b - a)U_n(a, b).$$

Proposition 1.2.2

If M is a sub-martingale with respect to its natural filtration then

$$(b - a)\mathbb{E}U_n(a, b) \leqslant \mathbb{E}\left[(M_n - a)^+ - (M_0 - a)^+\right].$$

We wanted to find a bound for our profit, but our profit is a stochastic quantity: so it's only natural to think about the expectation to give a bound to the number of expected up/downcrossing.
Observe that the number of upcrossings does not depend on the value of T_0 that we fix.

Proof

Choose $a = 0$. Consider hence the process $(M - a)^+$ that is sub-martingale (if M is a sub-martingale). Take $M \geqslant 0$ and let

$$F_n - \sum_{k=1}^{\infty} \mathbb{1}_{(S_k, T_k]}(n)$$

and consider

$$X = \int F \, dM$$

like in our example. We know that F is predictable since by definition $F_{k+1} \in \mathcal{F}_k$. Consider the expectation of the increment

$$\mathbb{E}[X_{k+1} - X_k | \mathcal{F}_k] = \mathbb{E}\left[F_{k+1}(M_{k+1} - M_k) | \mathcal{F}_k\right].$$

But since F_{k+1} is predictable we can take it out the expectation:

$$F_{k+1}\mathbb{E}\left[M_{k+1} - M_k | \mathcal{F}_k\right].$$

But since F_{k+1} is an indicator we know it is $\leqslant 1$:

$$\mathbb{E}[X_{k+1} - X_k | \mathcal{F}_k] \leqslant \mathbb{E}\left[M_{k+1} - M_k | \mathcal{F}_k\right].$$

Now take the expectation of both sides:

$$\mathbb{E}[X_{k+1} - X_k] \leqslant \mathbb{E}\left[M_{k+1} - M_k\right].$$

If we sum these inequalities over k we get:

$$\mathbb{E}\left[X_n - X_0\right] \leqslant \mathbb{E}\left[M_n - M_0\right].$$

So we now get that

$$bU_n(a, n) \leqslant \mathbb{E}\left[X_n - X_0\right] \leqslant \mathbb{E}\left[M_n - M_0\right]$$

but given that $a = 0$ we get that

$$bU_n(0, b) \leqslant \mathbb{E}[M_n - M_0].$$

Clearly we have to take the positive part. □

This characterizes our martingale and its boundedness. The number of oscillations of a sub-martingale is bounded! The next question is: can we say anything about the behaviour of maximum/minimum of a martingale or sub-martingale?

> **Remark**
>
> Consider a sequence $\{X_n\}$ of independent random variables with $\mathbb{E}X_n = 0$, $S_n = \sum X_i$. We proved that
>
> $$a^2 \mathbb{P}(\max_{k \leqslant n} |S_k| > a) \leqslant \operatorname{Var} S_n$$
>
> and we called this **Kolmogorov's inequality**.

What we are doing here is considering a random walk whose jumps have 0 mean. We wonder wether it is above the level a as seen in figure 1.8.
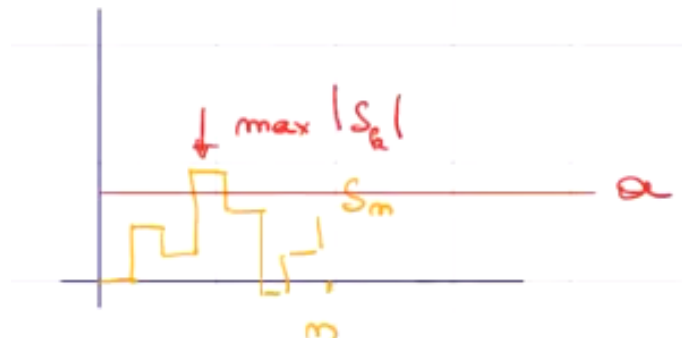

Figure 1.8: The maximum of the random walk.

About this, we know that $\{S_n\}$ is a $\mathcal{F}$-martingale but in proving the Kolmogorov's inequaility we never talked about the martingale property! We could improve this inequality using the **Doob's martingale inequality**. The

problem is that we can prove very general inequalities that hold true for any random variables but specifying more characteristic we can obtain stricter bounds. In this framework let's define

$$M_n^\star = \max_{k \leqslant n} M_k$$
$$m_n^\star = \min_{k \leqslant n} M_k$$

as current maximum and current minimum of M.

Theorem 1.2.6

Take M as a sub-martingale. For $b > 0$ it holds:

1. $b\mathbb{P}(M_n^\star \geqslant b) \leqslant \mathbb{E}\left[M_n \mathbb{1}_{\{M_n^\star \geqslant b\}}\right] \leqslant \mathbb{E}\left[M_n^+\right]$;

2. $b\mathbb{P}(m_n^\star \leqslant -b) \leqslant -\mathbb{E}M_0 + \mathbb{E}\left[M_n \mathbb{1}_{\{m_n^\star \geqslant b\}}\right] \leqslant \mathbb{E}M_n^+ - \mathbb{E}M_0$.

So we can further bound the result looking into the property of M_n.

Example 1.2.8

Prof. Sacerdote fucked up so we now need to define the brownian motion (or Weiner process).

Definition 1.2.6

A real-valued stochastic process $B = (B_t)_{t \geqslant 0}$ is called **brownian motion** if:

1. the index set is $\mathbb{R}^+$;

2. $B_0(\omega) = 0$ for almost all ω;

3. $B_{t_n} - B_{t_{n-1}}, \ldots, B_{t_1} - B_{t_0}$ are independent for $\forall\, 0 = t_0 < t_1 < \ldots < t_n < \infty$;

4. $B_t - B_s \sim B_{t+b} - B_{s+b}$ for every $0 \leqslant s < t < \infty \quad \forall n > -s$;

5. $B_t - B_s \sim N(0, t - s)$;

6. $t \mapsto B_t(\omega)$ are continuous for every ω.

The brownian motion is in itself a martingale, but there is a class of martingales strictly related to it:

$$M_t = \exp\left\{\lambda B_t - \frac{1}{2}\lambda^2 t\right\}, \qquad t \in \mathbb{R}^+.$$

Now we can get to the actual example: if B is a brownian motion, we have

$$\mathbb{P}(\sup_{0 \leqslant t \leqslant \theta} B_t \geqslant b) \leqslant \exp\left\{-\frac{b^2}{s\theta}\right\}.$$

The sample paths of brownian motions are extremely irregular. We are asking with which probability the max of our process will be over b at time θ.
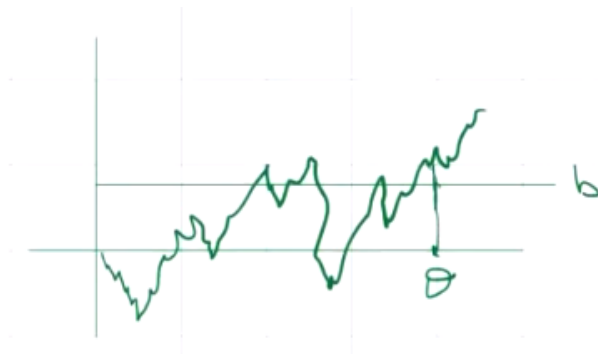

Figure 1.9: Check out Autechre. A staple in electronic experimental music.

I am basically asking if the maximum attained is above or below b. How can we use the Doob's inequality?

Consider

$$\mathbb{P}(\sup B_t \geqslant b) = \mathbb{P}(\sup e^{\lambda B_t} \geqslant e^{\lambda b})$$

and using Doob's inequality

$$\mathbb{P}(\sup B_t \geqslant b) = \mathbb{P}(\sup e^{\lambda B_t} \geqslant e^{\lambda b})$$

$$\leqslant \frac{\overbrace{\mathbb{E}\left[e^{\lambda B_\theta - \frac{\lambda^2}{2}\theta}\right]}^{\text{martingale}}}{e^{\lambda b}e^{-\frac{\lambda^2}{2}\theta}}$$

$$= e^{-\lambda b + \frac{\lambda^2}{2}\theta} \qquad \forall \lambda > 0$$

$\square$

Now we have to prove Doob's inequality.

We introduce 2 stopping times:

$$\begin{aligned} T &= \inf\{n \geqslant 0 : M_n \geqslant b\} \\ S &= \inf\{n \geqslant 0 : M_n \leqslant -b\}. \end{aligned}$$
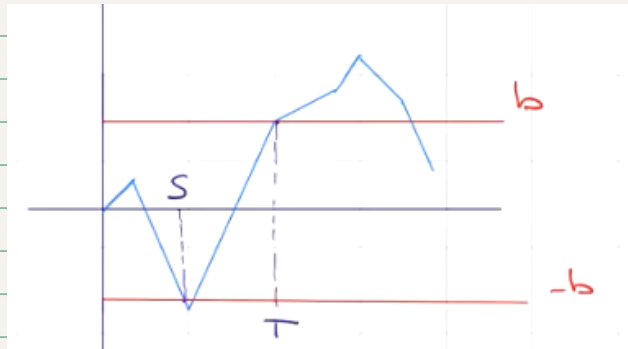

Figure 1.10: Also Aphex Twin. Really, Autechre and Aphex Twin are probably the best example of the boom of experimentation in popular electronic music between 1995 and 2005.

Consider the current maximum and minimum above the level b:

$$\{M_n^\star \geqslant b\} = \{T \leqslant n\} \qquad \{m_n^\star < -b\} = \{S \leqslant n\}.$$

Fix b and n. When $\{T \leqslant n\}$ we have

$$M_{T \wedge n} = M_T \geqslant b.$$

Multiply by $\mathbb{1}_{\{T \leqslant n\}}$:

$$b\mathbb{1}_{\{T \leqslant n\}} \leqslant M_{T \wedge m}\mathbb{1}_{\{T \leqslant n\}}.$$

Using Doob's stopping theorem we know that

$$\mathbb{E}\left[M_T - M_S | \mathcal{F}_S\right] \geqslant 0$$

and using $S = T \wedge n$ and $T = n$ we obtain

$$\mathbb{E}\left[M_n | \mathcal{F}_{T \wedge n}\right] > M_{T \wedge n}.$$

So summing it all up:

$$b\mathbb{1}_{\{T \leqslant n\}} \leqslant M_{T \wedge m}\mathbb{1}_{\{T \leqslant n\}}$$

$$\leqslant \mathbb{1}_{\{T \leqslant n\}}\mathbb{E}\left[M_n | \mathcal{F}_{T \wedge n}\right]$$

$$\mathbb{E}\left[M_n\mathbb{1}_{\{T \leqslant n\}} | \mathcal{F}_{T \wedge n}\right].$$

Now take the expectation:

$$b\mathbb{P}(T \leqslant n) = b\mathbb{P}(M_n^\star \geqslant b)$$
$$\leqslant \underbrace{\mathbb{E}\left[M_n \mathbb{1}_{\{T \leqslant n\}}\right]}_{\mathbb{E}\left[M_n \mathbb{1}_{\{M_n^\star \geqslant b\}}\right]}$$
$$\leqslant \mathbb{E}(M_n^+).$$

For the minimum we work on $\{S \leqslant n\}$ and we get

$$M_{S \wedge n} = M_S \mathbb{1}_{\{S \leqslant n\}} + M_S \mathbb{1}_{\{S > n\}}$$
$$\leqslant -b\mathbb{1}_{\{S \leqslant n\}} + M_n \mathbb{1}_{S \geqslant n}.$$

Now take the expectation:

$$\mathbb{E}M_{S \wedge n} \leqslant -b\mathbb{P}(m_n^\star \leqslant -b) + \mathbb{E}\left[M_n \mathbb{1}_{\{S > n\}}\right].$$

So that we get

$$\mathbb{P}(m_n^\star \leqslant -b) \leqslant -\mathbb{E}M_{S \wedge n} + \mathbb{E}[M_n \mathbb{1}_{\{S > n\}}].$$

Now use Doob's stopping theorem with $T = 0$ and $S = S \wedge n$ so that $\mathbb{E}M_0 \leqslant \mathbb{E}M_{\{S \wedge n\}}$. This gets us our result:

$$b\mathbb{P}(m_n^\star - b) \leqslant -\mathbb{E}M_0 + \mathbb{E}\left[M_n \mathbb{1}_{\{m_n^\star\}}\right]$$
$$\leqslant \mathbb{E}M_n^+ - \mathbb{E}M_0.$$

$\square$

Remark

If M is a martingale then $|M|^p$ is a sub-martingale for $p \geqslant 1$. If $M_n \in L^p \forall n$ we can apply Doob's inequality.

Corollary

If M is martingale in L^p for some $p \geqslant 1$ then for $b > 0$ we have that

$$b^p \mathbb{P}(\max_{k \leqslant n} |M_k| > b) \leqslant \mathbb{E}|M_n|^p.$$

There are also other bounds:

- $b\mathbb{P}(\max_{k \leqslant n} |M_k > b) \leqslant 2\mathbb{E}M_n^+ - 3M_0$;

- **Doob's norm inequality**: if M is a martingale in L^p, $p \geqslant 1$ and q is the exponent conjugate to p $\left(\frac{1}{p} + \frac{1}{q} = 1\right)$ then

$$\mathbb{E} \max_{k \leqslant n} |M_k|^p \leqslant q^p \mathbb{E}|M_n|^p.$$

- Consider L^2-bounded martingales characterized by final coordinate X with $\mathbb{V}\mathrm{ar}\, X = \sigma^2$ (that is I am fixing the variance of the last value I consider). We want to assess the variability of this process.

Theorem 1.2.7

Dubin & Schwartz 1998: it holds

1. $\mathbb{E}\left[\max_{0 \leqslant T \leqslant t} M_T\right] \leqslant \sigma$;

2. $\mathbb{E}\left[\max_{0 \leqslant T \leqslant t} |M_T|\right] \leqslant \sigma\sqrt{2}$.

Moreover, there exist suitable martingales for which this bound is attained and is scrict.

1.2.8 Limit of sequences of random variables

Consider a sequence of random variables $\{X_n\}$ with independence between the random variable. Then we know that if $\lim X_n$ exists then the limit is a constant. Why?

> **Remark**
>
> If the $\lim_n X_n$ exists than it belongs to the tail σ-algebra τ. Since the random variables are independent then we can apply the 0-1 Kolmogorov law, so the probability that the random variable is infinity is either 0 or 1 but if this is the case then it must be a constant.

This is about the behavior in the limits of a sequence of random variables with independence between $\{X_n\}$. Then if the limit exists it is a constant, since $\lim_n X_n$ belongs to the tail σ-algebra of the sequence. Independency between random variables tells us that we can apply Kolmogorov's 0-1 law to say that $\mathbb{P}(\lim_n X_n = \infty)$ is either 0 or 1... but if we somehow know that the limit exists then it can't be ∞ so its probability of being ∞ is 0 and therefore it is a constant almost surely! This is cool because it tells us that the limit is either infinite or finite. If $\{X_n\}$ is a submartingale consider the limit

$$\lim_{n \to \infty} X_n = X_\infty$$

What the fuck? What is this object? Consider the case in which $\mathbb{T} = \mathbb{N}$. We know that the expected value $\{\mathbb{E}X_n\}$ is an increasing sequence if the process is a sub-martingale. We have our first theorem about convergence:

> **Theorem 1.2.8**
>
> Let X be a sub-martingale. If (and note that is a sufficient condition)
>
> $$\sup_n \mathbb{E}X_n^+ < \infty$$
>
> Then
>
> 1. $\{X_n\}$ converges a.s.;
>
> 2. $\{X_n\}$ converges to an integrable random variable.

Before proving think about it.

1. The condition of this theorem requires that X^+ is in L^1 and is bounded. If X is a sub-martingale then also X^+ is a sub-martingale... But this means that $\mathbb{E}X_n^+$ is an increasing sequence. Our condition gives a bound to the increase of this sequence! But this means that I am trying to avoid the divergence of the expected value (which means: trying to control the rate of our martingale).

2. Now think about a negative sub-martingale: this would verify the condition so there is no problem about it.

3. Now think about a positive super-martingale: it is only a negative sub-martingale with another sign so it converges.

4. Consider now a positive or bounded martingale or a negative martingale. All of these cases converge.

5. $\sup \mathbb{E}X_n^+ < \infty \iff \sup \mathbb{E}|X_n| \leqslant \infty$. Let $\{X_n\}$ be a sub-martingale such that

$$\mathbb{E}X_n > \mathbb{E}X_0 \qquad \forall\, n.$$

But now

$$\mathbb{E}|X_n| = \mathbb{E}X_n^+ - \mathbb{E}X_n^- = \mathbb{E}X_n^+ + \mathbb{E}_n^+ - \mathbb{E}X_n$$

So

$$\mathbb{E}X_n^+ \leqslant \mathbb{E}|X_n| = 2\mathbb{E}X_n^+ - \mathbb{E}X_n$$
$$\leqslant 2 \underbrace{\mathbb{E}X_n^+}_{<\infty} - \mathbb{E}X_0 \leqslant \infty$$

So $\mathbb{E}|X_n| \leqslant \infty$.

We prove the theorem by contradiction. Pick an outcome ω and suppose that $\{X_n(\omega)\}$ is a numerical sequence that has not a limit. But if it doesn't have a limit, then

$$\exists \inf \lim \neq \sup \lim \qquad \inf \lim < \sup \lim.$$

So there exist at least 2 rationals $a < b$ such that

$$\inf \lim < a < b < \sup \lim.$$

The sequence $\{X_n(\omega)\}$ crosses (a, b) ∞ many times. Now take the union over rational a and b, $a < b$ of the sets

$$\{U(a, b) = \infty\}$$

with $U(a, b) = \lim_{n \to \infty} U_n(a, b)$. Our aim is now to show that $U(a, b) \leq \infty$ almost surely to get a contradiction.

Fix a, b. We know that $U_n(a, b)$ is increasing with n. Now consider

$$(b - a)\mathbb{E}U(a, b) = (b - a)\mathbb{E} \lim U_n(a, b)$$
$$= (b - a) \lim_{\text{monotone conv.}} \mathbb{E}U_n(a, b)$$
$$\leq \sup_{\text{upcross inequalities}} \mathbb{E}(X_n - a)^+$$
$$\leq \sup \mathbb{E}X_n^+ + |a|.$$

So this means that $\mathbb{E}U(a, b) < \infty$. But this is a contradiction, so it exists a limit $X_n = X_\infty$ a.s..
Now consider the second part of the theorem:

$$\mathbb{E}|X_\infty| = \mathbb{E} \lim \inf |X_n|$$
$$\leq \lim_{\text{Fatou's lemma}} \inf \mathbb{E}|X_n|$$
$$\leq 2 \sup \mathbb{E}X_n^+ - \mathbb{E}X_0 \leq \infty$$

so the limit is integrable. $\square$

Here everything is related to integrability conditions but this is a sufficient condition. What if we want a sufficient and necessary condition? Well, consider $\sup_n \mathbb{E}X_n^+ \leq \infty$. We want to control the tail uniformly not only on the tail but on the whole process.
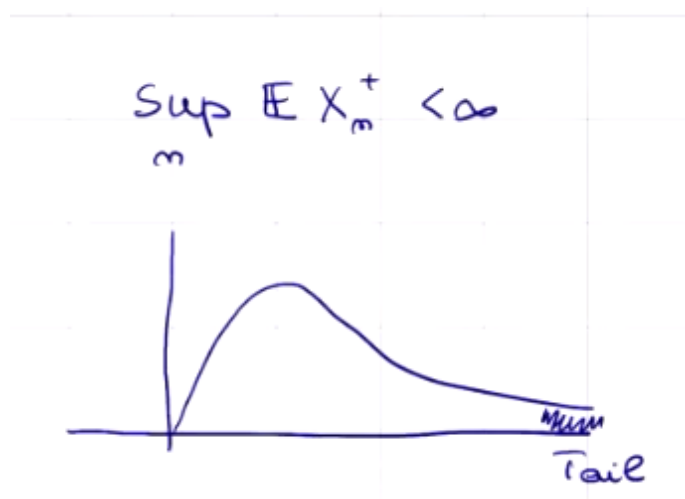

Figure 1.11: It's astonishing how much a bad teacher can have you hate a subject.

We recall that we already introduced a strong condition for uniform integrability. w We will need:

1. a collection $\mathcal{K}$ of real random variables is said to be uniformly integrable if

$$k(b) = \sup_{X \in \mathcal{K}} \mathbb{E}|X| \mathbb{1}_{\{X > b\}} \xrightarrow[b \to \infty]{} 0.$$

2. If $\mathcal{K}$ is dominated by an integrable random variable Z then it is uniformly integrable.

3. uniform integrability implies L^1-boundedness but not the converse.

4. If $\mathcal{K}$ is L^p-bounded for some $p > 1$ then it is uniformly integrable.

Lemma 1.2.1

Let Z be an integrable random variable. Then

$$\mathcal{K} = \{X : X = \mathbb{E}(Z|\mathcal{G})\}$$

for some sub-σ-algebra $\mathcal{G}$ of $\mathcal{H}$ is uniformly integrable.

Proposition 1.2.3

Let Z be an integrable random variable. Define

$$X_t = \mathbb{E}(Z|\mathcal{F}_t) \qquad t \in \mathbb{T}.$$

This means that $\{X_t\}$ is a uniformly integrable $\mathcal{F}$-martingale.

Theorem 1.2.9

Let $\{X_n\}$ be a sequence of real-valued random variables. The following are equivalent:

1. it converges in L^1;

2. it converges in probability and it is uniformly integrable.

We can now prove the theorem about the convergence of sub-martingales.

Theorem 1.2.10

Let X be a sub-martingale. We have that X converges almost surely and in L^1 if and only if it is uniformly integrable. Moreover, if it is so, setting
$$X_\infty = \lim X_n$$
extends X to a sub-martingale
$$\overline{X} = (X_n)_{n \in \overline{\mathbb{N}}}.$$

We only prove the first part of the theorem.

Proof

Necessity. If X converges in L^1 by the theorem above it is uniformly integrable.
Sufficiency. If X is uniformly integral then it is L^1-bounded for the property above. So our previous theorem holds and the martingale converges almost surely with X_∞ integrable. Furthermore, for the property above, it also converges in L^1. □

Is there a way to get a uniformly integrable sub-martingale?

Theorem 1.2.11

A process $M = (M_n)_{n \in \mathbb{N}}$ is a uniformly integrable martingale if and only if for some integrable random variable Z
$$M_n = \mathbb{E}[Z|\mathcal{F}_n] \qquad n \in \mathbb{N}. \tag{•}$$
If so it converges almost surely and in L^1 to the integrable random variable

$$M_\infty = \mathbb{E}[Z|\mathcal{F}_\infty].$$

Corollary

For every integrable random variable Z we have

$$\mathbb{E}(Z|\mathcal{F}_n) \xrightarrow{\text{a.s.}} \xrightarrow{L^1} \mathbb{E}(Z|\mathcal{F}_\infty).$$

Theorem 1.2.12

Let Z be an integrable random variable and let

$$M_n = \mathbb{E}(Z|\mathcal{F}_n)_{n \in \overline{\mathbb{N}}}.$$

For every stopping time T define

$$M_T = \mathbb{E}[Z|\mathcal{F}_T]$$

and for arbitrary stopping times S and T we get

$$\mathbb{E}[M_T|\mathcal{F}_S] = M_{S \wedge T}.$$

This lets us rethink Doob's theorem.

Theorem 1.2.13

If S and T are arbitrary stopping times such that $S \leqslant T$ then

$$\mathbb{E}[M_T|\mathcal{F}_S] = M_S$$

for an uniformly integrable martingale.

The dominated convergence theorem requires adaptness.

Theorem 1.2.14

Hunt's dominated convergence theorem.
Let $\{X_n\}$ be dominated by an integrable random variable and suppose that exists

$$X_\infty = \lim X_n$$

So $(\mathbb{E}_n X_n)_n$ converges to $\mathbb{E}X_\infty$ almost surely and in L^1.

To prove we should use Borel-cantelli somehow.

Chapter 2

Miscellanea: an integration (eheh)

2.1 Gaussian random variables

.

Definition 2.1.1

A **standard Gaussian random variable** (of one dimension) is a random variable Z with density

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \qquad z \in \mathbb{R}.$$

Definition 2.1.2

A **Gaussian random variable** with parameters μ and σ^2 is a random variable X with density

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbb{R}.$$

Proposition 2.1.1

We have $\mathbb{E}X = \mu$, $\operatorname{Var} X = \sigma^2$ and in particular $\mathbb{E}Z = 0$, $\operatorname{Var} Z = 1$.

Proposition 2.1.2

Let $Z \sim \mathcal{N}(0,1)$ and $\mu, \sigma \in \mathbb{R}, \ \sigma > 0$. Then

$$X := \mu + \sigma Z \sim \mathcal{N}(\mu, \sigma^2).$$

Proposition 2.1.3

The characteristic function of X is

$$\mathbb{E}\left[e^{i\lambda X}\right] = e^{i\lambda\mu} \cdot e^{-\frac{1}{2}\sigma^2\lambda^2}$$

if and only if we have $X \sim \mathcal{N}(\mu, \sigma^2)$.

Let's generalize this to n-dimensional random variables.

Definition 2.1.3

Let $X = (X_1, \ldots, X_n)$ be a n-dimensional vector of random variables. We say that X is Gaussian if for every vector $r \in \mathbb{R}^n$ we have that $r \cdot X = \sum_{i=1}^{n} r_i X_i$ is Gaussian.

> **Proposition 2.1.4**
>
> Let $Z = (Z_1, \ldots, Z_k)$ be k-dimensional vector of independent random variables $Z_i \sim \mathcal{N}(0,1)$. Let $\mu \in \mathbb{R}^n$ and A be a $(n \times k)$-dimensional matrix $A = (a_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,k}}$. We set
>
> $$\underset{(n \times 1)}{X} := \underset{(n \times 1)}{\mu} + \underset{(n \times k)}{A}\ \underset{(k \times 1)}{Z}.$$
>
> In particular,
>
> $$X_i := \mu_i + \sum_{j=1}^{k} a_{ij} Z_j.$$
>
> In this case, X is Gaussian.

The n-dimensional Gaussian random variables have the following properties:

1. the characteristic function of X is given by

$$\mathbb{E}\left[e^{ir \cdot X}\right] = \mathbb{E}\left[e^{i \sum_{i=1}^{n} r_i X_i}\right] = e^{ia(r) - \frac{1}{2}b(r)}, \qquad r \in \mathbb{R}^n$$

where

$$a(r) := \mathbb{E}[r \cdot X]$$
$$b(r) := \mathbb{V}\mathrm{ar}\,(r \cdot X).$$

Yeah, why not. Why not making my life even more miserable.

2. Setting $\mu_i := \mathbb{E}X_i$ and $\Gamma_{ij} = \mathbb{C}\mathrm{ov}\,(X_i, X_j)$, then

$$a(r) = r \cdot \mu \qquad \text{and} \qquad b(r) = \underset{(n \times 1)}{r} \cdot \underset{(n \times n)}{\Gamma}\ \underset{(n \times 1)}{r^\top}.$$

What the fuck does this mean? The Gaussian vector X is determined by its mean μ and covariance matrix Γ, because they determine the characteristic function of X. We write

$$X \sim \mathcal{N}(\mu, \Gamma)$$

where X is a n-dimensional vector (and also μ).

> **Proposition 2.1.5**
>
> Any Gaussian random variable X can be written as $X = \mu + \mathbf{A}Z$ with $\mu \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times k}$ and $Z = (Z_i, \ldots, Z_k)$ with $Z_i \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$

Let's now talk about independence. Since any Gaussian vector X can be written as $X = \mu + \mathbf{A}Z$ where the $Z_i \perp\!\!\!\perp Z_j, i \neq j, i, j = 1, \ldots, k$, what can we say about the independence of $X_i, X_j, i\, i, j = 1, \ldots n$? To see this, calculate $\mathbb{C}\mathrm{ov}\,(X_i, X_j)$:

$$
\begin{aligned}
\mathbb{C}\mathrm{ov}\,(X_i, X_j) &= \mathbb{E}\left[(Xi - \mathbb{E}X_i)(X_j - \mathbb{E}X_j)\right] \\
&= \mathbb{E}\left[(X_i - \mu_i)(X_j - \mu_j)\right] \\
&= \mathbb{E}\left[(\mu_i + \sum_{l=1}^{k} a_{il} z_l - \mu_i)(\mu_j + \sum_{m=1}^{k} a_{jm} z_m - \mu_j)\right] \\
&= \mathbb{E}[(\sum_{l=1}^{k} a_{il} z_l)(\sum_{m=1}^{k} a_{jm} z_m)] \\
&= \sum_{l=1}^{k} \sum_{m=1}^{k} a_{il} a_{jm} \mathbb{E}\,[Z_l Z_m] \\
&= \sum_{l=1}^{k} a_{il} a_{jl} \underbrace{\mathbb{E}(Z_l^2)}_{=1 \text{ since } Z_l \sim \mathcal{N}(0,1)} = (\mathbf{A}\mathbf{A}^\top)_{ij}.
\end{aligned}
$$

Since it is necessary (but not sufficient) that $\mathbb{C}\text{ov}\,(X_i, X_j) = 0$ to have $X_i \perp\!\!\!\perp X_j$ then X_i is not independent from X_j in general.

> **Note the notation!**
>
> We proved that $\mathbb{C}\text{ov}\,(X_i, X_j) = (\mathbf{AA}^\mathsf{T})_{ij}$, hence with the notation above we get
> $$\underset{n\times n}{\boldsymbol{\Gamma}} = \underset{n\times k}{\mathbf{A}}\,\underset{k\times n}{\mathbf{A}^\mathsf{T}}.$$

> **Proposition 2.1.6**
>
> If Γ is invertible (i.e. $\mathsf{rank}(\Gamma) = n$) then $X = \mu + \mathbf{A}Z$ has density given by
> $$f_X(x) = \frac{1}{(2\pi)^{\frac{k}{2}}} \cdot \frac{1}{\sqrt{\det(\Gamma)}} \cdot e^{-\frac{1}{2}\underset{(n\times 1)}{(x-\mu)}\cdot\,\underset{(n\times n)}{\Gamma^{-1}}\cdot\,\underset{(n\times 1)}{(x-\mu)}}$$

Note that if Γ is singular (i.e. $\mathsf{rank}(\Gamma) = r < n$) then the distribution of X is not absolutely continuous with respect to the Lebesgue Measure (and so X has no density). Indeed, the distribution lives on a hyperplane of $\mathbb{R}^n$ and at least on component of X is a linear combination of the others[1].

---

[1] What the fuck? This is absolutely heinous.