Ist Prob and Prob

Basteri Andrea

2019

Contents

1 Notation

A little section of notation that we used in all these notes.

- $\mathfrak{M}((\Omega, \mathscr{F}), (E, \mathscr{E})) := \{f : (\Omega, \mathscr{F}) \to (E, \mathscr{E})$ *measurable* $\}$. If the sigma algebras are clear, we indicate this set as $\mathfrak{M}(\Omega, E)$.

2 General

Let Ω be a set, let $\mathscr{F}$ be a σ-algebra of $\mathscr{P}(\Omega)$, and let $\mathbb{P}$ be a probability measure on $\mathscr{F}$.

Definition 1 (Probability Space)**.** $(\Omega,\mathscr{F},\mathbb{P})$ is a probability space.

Definition 2 (Measurable Space)**.** $(\Omega,\mathscr{F})$ is a measurable space.

Let E be a *Topological Space* (So we have the *Open Sets*).

Definition 3 (Borelian Sets)**.** $\mathfrak{B}(E)$ is the littlest σ-algebra that contains the open sets.

Let $Y : \Omega \to E$ be a function, with $(E,\mathscr{E})$ a measurable space.

Definition 4 (Aleatory Variable)**.** Y is an aleatory variable, or a random variable (r.v.), if Y is measurable, that is

$$\forall A \in \mathscr{E}, Y^{-1}(A) \coloneqq \{Y \in A\} \in \mathscr{F}$$

If we define

$$Y^{-1}(\mathscr{E}) \coloneqq \{Y^{-1}(A) | A \in \mathscr{E}\}$$

we can write definition (4) as $Y^{-1}(\mathscr{E}) \subset \mathscr{F}$. We sometimes indicate the $\sigma - algebrbas$ saying that Y is $(\mathscr{F}, \mathscr{E}) - measurable$, or if it is clear just as Y is $\mathscr{F} - measurable$.

Definition 5 (Law of a r.v)**.** Let $X : \Omega \to E$ a r.v. We define the *Law of X* as the probability P_X, defined as follow for all $A \in \mathscr{E}$.

$$\mathbb{P}_X(A) \coloneqq \mathbb{P}(X^{-1}(A)).$$

Definition 6 (Real Random Variable)**.** Y is a real random variable (r.r.v.) if $E = \mathbb{R}$, and $\mathscr{E} = \mathfrak{B}(\mathbb{R})$.

Definition 7 (Generated $\sigma - algebra$)**.** Let Ω be a set, and let $\mathcal{I} \subset \mathcal{P}(\Omega)$ be a class of set of Ω. We define as $\sigma(\mathcal{I})$ as the $\sigma - algebra$ generated by $\mathcal{I}$, that is

$$\sigma(\mathcal{I}) \coloneqq \bigcap_{\gamma \in \Gamma} \gamma, \;\; \Gamma \coloneqq \{\gamma \;|\; \gamma \;\, is \;\, a \;\, \sigma - algebra, \;\; \gamma \supset \mathcal{I}\}$$

that is $\sigma(\mathcal{I})$ is the smallest $\sigma - algebra$ that contains $\mathcal{I}$.

Definition 8 (Product $\sigma - algebra$)**.** Let $(\Omega, \mathscr{F})$ and $(E, \mathscr{E})$ be two measurable spaces. We define product $\sigma - algebra$ as the smallest $\sigma - algebra$ that contains the rectangles, that is

$$\mathscr{E} \otimes \mathscr{F} \coloneqq \sigma\left(\{A \times B| \;\; A \in \mathscr{E}, \;\; B \in \mathscr{F}\}\right).$$

Let $(\Omega, \mathscr{F})$ and $(E, \mathscr{E})$ be two measurable spaces, and let $X : (\Omega, \mathscr{F}) \to (E, \mathscr{E})$ be a r.v.

Proposition 2.1. *$X^{-1}(\mathscr{E})$ is a $\sigma - algebra$.*

Proof. Easy check. □

Definition 9 ($\sigma - algebra$ generated by a r.v.)**.** We define the $X^{-1}(\mathscr{E})$ above as the $\sigma - algebra$ generated by the r.v, and we denote it as $\sigma(X)$.

Remark 1. We observe that $\sigma(X)$ is the *smallest* $\sigma - algebra$ that make X measurable.

Definition 10 (Union $\sigma - algebra$)**.** Let Ω be a set and let $(\mathscr{F}_i)_{i\in I}$ a family of $\sigma - algebras$ of Ω indexed by a set I. We define the *smallest* $\sigma - algebras$ that contains every $\mathscr{F}_i$ as

$$\bigvee_{i\in I} \mathscr{F}_i \coloneqq \sigma\left(\bigcup_{i\in I} \mathscr{F}_i\right).$$

Now, we want to write down a trivial fact, but it may be useful in some observation.

- Let $(\Omega, \mathscr{F})$ and $(E, \mathscr{E})$ and $(T, \mathcal{T})$ be measurable spaces.
- Let us have X and Y measurable function defined in this way,

$$(\Omega, \mathscr{F}) \xrightarrow{X} (E, \mathscr{E}) \xrightarrow{Y} (T, \mathcal{T})$$

- Let us consider $Y \circ X$.

Proposition 2.2 (Trivial fact on $\sigma - algebra$ of r.v.)**.** *We have that $\sigma(Y \circ X) \subseteq \sigma(X)$.*

Proof. Let us have $A \in \sigma(Y \circ X)$. Then we have that we can find $B \in \mathcal{T}$ such that

$$A = (Y \circ X)^{-1}(B) = X^{-1}(Y^{-1}(B)) \in \sigma(X)$$

and we have finished. □

Corollary 2.3 (Second trivial fact on $\sigma - algebras$ of *r.v.*)**.** *Let us suppose that Y is invertible, and its inverse (that we call Z) is measurable. Then $\sigma(Y \circ X) = \sigma(X)$.*

Proof. We just need to use Proposition (2.2). We have

$$\sigma(Y \circ X) \subseteq \sigma(X) = \sigma(Z \circ (Y \circ X)) \subseteq \sigma(Y \circ X),$$

and this is the thesis. □

2.1 Independence

Let $(\Omega, \mathscr{F}, \mathbb{P})$ a fixed probability space.

Definition 11 (Independence of Events)**.** Let $A \in \mathscr{F}$ and $B \in \mathscr{F}$ be two events. We say that they are *independent* if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Let $\mathscr{F}_1, ..., \mathscr{F}_n$ be $\sigma - algebras$, and we suppose that $\mathscr{F}_i \subset \mathscr{F}$ for all i.

Definition 12 (Independence of $\sigma - algebras$)**.** The $\sigma - algebras$ $\mathscr{F}_1, .., \mathscr{F}_n$ are independent if for all $A_1 \in \mathscr{F}_1, ..., A_n \in \mathscr{F}_n$, we have that

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \mathbb{P}(A_i).$$

Remark 2. If we have $A \in \mathscr{F}$, then $\sigma(A) = \{A, A^c, \Omega, \emptyset\}$

Definition 13. The events $A_1 \in \mathscr{F}, .., A_n \in \mathscr{F}$ are independents if the $\sigma - algebras$ $\sigma(A_1), .., \sigma(A_n)$ are independents.

Let $X_i : (\Omega, \mathscr{F}) \to (E, \mathscr{E})$ be r.v for $i = 1, .., n$, and let $\mathbb{P}$ be a probability on $(\Omega, \mathscr{F})$.

Definition 14 (Independence of r.v.)**.** We say that the random variables $X_1, .., X_n$ are independent if $\sigma(X_1), .., \sigma(X_n)$ are independent.

2.2 Monotone Class Theorem and Its Consequences

Theorem 2.4 (Monotone Class Theorem)**.** *Let Ω be a set, let $\mathcal{I} \subset \mathcal{P}(\Omega)$ such that*

- $\Omega \in \mathcal{I}$,
- $A, B \in \mathcal{I} \implies A \cap B \in \mathcal{I}$.

Let $\mathcal{M} \supset \mathcal{I}$ such that

a) $\forall n \in \mathbb{N}$ *we have* $A_n \in \mathcal{M}$, $A_n \subseteq A_{n+1} \implies \bigcup_{n=0}^{+\infty} A_n \in \mathcal{M}$.

b) $A, B \in \mathcal{M}$ *and* $B \subseteq A \implies (A \setminus B) \in \mathcal{M}$.

c) $\mathcal{M}$ *is minimal, that is, if* $\mathcal{G}$ *is another class such that* $\mathcal{G} \supseteq \mathcal{I}$, *and* $\mathcal{G}$ *has the properties a) and b), then* $\mathcal{M} \subseteq \mathcal{G}$.

Then $\mathcal{M}$ is a $\sigma - algebra$, and $\mathcal{M} = \sigma(\mathcal{I})$.

Remark 3. This is a really simple but at the same time really important remark, because is the key to prove many corollaries of (2.4).

$$\mathcal{I} \subseteq \mathcal{A} \subseteq \mathcal{P}(\Omega),\ \mathcal{A}\ respects\ a)\ and\ b)\ of\ (2.4) \implies \sigma(\mathcal{I}) \subseteq \mathcal{A}.$$

The class $\mathcal{I}$ of (2.4) is really special, so we give it a special name.

Definition 15 ($\pi - system$)**.** Let $\mathcal{I} \subseteq \mathcal{P}(\Omega)$ be a class of sets. We say that $\mathcal{I}$ is a $\pi - system$ if

- $\Omega \in I$,
- $A, B \in \mathcal{I} \implies (A \cap B) \in \mathcal{I}$.

If a $\sigma - algebra$ $\mathscr{F}$ is given and $\sigma(\mathcal{I}) = \mathscr{F}$, then we say that $\mathcal{I}$ is a $\pi - system$ for $\mathscr{F}$.

Remark 4. From the definition above, it follows directly that a finite intersection of elements of $\mathcal{I}$ belongs to $\mathcal{I}$ (this property is called stability by intersection).

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probabilistic space, and let $\mathbb{Q}$ be another probability on the same space.

Corollary 2.5. *If $\mathbb{P}$ and $\mathbb{Q}$ coincides on a $\pi - system$ for $\mathscr{F}$, then $\mathbb{P}$ coincides to $\mathbb{Q}$ on $\mathscr{F}$.*

Proof. This is a standard strategy, so we write it just one time.

- Let $\mathcal{I}$ be a $\pi - system$ for $\mathscr{F}$. We observe that we just need that $\mathcal{I}$ is stable for intersection, because otherwise, since $\mathbb{P}(\Omega) = \mathbb{Q}(\Omega)$, we can take $\mathcal{I} \cup \{\Omega\}$.
- Let us set $\mathcal{A} := \{A \in \mathscr{F} | \ \mathbb{P}(A) = \mathbb{Q}(A)\}$. We observe that $\mathcal{I} \subseteq \mathcal{A} \subseteq \mathscr{F}$.
- Since $\mathcal{I}$ is a $\pi - system$, if $\mathcal{A}$ respects condition a) and b) of (2.4), then we have that $\mathscr{F} = \sigma(\mathcal{I}) \subseteq \mathcal{A} \subseteq \mathscr{F}$, so we have the equality. Let's check.
- a). Let us have $A_i \in \mathcal{A}$ for all $i \in \mathbb{N}$, and let us suppose that $A_i \subseteq A_{i+i}$ for all i. Let us set $A := \bigcup_{i=0}^{+\infty} A_i$. We have

  $$\mathbb{P}(A) = \lim_{n \to +\infty} \mathbb{P}(A_n) = \lim_{n \to +\infty} \mathbb{Q}(A_n) = \mathbb{Q}(A) \implies A \in \mathcal{A}.$$

  We have used the continuity of probability and that $A_n \in \mathcal{A}$ for all n, so condition a) is true.
- b). Let us have $A \in \mathcal{A}$ and $B \in \mathcal{A}$, with $B \subseteq A$. We can write

  $$\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(B) = \mathbb{Q}(A) - \mathbb{Q}(A) = \mathbb{Q}(A \setminus B) \implies (A \setminus B) \in \mathcal{A},$$

  so even condition b) holds true.
- Since we have check that a) and b) hold true, we have that $\mathcal{A} = \mathscr{F}$, and this is the thesis.

□

Let $(\Omega, \mathscr{F})$ and $(E, \mathscr{E})$ be two measurable spaces.

Definition 16. $C \in \mathscr{E} \otimes \mathscr{F}$. We define C_x as

$$C_x \coloneqq \{y \in F | (x, y) \in C\} = \pi_F(C \cap (\{x\} \times F)).$$

Corollary 2.6. *For every $C \in \mathscr{E} \otimes \mathscr{F}$, for every $x \in E$, we have that $C_x \in \mathscr{F}$.*

Let $(\Omega, \mathscr{F})$ and $(E, \mathscr{E})$ be two measurable space and let $X : (\Omega, \mathscr{F}) \to (E, \mathscr{E})$ be a function.

Proposition 2.7. *Let $\mathcal{D} \subseteq \mathscr{E}$ a family of subset of E, and let us suppose that*

- $\sigma(\mathcal{D}) = \mathscr{E}$,
- $\forall A \in \mathcal{D}$, *we have that* $X^{-1}(A) \in \mathscr{F}$.

Then X is $\mathscr{F} - measurable$, that is $\sigma(X) \subseteq \mathscr{F}$.

Proof. We can not apply directly the standard strategy because $\mathcal{D}$ is not a $\pi - system$ in general. This is not a big problem, because given $\mathcal{D}$ we can define a $\pi - system$ for $\mathscr{E}$. Let us set

$$\mathcal{L} \coloneqq \{A \in \mathscr{E} | \; \exists A_1, ..., A_n \in \mathcal{D} \; s.t. \; A = \bigcap_{i=1}^{n} A_i\} \cup \{\Omega\}.$$

□

It is easy to show that this is a $\pi - system$, that $\mathcal{D} \subseteq \mathcal{L}$ and $X^{-1}(\mathcal{L}) \subseteq \mathscr{F}$. Now we can follow the standard strategy, that is an easy check.

Let $(\Omega, \mathscr{F})$ and $(E, \mathscr{E})$ be two measurable space and let $X : (\Omega, \mathscr{F}) \to (E, \mathscr{E})$ be a function. We remember that X is $\sigma(X) - measurable$ by definition.

Proposition 2.8. *Let $\mathcal{I} \subseteq \mathscr{E}$ be a $\pi - system$ for $\mathscr{E}$. Then $X^{-1}(\mathcal{I})$ is a $\pi - system$ for $\sigma(X)$.*

Proof. The proof is an easy check that we sketch.

- $\sigma(X^{-1}(\mathcal{I})) = \sigma(X)$, that is $X^{-1}(\mathcal{I})$ generate $\sigma(X)$. Indeed, we just need to consider

  $$\mathcal{A} \coloneqq \{A \in \mathscr{E} | \; X^{-1}(A) \in \sigma(X^{-1}(\mathcal{I}))\}.$$

  We have that $\mathcal{I} \subseteq \mathcal{A}$, and we can verify that $\mathcal{A}$ verify the usual condition, so we have the searched equality.

- $X^{-1}(\mathcal{I})$ is a $\pi - system$. We have
  - $\Omega \in X^{-1}(\mathcal{I})$. We just observe that $\Omega = X^{-1}(E)$.
  - $A, B \in X^{-1}(\mathcal{I})$. Then

    $$A \cap B = X^{-1}(C) \cap X^{-1}(D) = X^{-1}(C \cap D) \in X^{-1}(\mathcal{I}),$$

    where C and D are elements of $\mathcal{I}$, so their intersection belong to $\mathcal{I}$. We have used the powerful property of the counter-images.

□

Now we want to enunciate a criterion to establish if a finite number of σ-algebras are independent. Let's begin with two $sigma - algebras$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ a fixed probability space. Let $\mathscr{F}_1 \subseteq \mathscr{F}$ and $\mathscr{F}_2 \subseteq \mathscr{F}$ be two $sigma - algebras$.

Proposition 2.9. *Let us have $\mathcal{I}_1 \subseteq \mathscr{F}_1$ that is a $\pi - system$ for $\mathscr{F}_1$ and $\mathcal{I}_2 \subseteq \mathscr{F}_2$ that is a $\pi - system$ for $\mathscr{F}_2$. Let us suppose that*

$$\forall A_1 \in \mathcal{I}_1,\ \forall A_2 \in \mathcal{I}_2,\ \mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2). \tag{1}$$

Then $\mathscr{F}_1$ and $\mathscr{F}_2$ are independents, that is equality (1) holds true for all $A_1 \in \mathscr{F}_1$ and for all $A_2 \in \mathscr{F}_2$.

Remark 5. As before, we just need that $\mathcal{I}_1$ and $\mathcal{I}_2$ were stable by intersection, because we can increase $\mathcal{I}_1$ and $\mathcal{I}_2$ by adding $\{\Omega\}$.

Proof. The proof is simple, let us see.

- Let us set $\mathcal{A} := \{A \in \mathscr{F}_1 |\ \forall B \in \mathcal{I}_2,\ \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)\}$. We want to check if $\mathcal{A}$ has condition a) and b) of (2.4), so by remembering *Remark* (3) we conclude.

  a) Obvious.

  b) $A, C \in \mathcal{A}$ and $C \subseteq A$. Then for all $B \in \mathcal{I}_2$

  $$\mathbb{P}((A \setminus C) \cap B) = \mathbb{P}((A \cap B) \setminus (C \cap B)) = \mathbb{P}(A \cap B) - \mathbb{P}(C \cap B) =$$
  $$\mathbb{P}(A)\mathbb{P}(B) - \mathbb{P}(C)\mathbb{P}(B) = (\mathbb{P}(A) - \mathbb{P}(C))\mathbb{P}(B) = \mathbb{P}(A \setminus C)\mathbb{P}(B) \implies (A \setminus C) \in \mathcal{A}.$$

  So, we have that $\mathcal{I}_1 \subseteq \mathcal{A} \subseteq \mathscr{F}_1 \implies \mathscr{F}_1 = \sigma(\mathcal{I}_1) \subseteq \mathcal{A} \subseteq \mathscr{F}_1$, and this implies the equality.

- Let us set $\mathcal{A}_2 := \{A \in \mathscr{F}_2 |\ \forall B \in \mathscr{F}_1,\ \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)\}$. The proof that $\mathcal{A}_2 = \mathscr{F}_2$ is equal to the one above.

□

Lemma 2.10. *Let us have*

- *$\mathscr{F}_1, .., \mathscr{F}_n$ $\sigma - algebras$, contained in a bigger $\sigma - algebra$ $\mathscr{F}$.*
- *for all i, $\mathcal{I}_i$ is a $\pi - system$ for $\mathscr{F}_i$,*
- *let us define $\mathcal{G} := \bigvee_{i=1}^{n} \mathscr{F}_i$.*

Then

$$\mathcal{L} := \{A \in \mathcal{G} |\ there\ exist\ A_1 \in \mathcal{I}_1, ..., A_n \in \mathcal{I}_n\ s.t.\ A = \bigcap_{i=1}^{n} A_i\}$$

is a $\pi - system$ for $\mathcal{G}$.

Proof. We just check that $\mathcal{L}$ has the property of a $\pi - system$ for $\mathcal{G}$.

- $\mathcal{L}$ generate, that is $\sigma(\mathcal{L}) = \mathcal{G}$.
  - Clearly, $\sigma(\mathcal{L}) \subseteq \mathcal{G}$.
  - for all i, $\mathcal{I}_i \subseteq \mathcal{L} \implies \mathscr{F}_i \subseteq \sigma(\mathcal{L}) \implies \mathcal{G} = \sigma(\bigcup_{i=1}^n \mathscr{F}_i) \subseteq \sigma(\mathcal{L})$, and this conclude.
- $\mathcal{L}$ is a $\pi - system$.
  - Clearly, $\Omega \in \mathcal{L}$ (we simply take $A_i = \Omega$).
  - $A, B \in \mathcal{L}$. Then $A = \bigcap_{i=1}^n A_i$ and $B = \bigcap_{i=1}^n B_i$, with $A_i \in \mathcal{I}_i$ and $B_i \in \mathcal{I}_i$ for all i. Then

$$A \cap B = \bigcap_{i=1}^{n} \left(\underbrace{A_i \cap B_i}_{\in \mathcal{I}_i} \right) \in \mathcal{L}$$

    and $A_i \cap B_i \in \mathcal{I}_i$ because $\mathcal{I}_i$ is closed by intersection.

With this last check we just have concluded. □

Corollary 2.11 (general criterion of independence)**.** *Let us have*

- $\mathscr{F}_1, .., \mathscr{F}_n$ *$\sigma - algebras$, contained in a bigger $\sigma - algebra$ $\mathscr{F}$.*
- *for all i, we have $\mathcal{I}_i$ a $\pi - system$ for $\mathscr{F}_i$,*
- *let us suppose that*

$$\forall A_i \in \mathcal{I}_i, \; \mathbb{P}\left(\bigcap_{1 \le i \le n} A_i \right) = \prod_{1 \le i \le n} \mathbb{P}(A_i). \tag{2}$$

Then

1. *$\mathscr{F}_n$ is independent of $\bigvee_{i=1}^{n-1} \mathscr{F}_i$.*
2. *$\mathscr{F}_1, .., \mathscr{F}_n$ are independent.*

Proof. We prove this corollary by induction.

- $n = 2$. This is just (2.9).
- $n > 2$.
  - Let us define $\mathcal{G} := \bigvee_{i=1}^{n-1} \mathscr{F}_i$. By (2.10), we have that

$$\mathcal{L} := \{A \in \mathcal{G} | \; there \; exist \; A_1 \in \mathcal{I}_1, ..., A_{n-1} \in \mathcal{I}_{n-1} \; s.t. \; A = \bigcap_{i=1}^{n-1} A_i\}$$

    is a $\pi - system$ for $\mathcal{G}$. Now we want to prove $\mathcal{G}$ and $\mathscr{F}_n$ are independent.

– We use again (2.9). We have $\mathcal{I}_n$ a $\pi - system$ for $\mathscr{F}_n$ and $\mathcal{L}$ a $\pi - system$ for $\mathcal{G}$. We have $\forall A_n \in \mathcal{I}_n \ and \ B \in \mathcal{L}$

$$\mathbb{P}(A_n \cap B) = \mathbb{P}\left(A_n \cap \bigcap_{i=1}^{n-1} \underbrace{A_i}_{\in \mathcal{I}_i}\right) = \prod_{1 \le i \le n} \mathbb{P}(A_i) = \mathbb{P}(A_n) \prod_{1 \le i \le n-1} \mathbb{P}(A_i) \underbrace{=}_{(*)}$$

$$= \mathbb{P}(A_n)\mathbb{P}\left(\bigcap_{i=1}^{n-1} A_i\right) = \mathbb{P}(A_n)\mathbb{P}(B),$$

where in $(*)$ we have used (2) with $A_n = \Omega$. So for our criterion (2.9), $\mathscr{F}_n$ and $\mathcal{G}$ are independents, that is

$$\forall A \in \mathscr{F}_n, \ \forall B \in \mathcal{G}, \ \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

– Now, let us set $B := \bigcap_{i=1}^{n-1} A_i$, with $A_i \in \mathscr{F}_i$. So $B \in \mathcal{G}$, and by inductive hypothesis, we have that

$$\mathbb{P}\left(\bigcap_{i=1}^{n-1} A_i\right) = \prod_{i=1}^{n-1} \mathbb{P}(A_i).$$

We conclude observing that for all $A_n \in \mathscr{F}_n$

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbb{P}(A_n \cap B) = \mathbb{P}(A_n)\mathbb{P}(B) = \prod_{i=1}^{n} \mathbb{P}(A_i)$$

and this is the thesis.

□

Now we have the following criterion,

Corollary 2.12. *Let us have* $n \ge 2$ *integer, and let us have*

- $\mathscr{F}_1, .., \mathscr{F}_n$ *that are* $\sigma - algebras$*, contained in a bigger* $\sigma - algebra$ $\mathscr{F}$*.*
- $\mathscr{F}_1, .., \mathscr{F}_{n-1}$ *are independent* $\sigma - algebras$*.*

Then the following are equivalent,

1. $\mathscr{F}_n$ *is independent of* $\bigvee_{i=1}^{n-1} \mathscr{F}_i$*,*
2. $\mathscr{F}_1, .., \mathscr{F}_n$ *are independent.*

Proof. Let us define

$$\mathcal{G} := \bigvee_{i=1}^{n-1} \mathscr{F}_i.$$

Let's see.

- 1) $\implies$ 2).

  Let $A_1 \in \mathscr{F}_1, ..., A_n \in \mathscr{F}_n$ be sets. Then we have

$$\mathbb{P}\left(A_n \cap \underbrace{(\bigcap_{i=1}^{n-1} A_i)}_{\in \mathscr{G}}\right) = \mathbb{P}(A_n)\mathbb{P}\left(\bigcap_{i=1}^{n-1} A_i\right) \underbrace{=}_{(*)} \prod_{i=1}^{n} \mathbb{P}(A_i),$$

  Where in $(*)$ we have used that $\mathscr{F}_1, ..., \mathscr{F}_{n-1}$ are independents, and the equality above is the definition of independence.

- 2) $\implies$ 1).

  We have that

  - for all $i = 1, .., n$, the $\sigma - algebra$ $\mathscr{F}_i$ is a $\pi - system$ for $\mathscr{F}_i$.
  - for all $A_i \in \mathscr{F}_i$, equality (2) hold since $\mathscr{F}_1, .., \mathscr{F}_n$ are independents.

  So by Corollary (2.11), we have that $\mathscr{F}_n$ is independent of $\mathscr{G}$, and this is the thesis.

□

2.2.1 Sigma Algebras Of R.V. and Independence

Now we write some corollaries of the theorems of the above section.

Let us have $(E, \mathscr{E})$ a measurable space, and let $n \geq 1$ be an integer.

Corollary 2.13. *Let us consider*

- $(E^n, \bigotimes_n \mathscr{E})$ *the product space with the product sigma algebra.*
- *Let $\mathcal{I}$ be a $\pi - system$ for $\mathscr{E}$.*

Then

$$\mathcal{G} := \{\times_{i=1}^{n} B_i \ : \ B_i \in \mathcal{I}\} = \mathcal{I}^n.$$

is a $\pi - system$ for $\bigotimes_n \mathscr{E}$

Proof. The proof is simple and follow from Corollary (2.10).

- We have that by definition

$$\bigotimes_n \mathscr{E} = \bigvee_{i=1}^{n} \sigma(p_i),$$

  with $p_i : E^n \to E$ such that $p(e_1, .., e_n) = e_i$ the canonical projection. In fact, the product $\sigma - algebra$ is the littlest $\sigma - algebra$ such that the canonical projection are measurable.

- Now, we have that
  - $\sigma(p_1), .., \sigma(p_n) \subseteq \bigotimes_n \mathscr{E}$ are sigma algebras,
  - for all i, we have $\mathcal{I}$ is a $\pi - system$ for $\mathscr{E} \implies p^{-1}(\mathcal{I})$ is a $\pi - system$ for $\sigma(p_i)$, thanks to $Corollary$ (2.8).

  So we have that

$$\mathcal{G} := \left\{ \cap_{i=1}^n A_i \ \ : \ \ A_i \in p_i^{-1}(\mathcal{I}) \right\}$$

  is a $pi - system$ for $\bigotimes_n \mathscr{E}$, and this is the thesis because every element of $\mathcal{G}$ is a product between elements of $\mathcal{I}$.

□

2.3 Product Probability

Let us have $(F, \mathscr{F}, \mathbb{Q})$ and $(E, \mathscr{E}, \mathbb{P})$, that are two probabilistic space. We would like a probability $\mathbb{R}$, which we denote with $\mathbb{P} \otimes \mathbb{Q}$, on the space $(E \times F, \mathscr{E} \otimes \mathscr{F})$ such that for all $A \in \mathscr{F}$ and $B \in \mathscr{E}$, we have

$$(\mathbb{P} \otimes \mathbb{Q})(A \times B) = \mathbb{P}(A)\mathbb{Q}(B).$$

Clearly, if such probability exists, it is unique.

Proposition 2.14. *Let us have $f : E \times F \to \mathbb{R}$ a measurable function such that $f \geq 0$. Then*

1. *$\forall x \in E$ the function $f_x : F \to \mathbb{R}$ such that $f_x(y) := f(x, y)$ is $\mathscr{F} - measurable$.*
2. *The function $g : E \to \mathbb{R}$ such that $g(x) := \int_F f_x(y) d\mathbb{Q}(y)$ is $\mathscr{E} - measurable$.*

Proof. Let us prove first 1. and after 2.

1. - $f(x, y) = I_C(x, y)$ with $C \in \mathscr{E} \otimes \mathscr{F}$.
     Because of (2.6), we have that the section $C_x \in \mathscr{F}$, and it is straightforward to show that for all $x \in E$ fixed, $I_C(x, y) = I_{C_x}(y) = f_x(y)$, so f_x is $\mathscr{F} - measurable$.
   - If f is a linear combination of indicator function (that is a simple function), we conclude by linearity.
   - f measurable, $f \geq 0$. We can find a sequence of simple function $(f_n)_{n \in \mathbb{N}}$ such that $f_n \uparrow f$, so we can conclude because $f = sup_n(f_n)$, that is $\mathscr{F} - measurable$ if x is fixed.
2. - $f(x, y) = I_{A \times B}(x, y)$, with $A \in \mathscr{F}$ and $B \in \mathscr{E}$. We have that $f_x(y) = I_A(x) I_B(y)$, so

$$g(x) = \int_F f_x(y) d\mathbb{Q}(y) = I_A(x)\mathbb{Q}(B),$$

     that is $\mathscr{E} - measurable$.

- We define $\mathcal{A} \coloneqq \{C \in \mathscr{E} \otimes \mathscr{F} | \int_F I_C(x, y)\ d\mathbb{Q}(y)\ is\ \mathscr{E} - measurable\}$, and we prove thanks to (2.4) that $\mathcal{A} = \mathscr{E} \otimes \mathscr{F}$.
- Thanks to linearity, we extend the above result to f simple.
- If f is measurable, positive, we find a sequence of simple, increasing function that approximate f and we conclude by Beppo Levi.

□

2.3.1 Independence thanks to Product Probability

- Let us have $(\Omega, \mathscr{F}, \mathbb{P})$ a probability space.
- Let $X : (\Omega, \mathscr{F}) \to (E_1, \mathscr{E}_1)$ and $Y : (\Omega, \mathscr{F}) \to (E_2, \mathscr{E}_2)$ be two r.v.
- We can consider the function $(X, Y) : (\Omega, \mathscr{F}) \to (E_1 \times E_2, \mathscr{E}_1 \otimes \mathscr{E}_2)$, such that

$$(X, Y)(\omega) = (X(\omega), Y(\omega)).$$

- All these function have a law, that are respectively $\mathbb{P}_X$ and $\mathbb{P}_Y$ and $\mathbb{P}_{(X,Y)}$.
- We observe that we can consider $(E_1, \mathscr{E}_1, \mathbb{P}_X)$ and $(E_2, \mathscr{E}_2, \mathbb{P}_Y)$ as probabilistic space, and we can build the probability space $(E_1 \times E_2, \mathscr{E}_1 \otimes \mathscr{E}_2, \mathbb{P}_X \otimes \mathbb{P}_Y)$.

Lemma 2.15. *X and Y are independent $\iff$ $\mathbb{P}_{(X,Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y$.*

Proof. We see first one implication, then the another.

- X and Y are independent.

  We observe that $\mathscr{E} \times \mathscr{F}$ is a $\pi - system$ for $\mathscr{E} \otimes \mathscr{F}$. Because of our hypothesis, $\mathbb{P}_{(X,Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y$ on $\mathscr{E} \times \mathscr{F}$, so they are equal on $\mathscr{E} \otimes \mathscr{F}$ thanks to (2.5).
- $\mathbb{P}_{(X,Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y$.

  We just need to observe that $\{(X, Y) \in A \times B\} = \{X \in A, Y \in B\}$, and the thesis is straightforward if we evaluated the identity in $A \times B \in \mathscr{E} \times \mathscr{F}$.

□

Remark 6. Of course, the argument is the same, even if we have a finite number (say n) of random variable.

2.4 Other Independence Criterion

2.4.1 Independence if we have a condition for every C 0 Bounded function

Let us have $(\Omega, \mathscr{F}, \mathbb{P})$ a probabilistic space, and let us have $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ Let $X : \Omega \to \mathbb{R}$ be a $r.r.v$, and let us have $\mathscr{G} \subseteq \mathscr{F}$ a $\sigma - algebra$.

Proposition 2.16. *Let us suppose that* $\forall A \in \mathcal{I}$*, and* $\forall \varphi \in C_B^0(\mathbb{R})$ *(continuous and bounded real functions) we have that*

$$\mathbb{E}[(\varphi \circ X) \cdot I_A] = \mathbb{E}[\varphi \circ X] \underbrace{\mathbb{E}[I_A]}_{\mathbb{P}(A)}, \tag{3}$$

with $\mathcal{I}$ *a* $\pi - system$ *for* $\mathcal{G}$*. Then* X *and* $\mathscr{G}$ *are independent, that is* $\sigma(X)$ *and* $\mathscr{G}$ *are independent.*

Proof. We just need to prove that given $A \in \mathcal{I}$ for all $B \in \mathfrak{B}(\mathbb{R})$ we have that

$$\mathbb{P}\left(\{X \in B\} \cap A\right) = \mathbb{P}(X \in B)\mathbb{P}(A).$$

because $\mathcal{I}$ is a $\pi - system$ for $\mathcal{G}$, and $\sigma(X)$ is a $\pi - system$ for itself, thanks to (2.9).

- Our strategy is to pass from the continuous and bounded functions to the indicator function of a $\pi - system$ of $\mathfrak{B}(\mathbb{R})$, then thanks to (2.4) we pass to the *Borel* of $\mathbb{R}$, so we have the thesis if we take the indicator function of such sets.
- Now, let us fix $A \in \mathcal{I}$.
- We define

  $$\mathcal{A} = \{\ (-\infty, x] :\ \ x \in \mathbb{R}\} \cup \{\mathbb{R}\}.$$

  We remember that $\mathfrak{B}(\mathbb{R}) = \sigma(\{open\ \ set\ \ of\ \ \mathbb{R}\}) = \sigma(\mathcal{A})$.
- It is immediate that $\mathcal{A}$ is stable by finite intersection, so it is a $\pi - system$ for $\mathfrak{B}(\mathbb{R})$.
- Let us test formula (3) on the sets of $\mathcal{A}$, that is we want to see if (3) holds true with $\varphi = I_B$, with $B \in \mathcal{A}$.
- If $\varphi = I_{\mathbb{R}}$, then (3) is trivially true because $\varphi \in C_B^0(\mathbb{R})$.
- We suppose now $\varphi = I_{(-\infty,b]}$, with $b \in \mathbb{R}$.
- Let us define

  $$\varphi_n(x) := I_{(-\infty,b]}(x) + I_{(b,b+\frac{1}{n}]} \cdot \left[n(b-x)+1\right].$$

  that is

  $$\varphi_n(x) = \begin{cases} 1\ \ if\ \ x \leq b \\ -nx + nb + 1\ \ if\ \ b < x \leq b + \frac{1}{n} \\ 0\ \ if\ \ x > b + \frac{1}{n}. \end{cases}$$

  It is immediate that φ_n is a bounded continuous function and, and $\varphi_n \overbrace{\to}^{n \to +\infty} \varphi$ point-wise.

- So, we have that

$$\forall \omega \in \Omega, \ \ \varphi_n(X(\omega)) \overbrace{\to}^{n\to+\infty} \varphi(X(\omega)),$$

and since $|\varphi_n \circ X| \leq I_\Omega \in L^1(\Omega)$, we have that

$$\mathbb{E}[\varphi_n \circ X] \overbrace{\to}^{n\to+\infty} \mathbb{E}[\varphi \circ X],$$

by dominated convergence.

- Now, we observe moreover that

$$\forall \omega \in \Omega, \ \ \varphi_n(X(\omega)) \cdot I_A(\omega) \overbrace{\to}^{n\to+\infty} \varphi(X(\omega)) \cdot I_A(\omega),$$

and since $|(\varphi_n \circ X) \cdot I_A| \leq I_\Omega \in L^1(\Omega)$, we have that

$$\mathbb{E}[(\varphi_n \circ X) \cdot I_A] \overbrace{\to}^{n\to+\infty} \mathbb{E}[(\varphi \circ X) \cdot I_A],$$

by dominated convergence.

- So, if we put everything together and we observe that for all n, we have $\varphi_n \in C^0_B(\mathbb{R})$, we obtain

$$\mathbb{E}[(\varphi \circ X) \cdot I_A] = \mathbb{E}[\lim_n (\varphi_n \circ X) \cdot I_A] = \lim_n \mathbb{E}[(\varphi_n \circ X) \cdot I_A] =$$
$$\lim_n \mathbb{E}[(\varphi_n \circ X)] \underbrace{\mathbb{E}[I_A]}_{\mathbb{P}(A)} = \mathbb{E}[\varphi \circ X]\mathbb{P}(A).$$

- So we have discover that (3) hold true for the elements of $\mathcal{A}$.

- Now, let us define

$$\mathcal{B} := \left\{ B \in \mathfrak{B}(\mathbb{R}) : \mathbb{E}[(I_B \circ X) \cdot I_A] = \mathbb{E}[(I_B \circ X)] \cdot \mathbb{E}[I_A] \ \right\}.$$

- We have that $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathfrak{B}(\mathbb{R})$, and $\mathcal{A}$ is a $\pi - system$ for $\mathfrak{B}(\mathbb{R})$. If we prove that $\mathcal{B}$ respect condition (a) and (b) of (2.4), we have finished.

- This is just a standard check, and we omit it, so we have the thesis.

□

2.5 Arbitrary Product

Let I be a set of index, and let us consider $(E_i, \mathscr{E}_i)_{i\in I}$ a family of measurable space. Let us define

$$E = \prod_{i\in I} E_i := \{f : I \to \bigcup_{i\in I} E_i \ \mid \ \forall i \in I, \ \ f(i) \in E_i\}.$$

We consider given $j \in I$ the function $\pi_j : E \to E_j$ that is the canonical projection on E_j. So, for all $f \in E$, we have $\pi_j(f) = \pi_j((e_i)_{i\in I}) = e_j = f(j)$.

We can even consider the following point of view. Given $\emptyset \subseteq J \subseteq I$, not empty, we can define

$$\pi_J : \prod_{i\in I} E_i \to \prod_{j\in J} E_j, \text{ such that } \pi_J(f) = f|_J.$$

So, if $J = \{i\} \subseteq I$, we can identify the element $e \in E$ as the function $\hat{e} : \{i\} \to E_i$ such that $\hat{e}(i) = e$.

Definition 17 (arbitrary product $\sigma - algebra$)**.** We define the product $\sigma - algebra$ on E, and we indicate it as $\otimes_{i\in I}\mathscr{E}$, as the smallest $\sigma - algebra$ on which the projection π_i are measurable, that is

$$\mathscr{E}^{\otimes_{i\in I}} = \otimes_{i\in I}\mathscr{E} := \sigma\left(\{\pi_i^{-1}(A_i) \mid A_i \in \mathscr{E}_i\}\right) = \bigvee_{i\in I} \sigma(\pi_i).$$

Remark 7. The above definition goes well even if we have a *finite* product of probabilities.

Remark 8. We follow the following notation. If we have $K \subseteq J \subseteq I$, we can consider

$$\pi_K^{(J)} : \prod_{j\in J} E_j \to \prod_{k\in K} E_k, \text{ such that } \pi_K^{(J)}(g) = g|_K$$

that is the projection from a suitable subset of $\prod_{i\in I} E_i$ to another. Of course, if we have $H \subseteq K \subseteq J \subseteq I$, we have the following identity

$$\underbrace{\prod_{j\in J} E_j \overbrace{\to}^{\pi_K^{(J)}} \prod_{k\in K} E_k \overbrace{\to}^{\pi_H^{(K)}} \prod_{h\in H} E_h}_{\pi_H^{(J)}}$$

that is $\pi_H^{(K)} \circ \pi_K^{(J)} = \pi_H^{(J)}$.

We remember the following easy equality that hold true in general and sometimes it is useful. Let $A \to^f B \to^g C$ be two function, and let $D \subseteq C$ be a set. Then the following equality hold,

$$(g \circ f)^{-1}(D) = f^{-1}(g^{-1}(D)).$$

What are the rectangles in this definition? Given our idea in finite dimension, one can discover by heuristic that the rectangles are the elements of the form

$$\prod_{i\in I} A_i = \bigcap_{i\in I} \pi_i^{-1}(A_i),$$

with $(A_i)_{i\in I}$ a sequence of elements such that for all $i \in I$, we have $A_i \in \mathscr{E}_i$. We observe that in general I is not countable, so in general it is **NOT** true that $\prod_{i\in I} A_i \in \bigotimes_{i\in I} \mathscr{E}_i$. In fact, intersection of too sets can not stay in the $\sigma - algebra$, even though the single elments stay in it.

- Now, let $J \subseteq I$ be a subset, not empty.
- $\forall j \in J$, let us consider $A_j \subseteq E_j$. It would be better $A_j \in \mathscr{E}_j$.
- Let us define

$$A := \{g : J \to \bigcup_{j\in J} E_j : \ \ \forall j \in J,\ \ g_j = g|_j \in A_j\} \subseteq \prod_{j\in J} E_j.$$

Proposition 2.17. *Then we have*

$$A = \bigcap_{j\in J} (\pi_j^{(J)})^{-1}(A_j).$$

Proof. It is just a chain of implication. In fact

$$\begin{aligned} &f \in A \iff \forall j \in J,\ \ f|_j = \pi_j^{(J)}(f) \in A_j \iff \\ &\forall j \in J,\ \ f \in (\pi_j^{(J)})^{-1}(A_j) \iff f \in \bigcap_{j\in J} (\pi_j^{(J)})^{-1}(A_j). \end{aligned}$$

□

Remark 9. It is straightforward that

$$(\pi_J^{(I)})^{-1}(A) = \bigcap_{j\in J} (\pi_J^{(I)})^{-1}\left((\pi_j^{(J)})^{-1}(A_j)\right) = \bigcap_{j\in J} (\pi_j^{(J)} \circ \pi_J^{(I)})^{-1}(A_j) = \bigcap_{j\in J} (\pi_j^{(I)})^{-1}(A_j) =$$

if it is not ambiguous, for all $J \subseteq I$ we denote $\pi_J^{(I)}$ as π_J, that is we don't write the starting space if it is I.

Now, let $X : (\Omega, \mathscr{F}) \to (E, \otimes_{i\in I}\mathscr{E})$ be a function.

Proposition 2.18 (Measurability of a r.v. in the product space)**.** *The following fact are equivalent,*

1. *X is measurable.*
2. *$\forall i \in I$, $X_i := \pi_i \circ X : (\Omega, \mathscr{F}) \to (E_i, \mathscr{E}_i)$ is measurable.*

Moreover, we have that $\sigma((X_i)_{i\in I}) = \sigma(X)$, that is the smallest $\sigma - algebra$ that make every X_i measurable and the smallest $\sigma - algebra$ that make X measurable are the same.

Proof. Let's see.

- 1) $\implies$ 2). It's clear, it is composition of measurable functions, and from this it follows immediately that $\sigma(X_i : i \in I) \subseteq \sigma(X)$.
- 2) $\implies$ 1). We would like to use (2.7), so we have to find $\mathcal{A} \subseteq \bigotimes_i \mathscr{E}_i$ such that
  - $\sigma(\mathcal{A}) = \bigotimes_i \mathscr{E}_i$.
  - $\forall A \in \mathcal{A}$, we have that $X^{-1}(A) \in \mathscr{F}$.
- Let us consider

$$\mathcal{A} \coloneqq \{\pi_i^{-1}(A_i)|\ i \in I\ \ and\ \ A_i \in \mathscr{E}_i\} = \bigcup_{i \in I} \pi_i^{-1}(\mathscr{E}_i). \tag{4}$$

  We have that $\bigotimes_i \mathscr{E}_i = \sigma(\mathcal{A})$.
- Now, let $A \in \mathcal{A}$ be a set. Then there exists $i \in I$ and $A_i \in \mathscr{E}_i$ such that we have $A = \pi_i^{-1}(A_i)$. So we have that

$$X^{-1}(A) = X^{-1}(\pi_i^{-1}(A_i)) = (\pi_i \circ X)^{-1}(A_i) = X_i^{-1}(A_i) \in \mathscr{F}.$$

  so we can apply Proposition (2.7), and we have concluded.

  In particular if we are more accurate we can observe that

$$X^{-1}(A) \in \sigma(X_i) \subseteq \sigma(X_i : i \in I),$$

  so the truth is that X is $\sigma(X_i : i \in I)$ measurable, that is $\sigma(X) \subseteq \sigma(X_i : i \in I)$, and this conclude.

□

Remark 10. We observe that, given the set $\mathcal{A}$ defined in (4), we have that

$$\mathcal{B} \coloneqq \bigcup_{n \in \mathbb{N}} \{\ \cap_{i=1}^n A_i|\ \forall i,\ \ A_i \in \mathcal{A}\ \}$$

is a $\pi - system$ for $\bigotimes_i \mathscr{E}_i$.

3 Integrals and Convergence Theorems

Let us have a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Proposition 3.1 (Integration with respect to a Probability Law). *Let us have the following setting,*

- *let* $X : (\Omega, \mathscr{F}, \mathbb{P}) \to (E, \mathscr{E})$ *be a r.v,*
- *let* μ_X *be the law of* X*, that is for all* $A \in \mathscr{E}$*,* $\mu_X(A) = \mathbb{P}(X \in A)$*,*

- *let* $f : (E, \mathscr{E}) \to (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ *be a measurable function.*

  Then f *is* $\mu_X -$ *integrable if, and only if* $f \circ X$ *is* $\mathbb{P} -$ *integrable, and in this case we have*

$$\int_E f(x) \ \ \mu_X(dx) = \int_\Omega f \circ X(\omega) \ \ P(d\omega).$$

Let us have $(A_n)_{n\geq 1}$ a sequence of events in $(\Omega, \mathscr{F}, \mathbb{P})$. Let us define

$$\limsup_{n\to+\infty} A_n = A := \bigcap_{n=1}^{+\infty} \left(\bigcup_{k\geq n} A_k \right),$$

that is

$$A := \{\omega : \omega \in A_n \ \ for \ \ infinitely \ \ many \ \ indices \ \ n\}.$$

Then the following is true,

Lemma 3.2 (Borel-Cantelli Lemma)**.** *We have the following two condition.*

1. *If* $\sum_{n=1}^{+\infty} \mathbb{P}(A_n) < +\infty$, *then* $\mathbb{P}(A) = 0$.
2. *If* $\sum_{n=1}^{+\infty} \mathbb{P}(A_n) = +\infty$ *and* A_n *are pairwise independent, then* $\mathbb{P}(A) = 1$.

3.1 Convergence Theorem

Lemma 3.3 (Approximating by Simple Function)**.** . *Let* $f : (\Omega, \mathscr{F}) \to (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ *be a positive measurable function. Then we can find a sequence of simple function* $(f_n)_{n\in\mathbb{N}}$ *such that* $f_n \uparrow f$.

3.2 Switch Limit, Integral Theorem

Lemma 3.4 (Fatou's Lemma)**.** *Let us have*

- $(X_n)_{n\in\mathbb{N}}$ *a sequence of r.r.v. on our probability space.*
- $\forall n \in \mathbb{N}$, $X_n \geq 0$

Then

$$\int_\Omega (\liminf X_n) \ \ d\mathbb{P} \leq \liminf \left(\int_\Omega X_n \ \ d\mathbb{P} \right).$$

3.3 Measure Defined By a Density

Let μ be a probability on $(\mathbb{R}^m, \mathfrak{B}(\mathbb{R}^m))$.

Definition 18 (Density)**.** μ admits a *density* if there exists a function $f : (\mathbb{R}^m, \mathfrak{B}(\mathbb{R}^m)) \to (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ measurable, $f \geq 0$, such that for every $A \in \mathfrak{B}(\mathbb{R})$, we have

$$\mu(A) = \int_A f(x)dx.$$

Let us have $X : (\Omega, \mathscr{F}) \to (\mathbb{R}^m, \mathfrak{B}(\mathbb{R}^m))$ a r.v, let μ be the law of X and we suppose that μ has density f.

Lemma 3.5. *Let $a \in \mathbb{R}^m$, and let A a $m \times m$ invertible matrix. Let us set $Y := AX + a$. Then Y has density, and the density of Y is given by*

$$g(y) = \frac{1}{|\det A|} f(A^{-1}(y - a)), \ \ \forall y \in \mathbb{R}^m.$$

that is for all $A \in \mathfrak{B}(\mathbb{R}^m)$, we have $\mu_Y(A) = \mathbb{P}(Y \in A) = \int_A g(y)dy$.

4 Conditional Mean

5 Theorem on r.v.

5.1 Sum of r.r.v.

- Let us have $(\Omega, \mathscr{F}, \mathbb{P})$ a probabilistic space.
- Let us have $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$.
- Let $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$ be random variable.

Proposition 5.1. *We have that*

$$\sigma(X + Y) \subseteq \sigma(X, Y) = \sigma\left(\sigma(X) \cup \sigma(Y)\right),$$

that is the sigma algebra generated by the sum is contained in the sigma algebra generated by both random variables. In particular if X and Y are r.v, then even the sum is a r.v.

Proof. The proof is the following.

- Let us set $Z = X + Y$.
- We just need to prove thanks to (2.7) that $Z^{-1}(A) \in \mathscr{F}$, for all A in a set of generator of $\mathfrak{B}(\mathbb{R})$.
- Let us define

  $$\mathcal{A} := \{(-\infty, x) : x \in \mathbb{R}\}.$$

  This is our sets of generator.

- We have that

$$Z^{-1}((-\infty, c)) = \{X + Y < c\} = \bigcup_{q \in \mathbb{Q}} \{X + q < c\} \cap \{Y < q\},$$

and the last inequality hold true because we have that for all $a \in \mathbb{R}$ and $b \in \mathbb{R}$ and $c \in \mathbb{R}$,

$$a + b < c \iff \exists\ q \in \mathbb{Q}\ \ s.t.\ \ a + q < c\ \ and\ \ b < q.$$

- So $Z^{-1}((-\infty, c)) \in \sigma(X, Y)$ for all $c \in \mathbb{R}$, so we have the thesis.

$\square$

5.2 Topological Results (for vector normed spaces)

- Let us have $(E, \|\cdot\|)$ a Banach space, but completeness.
- E have a topology that is induced by the norm. We indicate as

$$B(e, r) \coloneqq \{x :\ \|x - e\| < r\} = \{\ \|x - e\| < r\ \}$$

the ball centered in $e \in E$ of radius r, and it is well known that it is a base of the topology. we indicate the topology generated by the ball as $\mathcal{B}$ (that is the set of the open sets).

- Let us consider now E^n. It has by definition a topology, that is the product topology.
- This product topology is defined in this way. Let us consider p_i the canonical projection.
- Let us define

$$\mathcal{A} \coloneqq \{\mathcal{T}\ \ topology\ \ s.t.\ \ p_1, .., p_n : (E^n, \mathcal{T}) \to (E, \mathcal{B})\ \ are\ \ continuous\}.$$

- Now, let us define

$$\mathcal{P} \coloneqq \bigcap_{\mathcal{T} \in \mathcal{A}} \mathcal{T}$$

Well, $\mathcal{P}$ is the product topology.

- We can indicate more generally as

$$\mathcal{T}(\mathcal{R})$$

the littlest *topology* that contain $\mathcal{R}$, with $\mathcal{R} \subseteq \mathcal{P}(E)$, in analogy of what we did in Definition (7).

- Since we have a finite product, it is immediate to show that

$$\mathcal{P} = \mathcal{T}\left(\{\times_{i=1}^{n} A_i,\ \ A_i \in \mathcal{B}\}\right),$$

  that is the product topology is the smallest topology which contains the product between the open sets of E.

- Now we want to prove that the product topology on E^n in this case is given by a norm.

  Proposition 5.2. *Let us define*

$$n : E^n \to \mathbb{R},\ \ n(e_1, .., e_n) = max_{i=1,..,n} \|e_i\|\,.$$

  Then n is a norm.

  Proof. Obvious. □

  We indicate as $\|\cdot\|_\infty$ this norm.

- Let us denote as $\mathcal{P}_{\|\cdot\|_\infty}$ the topology induced by this norm (the one defined by the ball on E^n).

- **Proposition 5.3.** $\mathcal{P} = \mathcal{P}_{\|\cdot\|_\infty}$.

  Proof. The proof is not hard, we prove the double inclusion.

  - $\subseteq$:
    We just need to prove that projections are continuous functions. Since we have in this case for all $i = 1, .., n$ that

$$p_i : (E^n, \mathcal{P}_{\|\cdot\|_\infty}, \|\cdot\|_\infty) \to (E, \mathcal{B}, \|\cdot\|)$$

    is a function between normed spaces, we just need to check the $\epsilon - \delta$ definition, and this is immediate.
    So we obtain one inclusion.
  - " $\supseteq$ ":
    We just need to see if the ball with respect to $\|\cdot\|_\infty$ are elements of $\mathcal{P}$. We have for all $e = (e_1, .., e_n) \in E^n$ and for all $r > 0$ real,

$$\begin{aligned} &B_{\|\cdot\|_\infty}(e, r) = \{x : \|x - e\|_\infty < r\} = \\ &= \{x = (x_1, .., x_n) : \forall i = 1, .., n,\ \|x_i - e_i\| < r\} = \\ &= \{x_1 : \|x_1 - e_1\| < r\} \times \cdots \times \{x_n : \|x_n - e_n\| < r\} = \\ &= \bigcap_{i=1}^{n} p_i^{-1}(\{y \in E : \|y - e_i\| < r\}) \in \mathcal{P} \end{aligned}$$

    and the last set belong to $\mathcal{P}$ because is intersection of a finite number of open sets.

  □

5.3 Relation product sigma algebra and sigma algebra induced by the topology

- Let us consider again $(E,\|\cdot\|)$, that is a vectorial normed space.
- Let us consider $Open(E)$ the topology induced by the norm.
- Let us consider $Open(E^n)$ the *topology* on E^n induced by the norm $\|\cdot\|_\infty$, that is the same as the canonical topology on the product space E^n.
- As always we denote as $\mathfrak{B}(E) := \sigma(Open(E))$, that is $\mathfrak{B}(E)$ is the smallest $\sigma - algebra$ which contains the open sets of E.
- Now, we want to compare

  $$\mathfrak{B}(E^n) := \sigma(Open(E^n))$$

  and

  $$\bigotimes_n \mathfrak{B}(E) := \bigvee_{i=1}^n p_i^{-1}(\mathfrak{B}(E)) = \sigma\left(\{\times_{i=1}^n A_i \ : \ A_i \in \mathfrak{B}(E)\}\right) = \sigma(\mathfrak{B}(E)^n).$$

Proposition 5.4. *Let us suppose that E is separable. Then $\mathfrak{B}(E^n) = \bigotimes_n \mathfrak{B}(E)$.*

Proof. We show the double inclusion.

- " $\supseteq$ "

  Since E is separable, it is base numerable (it's metric), that is

  $$\exists N \subseteq Open(E), \ countable: \ \forall A \in Open(E), \ \exists J \subseteq N: \ A = \bigcup_{B \in N} B.$$

- Now, N base for $Open(E) \implies N^n$ is a base for $Open(E^n)$, because the product of basis is a base for the product space (it is immediate to show).
- So, we have that $Open(E) \subseteq \sigma(N^n) \implies \mathfrak{B}(E^n) = \sigma(N^n)$.
- We define

  $$\mathcal{A} := \left\{\times_{i=1}^n A_i \ : \ A_i \in \mathfrak{B}(E)\right\} = \mathfrak{B}(E)^n,$$

  so we have that $\otimes_n \mathfrak{B}(E) = \sigma(\mathcal{A})$.
- Since $N \subseteq Open(E) \subseteq \mathfrak{B}(E)$, we have that $N^n \subseteq \mathcal{A}$, so we conclude seeing that $\mathfrak{B}(E^n) = \sigma(N^n) \subseteq \sigma(\mathcal{A}) = \otimes_n \mathfrak{B}(E)$.
- " $\subseteq$ "

  Since $Opens(E)$ is a $\pi - system$ for $\mathfrak{B}(E)$, we have that $Open(E)^n$ is a $\pi - system$ for $\otimes_n \mathfrak{B}(E)$ because of Corollary (2.13).
- Moreover, $Open(E)^n \subseteq Open(E^n)$, so $\otimes_n \mathfrak{B}(E) = \sigma(Open(E)^n) \subseteq \sigma(Open(E^n)) = \mathfrak{B}(E^n)$, so we have finished.

□

5.4 Measurability of the sum

Now we can prove some useful theorem.

- Let us have a fixed *Measurability Space* $(\Omega . \mathscr{F})$.
- Let us consider $(E, \mathfrak{B}(E))$, with $(E, \|\cdot\|)$ a separable normed space.
- Let $X_i : \Omega \to E$ be a measurable function, for all $i = 1, .., n$.
- Let us consider the function X defined in this way,

$$\begin{array}{rcl} X : (\Omega, \mathscr{F}) & \to & (E^n, \bigotimes_n \mathfrak{B}(E)) \\ \omega & \to & (X_1(\omega), .., X_n(\omega)). \end{array}$$

  We know that such function is $(\mathscr{F}, \bigotimes_n \mathfrak{B}(E)) - \textit{measurable}$ and $\sigma(X_1, .., X_n) = \sigma(X)$ from Proposition (2.18).

- Let us consider $(H, Open(H))$ another topological space.
- Let us consider $f : (E, Open(E)) \to (H, Open(H))$ a continuous function.

Theorem 5.5. *We have that $f \circ X$ is $(\mathscr{F}, \mathfrak{B}(H))-$measurable.*

Proof. Given our work before, the proof is a simple path.

- We firstly show that f is $(\mathfrak{B}(E^n), \mathfrak{B}(H))$ measurable. Let's see.
  - We have that $\sigma(Open(H)) = \mathfrak{B}(H)$ and
  - $f^{-1}(Open(H)) \subseteq Open(E^n) \subseteq \mathfrak{B}(E^n)$.
  - Condition above implies that $f^{-1}(\mathfrak{B}(H)) = \sigma(f) \subseteq \mathfrak{B}(E^n)$ thanks to (2.7).
  - So we have proved that f is $(\mathfrak{B}(E^n), \mathfrak{B}(H))$-measurable, as we wanted.
- Moreover, we have that $\mathfrak{B}(E^n) = \bigotimes_n \mathfrak{B}(E)$ thanks to Proposition (5.4), so we can write that f is also $(\bigotimes_n \mathfrak{B}(E), \mathfrak{B}(H))$ measurable.
- So we have the following composition of function,

$$(\Omega, \mathscr{F}) \xrightarrow{X} (E^n, \otimes_n \mathfrak{B}(E)) \xrightarrow{f} (H, \mathfrak{B}(H)),$$

  and since f and X are measurable, we have that $f \circ X$ is measurable, as we wanted.

□

- Let us suppose now that $H = E^n$, with the same topology that we had at the beginning.
- Now f will be a function like $f = (f_1, .., f_n)$, with $f_i = \pi_i \circ f$, with $\pi : E^n \to E$ the canonical projection such that $\pi_i(e_1, .., e_n) = e_i$.
- In this case we have that $f \circ X = (f_1 \circ X, .., f_n \circ X)$.

Corollary 5.6. *Let us suppose that f is continuos, invertible, and its inverse is a continuous function. Then*

$$\sigma(X_1, .., X_n) = \sigma(f_1 \circ X, .., f_n \circ X).$$

Proof. We observe that f and f^{-1} are measurable function. So we have thanks to Corolllary (2.3) and Proposition (2.18)

$$\sigma(X_1, .., X_n) = \sigma(X) = \sigma(f \circ X) = \sigma(\pi_1 \circ (f \circ X), .., \pi_n \circ (f \circ X)) = \sigma(f_1 \circ X, ..., f_n \circ X),$$

as we wanted. □

Corollary 5.7 (measurability of the sum)**.** *Let us suppose that E is a vectorial space. Then the random variable*

$$Z := \sum_{i=1}^{n} X_i,$$

is $(\mathscr{F}, \mathfrak{B}(E))$ – measurable.

Proof. Let us consider

$$f : (E^n, Open(E^n)) \to (E, Open(E)), \;\; f(e_1, .., e_n) = \sum_{i=1}^{n} e_i.$$

It is immediate that this is continuous, so we can apply the *Theorem* above. □

6 Characteristic Function

Let $X : (\Omega, \mathscr{F}, \mathbb{P}) \to (\mathbb{R}^m, \mathfrak{B}(\mathbb{R}^m))$ be a r.v, let us set $\mu := \mathbb{P}_X$.

Definition 19. We define for all $\theta \in \mathbb{R}^m$,

$$\hat{\mu}(\theta) = \int_{\mathbb{R}^m} e^{i\langle \theta, x\rangle} \mu(dx) = \mathbb{E}[\; e^{i\langle \theta, X\rangle} \;].$$

6.1 Property

1. X and Y are r.v. independents. Then $\hat{\mu}_{X+Y}(\theta) = \hat{\mu}_X(\theta)\hat{\mu}_Y(\theta)$.
2. $\forall \theta \in \mathbb{R}^n$ we have $\;\hat{\mu}(\theta) = \hat{\nu}(\theta) \implies \mu \equiv \nu$.
3. $X_1, .., X_n$ are $r.v.'s$ respectively with law $\mu_1, ..., \mu_n$.
   We denote as μ the law of $(X_1, .., X_n)$.
   Then $X_1, .., X_n$ are independent if, and only if for all $\theta = (\theta_1, .., \theta_n) \in \mathbb{R}^n$, we have

$$\hat{\mu}(\theta) = \prod_{i=1}^{n} \hat{\mu}_i(\theta_i).$$

4. Let $b \in \mathbb{R}^k$ be a vector, and let $A \in \mathbb{R}^{k\times m}$ a matrix. Let us set $Y := AX + b$. Then for all $\theta \in \mathbb{R}^k$ we have

$$\hat{\mu}_Y(\theta) = \mathbb{E}[e^{i\langle \theta, Y\rangle}] = \mathbb{E}[e^{i\langle \theta, AX+b\rangle}] = e^{i\langle \theta, b\rangle}\mathbb{E}[e^{i\langle A^*\theta, X\rangle}] = e^{i\langle \theta, b\rangle}\hat{\mu}_X(A^*\theta).$$

   We remember that A^* is the transpose of A.

5. $X = (X_1, .., X_n)$ a random vector. Then for all $i = 1, .., n$ we have that

$$\hat{\mu}_{X_i}(\theta) = \hat{\mu}_X(0, .., \underbrace{\theta}_{i-th}, .., 0).$$

We prove now a little lemma. Let $Y : (\Omega, \mathscr{F}, \mathbb{P}) \to (\mathbb{R}^m, \mathfrak{B}(\mathbb{R}^m))$ be a r.v, let us set $\mu := \mathbb{P}_Y$.

Lemma 6.1. *We have that Y is constant and $Y = a$ almost certain, with $a \in \mathbb{R}^m$ if, and only if for all $\theta \in \mathbb{R}^m$ we have that $\hat{\mu}_Y(\theta) = e^{i\langle \theta, a\rangle}$.*

Proof. We prove both implication. We denote as μ_Z the law of Z and as μ_Y the law of Y.

- If $Y = a$ almost certain, then the thesis is obvious.
- For the other, let us set $Z(\omega) = a$ for all $\omega \in \Omega$.
- Then we have that $\hat{\mu}_Z \equiv \hat{\mu}_Y$, and this implies for our property (6.1) that $\mu_Z \equiv \mu_Y$.
- But then, we have that

$$1 = \mathbb{P}(Z = a) = \mu_Z(a) = \mu_Y(a) = \mathbb{P}(Y = a)$$

  so $Y = a$ almost certain, and this means that it is constant almost certain.

□

6.2 Gaussian Law

Let us have μ a probability on $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$.

Definition 20. We say that μ is $N(a, \sigma^2)$ (normal, with mean a and variance σ^2), with $\sigma > 0$ and $a \in \mathbb{R}$, if it has density with respect to Lebesgue measure given by

$$f_{a,\sigma^2}(x) := \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Remark 11. If we have

$$f_{0,1}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

then we have the following relation,

$$f_{a,\sigma^2}(x) = \frac{1}{\sigma} f_{0,1}\left(\frac{x-a}{\sigma}\right).$$

at this point, we have a probability μ on $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$, that is

$$\forall A \in \mathfrak{B}(\mathbb{R}), \ \ \mu(A) = \int_A f_{a,\sigma^2}(x)dx.$$

We want to compute its characteristic function $\hat{\mu}$. We have

Proposition 6.2.

$$\forall \theta \in \mathbb{R}, \ \ \hat{\mu}(\theta) = e^{i\theta a}e^{-\frac{1}{2}\sigma^2\theta^2}.$$

Remark 12. We observe that, if $X \sim N(a, \sigma^2)$ and $Y \sim N(b, \tau^2)$ and they are independent, then $X + Y \sim N(a + b, \sigma^2 + \tau^2)$. In fact if we denote respectively as μ_X and μ_Y and μ_{X+Y} their law, we have

$$\hat{\mu}_{X+Y}(\theta) = \hat{\mu}_X(\theta)\hat{\mu_Y}(\theta) = e^{i\theta(a+b)}e^{-\frac{1}{2}(\sigma^2+\tau^2)\theta^2}$$

and this last one is the characteristic function of a r.v. $N(a+b, \sigma^2+\tau^2)$, so $(X+Y)$ it have to be a variable $N(a + b, \sigma^2 + \tau^2)$. We have used our properties in Section (6.1).

Now let's talk about *Gaussian Vectors.* We suppose to have a fixed probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Definition 21 (Standard Gaussian Vectors)**.** Let $Z := (Z_1, .., Z_n)$ be a random vector. We say that Z has the *Standard Gaussian Distribution* if $\mathbb{P}_Z$ has density

$$f_{\mathbf{0},I}(x_1, .., x_n) := \frac{1}{(2\pi)^{\frac{n}{2}}}e^{-\frac{1}{2}\|x\|_2^2} = \frac{1}{(2\pi)^{\frac{n}{2}}}e^{-\frac{1}{2}\langle x,x\rangle}.$$

for all $x \in \mathbb{R}^n$.

Remark 13. The definition above is equivalent to ask that the r.v. $Z_1, ..., Z_n$ are independent and for all i, we have that $Z_i \sim N(0, 1)$.

Remark 14. $Z_i \sim N(0, 1) \implies$ for all $\theta \in \mathbb{R}$, we have that $\hat{\mu}_{Z_i}(\theta) = e^{-\frac{1}{2}\theta^2}$.

Remark 15. For all $\theta = (\theta_1, ..., \theta_n) \in \mathbb{R}^n$, properties (6.1) implies that

$$\hat{\mu}_Z(\theta) = \prod_{i=1}^{n} \mu_{Z_i}(\theta_i) = \prod_{i=1}^{n} e^{-\frac{1}{2}\theta_i^2} = e^{-\frac{1}{2}\|\theta\|_2^2} = e^{-\frac{1}{2}\langle\theta,\theta\rangle}.$$

Definition 22. A random vector $Y = (Y_1, .., Y_m)$ is said to be *Gaussian* if we can write

$$Y = AZ + b$$

where A is a $m \times n$ matrix, $b \in \mathbb{R}^m$, $Z = (Z_1, .., Z_n)$ is a $n-$*dimensional* standard gaussian vector.

Now let us suppose $n = m$, and $\det(A) \neq 0$. Then (3.5) hold true, and we have that Y has density. We observe preliminary that for all $z, w \in \mathbb{R}^n$

$$\langle A^{-1}z, A^{-1}w\rangle = \langle (A^{-1})^* A^{-1}z, w\rangle = \langle (A^*)^{-1}A^{-1}z, w\rangle = \langle \Gamma^{-1}z, w\rangle.$$

where we have set $\Gamma = AA^*$ in the above equality, so $\Gamma^{-1} = (A^*)^{-1}A^{-1}$. Given this, we can write down the density of Y as

$$f_{b,\Gamma}(y) := \frac{1}{|\det A|} f_{\mathbf{0},I}(A^{-1}(y-b)) = \tag{5}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{|\det A|} e^{-\frac{1}{2}\langle \Gamma^{-1}(y-b),(y-b)\rangle} = \tag{6}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{|\det \Gamma|^{\frac{1}{2}}} e^{-\frac{1}{2}\langle \Gamma^{-1}(y-b),(y-b)\rangle} \tag{7}$$

Lemma 6.3. *Let $Y = (Y_1, .., Y_m) = AZ + b$ be a Gaussian Vector, and A is an $m \times n$ matrix, Z is an n-dimensional Standard Gaussian Vector. Then*

- $b = (\mathbb{E}[Y_1], .., \mathbb{E}[Y_m])$.
- $Cov(Y_i, Y_j) = (A \cdot A^*)_{i,j} = A_i \cdot (A^*)^j$.

Proof. This is a simple check.

- $b = \mathbb{E}[Y]$ is obvious since $Z_i \sim N(0, 1)$.
- This is a simple check. Let $i \in \{1, .., m\}$ and $j \in \{1, .., m\}$ be two numbers. Then

$$\begin{aligned}
Cov(Y_i, Y_j) =& Cov\left(\sum_{k=1}^{n} A_{i,k}Z_k + b_i,\ \sum_{h=1}^{n} A_{j,h}Z_h + b_j\right) \underbrace{=}_{bilinearity} \\
&= \sum_{k,h=1}^{n} A_{i,k}A_{j,h} \underbrace{Cov(Z_k, Z_h)}_{=0 \ \ if \ \ h \neq k} = \\
&= \sum_{k=1}^{n} A_{i,k}A_{j,k} \underbrace{Cov(Z_k, Z_k)}_{=Var(Z_k)=1} \\
&= A_i \cdot (A_j)^* = A_i \cdot (A^*)^j = (AA^*)_{i,j}.
\end{aligned}$$

□

Remark 16. Given Y Gaussian Vector, we denote as $Cov(Y) = A \cdot A^*$. We observe that it is positive definite.

Now, given Γ $m \times m$ matrix positive definite and b a vector in $\mathbb{R}^m$, we ask ourselves if we can find a Gaussian random variable $Y = (Y_1, .., Y_m)$ such that $Cov(Y) = \Gamma$ and $\mathbb{E}[Y] = b$.

Lemma 6.4. *Let us have*

- Γ *$m \times m$ positive definite matrix.*
- *b vector in $\mathbb{R}^m$*

Then we can find a Gaussian Variable X such that $Cov(X) = \Gamma$ and $\mathbb{E}[X] = b$.

Proof. This is a linear algebra exercise, because we just need to find a matrix A such that $A \cdot A^* = \Gamma$. This is possible, and we can find it symmetric. Let's call such matrix $\sqrt{\Gamma}$. We obtain the thesis if we set

$$X := \sqrt{\Gamma} Z + b$$

with Z a Standard Gaussian Vector. □

Remark 17. We have seen that given a matrix Q positive definite and a vector $b \in \mathbb{R}^m$, we can find a r.v. $Y = (Y_1, .., Y_m)$ that is *Gaussian*, that $Cov(Y) = Q$ and $\mathbb{E}[Y] = b$. We denote this fact saying that $Y \sim N(b, Q)$, that is Y is normal, with mean b and covariance matrix Q.

Remark 18. A $n-$*dimensional* standard Gaussian Vector $Z = (Z_1, .., Z_n)$ is denoted by $Z \sim N(\mathbf{0}, I)$, with $\mathbf{0}$ the null vector of $\mathbb{R}^n$ and I the identical matrix $n \times n$.

Now, let us calculate the characteristic function of $Y = (Y_1, .., Y_k)$, with Y a Gaussian Vector. So we have that $Y = AZ + b$, with A matrix $k \times m$ and $b \in \mathbb{R}^k$ and Z $m-$*dimensional* standard Gaussian Vector. Thanks to Properties (6.1), we have for all $\theta \in \mathbb{R}^k$.

$$\hat{\mu}_Y(\theta) = e^{i\langle \theta, b\rangle} \hat{\mu}_Z(A^*\theta) = e^{i\langle \theta, b\rangle} e^{-\frac{1}{2}\langle A^*\theta, A^*\theta\rangle} = e^{i\langle \theta, b\rangle} e^{-\frac{1}{2}\langle \Gamma\theta, \theta\rangle}. \tag{8}$$

where we have set $\Gamma := A \cdot A^*$, that is $\Gamma = Cov(Y)$.

Let us have $Y = (Y_1, .., Y_m) : (\Omega, \mathscr{F}, \mathbb{P}) \to (\mathbb{R}^m, \mathfrak{B}(\mathbb{R}^m))$ a r.v. Let μ_Y be the law of Y.

Proposition 6.5. *The following statements are equivalent.*

1. *$Y \sim N(b, \Gamma)$, that is Y is Gaussian, $b = \mathbb{E}[Y]$ and $\Gamma = Cov(Y)$*
2. *$\forall \theta \in \mathbb{R}^m$, $\hat{\mu}_Y(\theta) = e^{i\langle \theta, b\rangle} e^{-\frac{1}{2}\langle \Gamma\theta, \theta\rangle}$, with $b \in \mathbb{R}^m$ and Γ a $m \times m$ positive semi-definite matrix.*
3. *$\forall \theta \in \mathbb{R}^m$, we have that $\langle \theta, Y\rangle$ is a Gaussian r.r.v, that is $\langle \theta, Y\rangle \sim N(a_\theta, \sigma_\theta^2)$, with $a_\theta \in \mathbb{R}$ and $\sigma_\theta^2 \geq 0$.*

Proof. It is easy.

- 1) $\implies$ 2). This is the count above in (8).
- 2) $\implies$ 1). We suppose $\det \Gamma \neq 0$. The case $\det \Gamma = 0$ is degenerate and we study it another time.
  - We can find thanks to (6.4) a Gaussian Variable $X = AZ + b$, with A $m \times m$ matrix such that $AA^* = \Gamma$ and $Z \sim N(\mathbf{0}, I)$ and $b \in \mathbb{R}^m$.

- We see immediately that for all $\theta \in \mathbb{R}^m$, we have
$$\hat{\mu}_X(\theta) = \hat{\mu}_Y(\theta).$$
and thanks to properties (6.1) we have that $\mu_X = \mu_Y$.
- Now we know that X has density that is given by (5), so even Y has the same density because $\mu_X = \mu_Y$, and this density is $f_{b,\Gamma}$.
- Let us consider $W := A^{-1}Y - A^{-1}b = A^{-1}(Y - b)$. Thanks to (3.5), we know that even W has density, and this density is given by
$$\begin{aligned} g(x) &= \frac{1}{|\det(A^{-1})|} f_{b,\Gamma}(A(x + A^{-1}b)) = \\ &= \frac{1}{|\det(A^{-1})|} f_{b,\Gamma}(Ax + b) = \\ &= \frac{1}{|\det(A^{-1})|}\frac{1}{|\det(A)|} f_{\mathbf{0},I}(A^{-1}([Ax + b] - b)) \\ &= f_{\mathbf{0},I}(x), \end{aligned}$$
so for our definition (21), we have that W is a standard Gaussian Vector, that is $W \sim N(\mathbf{0}, I)$.
- Then we have
$$Y = \underbrace{A}_{matrix} \underbrace{(A^{-1}[Y - b])}_{N(\mathbf{0},\mathbf{I})} + \underbrace{b}_{vector},$$
that is $Y \sim N(b, \Gamma)$, and that is the thesis.
- Now we suppose $\det \Gamma = 0$.
- If $\Gamma = 0$, then by Lemma (6.1) we obtain that Y is constant, so it is Gaussian (by definition (?, that is we don't know)).
- Otherwise, if $rk(\Gamma) > 0$ we firstly suppose that Γ is diagonal, so we have
$$\Gamma = diag(\gamma_1^2, .., \gamma_k^2, 0, .., 0) \;\; with \;\; 1 \leq k < m \;\; and \;\; \gamma_i > 0.$$
- One of the Properties in (6.1) say us that the marginal law of Y are Gaussian 1-dimensional, so we have that
$$\begin{cases} Y_i \sim N(b_i, \gamma_i^2) \;\; if \;\; 1 \leq i \leq k \\ Y_i \equiv b_i, \;\; that \;\; is \;\; Y_i \sim N(b_i, 0) \;\; if \;\; k < i \leq m. \end{cases}$$
We have used even *Lemma* (6.1).
- Let use set
$$Z_i = \frac{Y_i - b_i}{\gamma_i} \;\; for \;\; 1 \leq i \leq k,$$
$$Z = (Z_1, .., Z_k)^*,$$
$$b = (b_1, .., b_m)^*,$$
$$A = \left[\begin{array}{c} B \\ \hline C \end{array}\right], \;\; with \;\; B = \begin{bmatrix} \gamma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \gamma_k \end{bmatrix} \in \mathbb{R}^{k \times k} \;\; and \;\; C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{(m-k) \times k}.$$

We remember that $*$ means *transpose*.

- Now it is a little count to show that $Y = AZ + b$, and Z is a $k-dimensional$ Standard Gaussian Vector, so Y is a Gaussian Vector.
- If Γ is not diagonal or it is not in our form, we can find an orthogonal matrix O such that $O\Gamma O^*$ is diagonal as we want, because by definition Γ is symmetric. We define $\tilde{\Gamma} = O\Gamma O^*$.
- Let us define $T = OY$. We have that the characteristic function of μ_T become, thanks to our property (6.1),

$$\hat{\mu}_T(\theta) = \hat{\mu}_Y(O^*\theta) = e^{i\langle O^*\theta, b\rangle}e^{-\frac{1}{2}\langle \Gamma O^*\theta, O^*\theta\rangle} = e^{i\langle \theta, Ob\rangle}e^{-\frac{1}{2}\langle \tilde{\Gamma}\theta, \theta\rangle}, \ \ \forall \theta \in \mathbb{R}^m.$$

- So, we have discovered that T is a vector that has a characteristic function as the one in hypothesis 2, with $Ob \in \mathbb{R}^m$ and $\tilde{\Gamma} = O\Gamma O^*$ a $m \times m$, positive semi-definite matrix. So we have, because of what we have just proved that

$$OY = AZ + Ob,$$

with A matrix $m \times k$, and Z a $k - dimensional$ Standard Gaussian Vector, with $k = rk(\Gamma)$.

- From the last identity, we discover that

$$Y = [(O^*)A]Z + b,$$

that is Y is a Gaussian Vector.

- 1) $\implies$ 3). Obvious, we just need to observe that a sum of real Gaussian variable is still Gaussian.
- 3) $\implies$ 2). We calculate the characteristic function of Y. We proceed in this way.
  - We observe that Y_i is Gaussian for all i (take $\theta = (0, , \underbrace{1}_{i-th}, ..0)$).
  - This implies that $\mathbb{E}[Y_i]$ is well defined (we need $Y_i \in L^1$), as well $Cov(Y_i, Y_j)$ (we need $Y_i \in L^2$ and $Y_j \in L^2$).
  - Let us set $b := (\mathbb{E}[Y_1], .., \mathbb{E}[Y_m])$ and $\Gamma := Cov(Y) = [Cov(Y_i, Y_j)]_{i,j=1,..,m}$.
  - It is well known that $Cov(Y)$ is a positive semi-definite matrix.
  - Also, we have that for all $\theta \in \mathbb{R}^m$,

$$Var(\langle Y, \theta\rangle) = Cov(\langle Y, \theta\rangle, \ \langle Y, \theta\rangle) = \langle Cov(Y)\theta, \ \theta\rangle = \langle \Gamma\theta, \ \theta\rangle,$$
$$\mathbb{E}[\langle Y, \theta\rangle] = \langle \mathbb{E}[Y], \theta\rangle = \langle b, \theta\rangle.$$

  - So, if we recall (6.2), we have for all $\theta \in \mathbb{R}^m$ that

$$\hat{\mu}_Y(\theta) = \mathbb{E}[e^{i\langle Y, \theta\rangle \cdot 1}] = \mu_{\langle Y, \theta\rangle}(1) = e^{i\cdot 1\mathbb{E}[\langle Y, \theta\rangle]}e^{-\frac{1}{2}Var(\langle Y, \theta\rangle)} = e^{i\cdot\langle b, \theta\rangle}e^{-\frac{1}{2}\langle \Gamma\theta, \ \theta\rangle}$$

and this is the thesis. We have used the integration with respect to a probability law, that is (3.1).

□

Proposition 6.6 (Normal Law under Affine Trasformation)**.** . *Let $L(x) := Ax + b$ be an affine transformation with C a matrix $k \times m$ and $b \in \mathbb{R}^k$ a vector and let $Y = (Y_1, .., Y_m) \sim N(a, \Gamma)$ be a Gaussian variable. Then $X := L \circ Y$ is still Gaussian, and its law is given by $N(Aa + b, A\Gamma A^*)$.*

Proof. We just need to compute the characteristic function of X. Since L is affine, using a property in (6.1) we have for all $\theta \in \mathbb{R}^k$,

$$\hat{\mu}_X(\theta) = e^{i\langle b,\theta\rangle}\hat{\mu}_Y(A^*\theta) = e^{i\langle b,\theta\rangle}\left(e^{i\langle a,A^*\theta\rangle}e^{-\frac{1}{2}\langle \Gamma A*\theta, A^*\theta\rangle}\right) = e^{i\langle Aa+b,\theta\rangle}e^{-\frac{1}{2}\langle A\Gamma A^*,\theta\rangle},$$

so $X = AY + b \sim N(Aa + b, A\Gamma A^*)$. □

Remark 19. We observe that given $Y = (Y_1, .., Y_m) \sim N(a, \Gamma)$, we have that its $k - th$ marginal is given by

$$\mu_{Y_k}(\theta) = e^{ia_k\theta}e^{-\frac{1}{2}\Gamma_{k,k}\theta^2}, \ \ \forall\theta \in \mathbb{R}$$

thanks to property (6.1) and the formula (8).

From this remark and from property (6.1), it is straightforward the following

Proposition 6.7. *Let $Y = (Y_1, .., Y_m) \sim N(b, \Gamma)$ be a Gaussian Vector. Then the following statement are equivalent,*

1. *$Y_1, .., Y_m$ are independent.*
2. *$Y_1, .., Y_m$ are uncorrelated, that is Γ is diagonal.*

Proof. we have

- 1) $\implies$ 2). It is always true.
- 2) $\implies$ 1).
  - We just need to check if for all $\theta = (\theta_1, .., \theta_m) \in \mathbb{R}^m$, we have

    $$\hat{\mu}_Y(\theta) = \prod_{k=1}^{m}\hat{\mu}_{Y_k}(\theta_k)$$

    because we can conclude thanks to one of our property (6.1).
  - But, if Γ is diagonal, then

    $$\hat{\mu}_Y(\theta) = e^{i\langle\theta,b\rangle}e^{-\frac{1}{2}\langle\Gamma\theta,\theta\rangle} = e^{\sum_{k=1}^m i\theta_k b_k}e^{\sum_{k=1}^m -\frac{1}{2}\Gamma_{k,k}\theta_k^2} =$$
    $$= \prod_{k=1}^{m} e^{ib_k\theta_k}e^{-\frac{1}{2}\Gamma_{k,k}\theta_k^2} = \prod_{k=1}^{m}\hat{\mu}_{Y_k}(\theta_k)$$

    and for what we have said, this implies the thesis.

□

7 Stochastic Process

Definition 23 (Stochastic Process)**.** Let $X : \Omega \times T \to E$ be a function.
X is a *Stochastic Process* (S.P.) if for all $t \in T$, we have that

$$X_t \coloneqq X|_{\Omega \times \{t\}} : (\Omega, \mathscr{F}) \to (E, \mathscr{E})$$

is $\mathscr{F}$-measurable.

We can indicate X as $(X_t)_{t \in T}$.

7.1 Another Point of View, and the Law of a S.P.

We observe that there is another point of view connected to our definition, that is the following. We denote as

$$E^T \coloneqq \prod_{t \in T} E_t, \quad \mathscr{E}^{\otimes T} \coloneqq \bigotimes_{t \in T} \mathscr{E}_t$$

where $E_t = E$ for all t and $\mathscr{E}_t = \mathscr{E}$ for all t. We remember that $E^T \coloneqq \{f : T \to E\}$, and for all $t \in T$, the *projection* on E_t is defined as $\pi_t(f) \coloneqq f(t) = f|_t$, for all $f \in E^T$. We can see a *S.P.* X as a $\mathscr{F} - measurable$ function

$$\Phi_X : (\Omega, \mathscr{F}) \to (E^T, \mathscr{E}^{\bigotimes T})$$

defined as $\Phi_X(\omega)(t) \coloneqq X(\omega, t)$. The definitions are equivalent in the following sense.
Let $(\Omega, \mathscr{F})$ and $(E, \mathscr{E})$ be two measurable spaces. Let T be a set of index.

Proposition 7.1. *The following statement hold true.*

1. *Let us consider $X : T \times \Omega \to E$ a S.P, that is $\forall t \in T$, we have that $X_t = X|_{\{t\} \times \Omega}$ is $\mathscr{F} - measurable$.*

   Then the function

$$\begin{array}{rcc} \Phi : (\Omega, \mathscr{F}) & \to & (E^T, \mathscr{E}^{\otimes T}) \\ \omega & \to & \Phi(\omega) : T \to E \\ & & t \to X(t, \omega) \end{array}$$

   is $\mathscr{F} - measurable$.

2. *Let us consider a function*

$$\Phi : (\Omega, \mathscr{F}) \to (E^T, \mathscr{E}^{\otimes T})$$

   that is $\mathscr{F} - measurable$.

   Then the function

$$\begin{array}{rcc} X : T \times \Omega & \to & E \\ (\omega, t) & \to & (\Phi(\omega))(t) \end{array}$$

   is a S.P.

Proof. The proof are simple.

- Let's start with 1.

  For our Lemma (2.18), we just need to check that for all $t \in T$, the function $(\pi_t \circ \Phi)$ is $\mathscr{F} - measurable$. We have

  $$(\pi_t \circ \Phi)(\omega) = \pi_t(\Phi(\omega)) = (\Phi(\omega))(t) = X(t, \omega) = X_t(\omega)$$

  and this last one is measurable because of our hypothesis, so Φ is $\mathscr{F} - measurable$.

- Now we prove 2.

  We just need to observe that for all $\omega \in \Omega$,

  $$X_t(\omega) = X|_{\{t\}\times\Omega}(t, \omega) = X(t, \omega) = (\Phi(\omega))(t) = \pi_t(\Phi(\omega)) = (\pi_t \circ \Phi)(\omega),$$

  and the last function is $\mathscr{F} - measurable$ because it is composition of measurable function, so for all $t \in T$ we have that X_t is measurable, so X is a *S.P.*

□

Remark 20. Given a *S.P.*, let's say X, we denote the function Φ defined above as Φ_X.

Now, let us suppose that we have a probability measure on $(\Omega, \mathscr{F})$, let's say $\mathbb{P}$.

Definition 24 (Law of a S.P.)**.** Let X be a *S.P.* We define the *Law* of the process as the measure of probability $\mathbb{P}_{\Phi_X}$ on the measurable space $(E^T, \mathscr{E}^{\otimes T})$.

Remark 21. So, given $A \subseteq E^T$, with $A \in \mathscr{E}^{\otimes T}$, we have that

$$\mathbb{P}_{\Phi_X}(A) = \mathbb{P}\left(\{\omega |\ \Phi_X(\omega) \in A\}\right),$$

that is $\mathbb{P}_{\Phi_X}$ calculate the probability that the function Φ_X maps some elements of Ω in a bunch of fixed function, that is A.

Remark 22. Let us consider $T = [0, +\infty)$, or $[0, t_0]$ with t_0 a real number, or let T be a subset of $\mathbb{N}$. We observe that definition (23) implies that for all $t_1 < t_2 ... < t_n \in T$, we have

$$(X_{t_1}, ..., X_{t_n}) : (\Omega, \mathscr{F}) \to (E^n, \otimes^n \mathscr{E})$$
$$(X_{t_1}, ..., X_{t_n})(\omega) := (X_{t_1}(\omega), ..., X_{t_n}(\omega))$$

is $\mathscr{F}$-measurable.

Let us consider

$$\mathcal{S} := \{(t_1, t_2, ..., t_n) \ | \ t_1 < t_2 < ..., < t_n, \ t_i \in T\}. \tag{9}$$

We define for a generic $(t_1, ..., t_n) := \bar{t} \in \mathcal{S}$ the r.v

$$X_{\bar{t}} := (X_{t_1}, ..., X_{t_n}) \tag{10}$$

Definition 25 (Finite Dimensional Distribution)**.** The set

$$\{ \ \mathbb{P}_{X_{\bar{t}}} \ | \ \bar{t} \in \mathcal{S} \ \}$$

Is the set of finite dimensional distribution of the process X. We observe that $\mathbb{P}_{X_{\bar{t}}}$ is a probability on the measurable space $(E^n, \mathscr{E}^{\otimes^n})$, and for all $A \in \mathscr{E}^{\otimes^n}$, we have that $\mathbb{P}_{X_{\bar{t}}}(A) = \mathbb{P}$

$$\mathbb{P}_{X_{\bar{t}}}(A) = \mathbb{P}(X_{\bar{t}} \in A) = \mathbb{P}((X_{t_1}, .., X_{t_n}) \in A).$$

Definition 26 (Realization (or Trajectory) of a S.P.)**.** Let $X : \Omega \times T \to E$ a S.P.

For all $\omega \in \Omega$, the function

$$\begin{aligned} X_\omega \coloneqq X|_{\{\omega\}\times T} : T &\to E \\ t &\to X_t(\omega) \end{aligned}$$

is a realization (or a trajectory) of the S.P.

Definition 27 (Continuity of a S.P.)**.** Let us suppose now that T and E are topological space, and $\mathcal{T} = \mathfrak{B}(T)$ and $\mathscr{E} = \mathfrak{B}(E)$. Let X be a S.P as above.

X is continuous if every trajectory is a continuous function. It is a.c (almost certain) continuous if $\mathbb{P}(\{\omega| \ X_\omega \ \textit{is continuous}\}) = 1$.

Definition 28 (ca'dla'g)**.** If $T \subset [0, +\infty)$, a SP is ca'dla'g if his trajectory are right-C^0, and the limit exists and is bounded on the left.

Definition 29 (Measurable Process)**.** Let $(T, \mathcal{T})$ be a measurable space. Let X be a *S.P.*. We say that X is measurable if the function

$$X : (T \times \Omega, \mathcal{T} \otimes \mathscr{F}) \to (E, \mathscr{E})$$

is $\mathcal{T} \otimes \mathscr{F} - measurable$.

Now, let X and Y be two *S.P.*

Definition 30 (Equivalent Process)**.** We say that X and Y are equivalent if they have the same finite dimensional distributions.

Definition 31 (Modification)**.** We say that X and Y are modification one of the another if for all $t \in T$, we have

$$\mathbb{P}(X_t = Y_t) = 1$$

Definition 32 (Indistinguishability)**.** We say that X and Y are indistinguishable if

$$\mathbb{P}(\forall t \in T, \ \ X_t = Y_t) = 1$$

The finite dimensional distribution are linked to the law of the *S.P.* in the following sense. Let $(\Omega_1, \mathscr{F}_1, \mathbb{P})$ and $(\Omega_2, \mathscr{F}_2, \mathbb{Q})$ be two probabilistic spaces, let $(E, \mathscr{E})$ be a measurable space and let T be a set of indexes. Let

$$\begin{aligned} X : T \times \Omega_1 \to E \\ Y : T \times \Omega_2 \to E \end{aligned}$$

be two $S.P.$ Let us set

$$\begin{aligned}
&\mathcal{S} = \{(t_1,..,t_n)|\ \forall i,\ \ t_i \in T \ \ and \ \ t_i \neq t_j \iff i \neq j\}, \\
&\mathcal{S}_X = (\mathbb{P}_{X_{\bar{t}}})_{\bar{t}\in\mathcal{S}} \\
&\mathcal{S}_Y = (\mathbb{Q}_{Y_{\bar{t}}})_{\bar{t}\in\mathcal{S}}
\end{aligned}$$

that is $\mathcal{S}$ is the set of every tuples of elements of T that have every element distinct from each other, $\mathcal{S}_X$ is the sequence of finite dimensional distribution of X indexed by $\mathcal{S}$ and $\mathcal{S}_\mathcal{Y}$ is the sequence of finite dimensional distribution of Y indexed by $\mathcal{S}$. Let

$$\mathbb{P}_{\Phi_X}$$
$$\mathbb{Q}_{\Phi_Y}$$

be respectively the law of the process X and the process Y on the measurable space $(E^T, \mathscr{E}^{\otimes T})$. This law is defined in Definition (24).

Lemma 7.2. *Let us have $(E^T, \mathscr{E}^{\otimes T})$ the product space and let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probabilistic space. Let us have*

- *$Z : T \times \Omega \to E$ a S.P.*
- *$A_1, .., A_n \in \mathscr{E}$,*
- *$t_1, .., t_n \in T$ that are distinct, that is $t_i = t_j \iff i = j$, and let us set $\bar{t} := (t_1, .., t_n)$,*
- *for all i, let us take $\pi_{t_i} : E^T \to E$ the canonical projection, that is $\pi_{t_i}(f) := f(t_i) = f|_{t_i}$.*
- *let us define*

$$A := \bigcap_{i=1}^{n} \pi_{t_i}^{-1}(A_i)$$

Then we have the following equality,

$$\{\omega|\ \Phi_Z(\omega) \in A\} = \{\omega|\ Z_{\bar{t}}(\omega) \in \bigtimes_{i=1}^{n} A_i\},$$

with $Z_{\bar{t}}(\omega) := (Z_{t_1}(\omega), .., Z_{t_n}(\omega))$.

Proof. Let's start.

- We remember the following easy equality. Given $(f_1, .., f_n) : \Omega \to C_1 \times .. \times C_n$ a function and given $A_1 \subseteq C_1, ..., A_n \subseteq C_n$, we have

$$\begin{aligned}
&(f_1(\omega), .., f_n(\omega)) \in \bigtimes_{i=1}^{n} A_i \iff \forall i = 1, .., n\ \ , f_i(\omega) \in A_i \iff \\
&\forall i = 1, .., n\ \ , \omega \in f_i^{-1}(A_i) \iff \omega \in \bigcap_{i=1}^{n} f_i^{-1}(A_i).
\end{aligned}$$

- Now, we have the following chain of implication,

$$\Phi_Z(\omega) \in A \iff \forall i = 1, .., n \;\; \pi_{t_i}(\Phi_Z(\omega)) \in A_i \iff$$
$$\forall i = 1, .., n \;\; (\Phi_Z(\omega))(t_i) \in A_i \iff \forall i = 1, .., n \;\; Z(t_i, \omega) = Z_{t_i}(\omega) \in A_i \iff$$
$$i = 1, .., n, \;\; \omega \in Z_{t_i}^{-1}(A_i) \iff \omega \in \bigcap_{i=1}^{n} Z_{t_i}^{-1}(A_i) \iff \underbrace{(Z_{t_1}, .., Z_{t_n})}_{Z_{\bar{t}}}(\omega) \in \bigtimes_{i=1}^{n} A_i,$$

so

$$\Phi_Z(\omega) \in A \iff Z_{\bar{t}}(\omega) \in \bigtimes_{i=1}^{n} A_i,$$

and this conclude.

□

Proposition 7.3. *We have that the following statements are equivalent,*

1. $\mathbb{P}_{\Phi_X} = \mathbb{Q}_{\Phi_Y}$.
2. $\mathcal{S}_X = \mathcal{S}_Y$*, that is* $\forall \bar{t} \in \mathcal{S}$*, we have that* $\mathbb{P}_{X_{\bar{t}}} = \mathbb{Q}_{Y_{\bar{t}}}$.

That is, the law of a S.P. is uniquely defined by the finite dimensional distribution, and $vice - versa$.

Proof. The proof follows from the previous lemma.

- 2) $\implies$ 1).
  - Let us define
  $$\mathcal{A} := \{\pi_t^{-1}(A) | \;\; t \in T \;\; and \;\; A \in \mathscr{E}\},$$
  $$\mathcal{B} := \bigcup_{n=1}^{+\infty} \{\cap_{i=1}^{n} A_1 | \;\; A_i \in \mathcal{A}\}.$$
  We remember that $\mathscr{E}^{\otimes T} := \sigma(\mathcal{A}) = \sigma(\mathcal{B})$, and $\mathcal{B}$ is a $\pi - system$ for $\mathscr{E}^{\otimes T}$.
  - So, since we have the probability spaces $(E^T, \mathscr{E}^{\otimes T}, \mathbb{P}_{\Phi_X})$ and $(E^T, \mathscr{E}^{\otimes T}, \mathbb{Q}_{\Phi_Y})$, we have that $\mathbb{P}_{\Phi_X} = \mathbb{Q}_{\Phi_Y}$ if they coincide on a $\pi - system$ for Corollary (2.5).
  - Now, let's take $A \in \mathcal{B}$. So by definition, $A = \cap_{i=1}^{n} \pi_{t_i}^{-1}(A_i)$, with $t_i \in T$ and $A_i \in \mathscr{E}$, and we have

$$\mathbb{P}_{\Phi_X}(A) = \mathbb{P}\left(\{\Phi_X \in A\}\right) \underbrace{=}_{(7.2)} \mathbb{P}\left(\{X_{\bar{t}} \in \bigtimes_{i=1}^{n} A_i\}\right) = \tag{11}$$
$$\mathbb{P}_{X_{\bar{t}}}\left(\bigtimes_{i=1}^{n} A_i\right) \underbrace{=}_{hypothesis} \mathbb{Q}_{Y_{\bar{t}}}\left(\bigtimes_{i=1}^{n} A_i\right) = \mathbb{Q}\left(\{Y_{\bar{t}} \in \bigtimes_{i=1}^{n} A_i\}\right) =$$
$$\mathbb{Q}(\{\Phi_Y \in A\}) = \mathbb{Q}_{\Phi_Y}(A),$$

  so $\mathbb{P}_{\Phi_X} = \mathbb{Q}_{\Phi_Y}$, and this is the thesis.

- 1) $\implies$ 2).
  - Let us fix $\bar{t} = (t_1, .., t_n) \in \mathcal{S}$. We have the probabilistic space

$$(E^n, \mathscr{E}^{\otimes n}, \mathbb{P}_{X_{\bar{t}}}),$$
$$(E^n, \mathscr{E}^{\otimes n}, \mathbb{Q}_{Y_{\bar{t}}})$$

    and we want to prove that $\mathbb{P}_{X_{\bar{t}}} = \mathbb{Q}_{Y_{\bar{t}}}$.
  - We remember that

$$\mathcal{C} := \{\times_{i=1}^{n} A_i |\ A_i \in \mathscr{E}\}$$

    is a $\pi-system$ for $\mathscr{E}^{\otimes n}$. This is true because we have a finite product of measurable spaces.
  - So, given $B = \times_{i=1}^{n} A_i \in \mathcal{C}$, we can do as above in (11), and if we set

$$A = \cap_{i=1}^{n} \pi_{t_i}^{-1}(A_i),$$

    ($\pi_{t_i} : E^T \to E$ is the projection) we obtain

$$\mathbb{P}_{X_{\bar{t}}}(B) = \mathbb{P}_{\Phi_X}(A) \underbrace{=}_{hypothesis} \mathbb{Q}_{\Phi_Y}(A) = \mathbb{Q}_{Y_{\bar{t}}}(B).$$

□

7.2 Kolmogorv's Theorems

- Let T be a not empty set.
- Let us denote as $\mathcal{S} := \{(t_1, .., t_n)|\ t_i \in T\ and\ t_i \neq t_j \iff i \neq j\}$ the set of tuples with every element different.
- Let us have $(E, \mathscr{E})$ a measurable space.
- Let us consider $(E^T,\ \mathscr{E}^{\bigotimes T})$ the product space.
- Let us have

$$\{\mu_\tau |\ \tau \in \mathcal{S}\}.$$

  a family of *probability*. We intend that, if $\tau = (t_1, ..., t_n)$, then μ_τ is a probability on $(E^n, \mathscr{E}^{\bigotimes_n})$.
- **QUESTION:** can we find
  - $(\Omega, \mathscr{F}, \mathbb{P})$ a probabilistic space,
  - $X : T \times \Omega \to E$ a *S.P.*

  such that

$$\forall \tau \in \mathcal{S},\ \ \mathbb{P}_{X_\tau} = \mu_\tau,$$

  that is, X has how finite dimensional distribution the family $\mu_\tau | \tau \in \mathcal{S}$.

8 Filtration

Setting 1. Let us suppose to be in the following setting.

* Let $(\Omega,\mathscr{F},\mathbb{P})$ be a probabilistic space and $(E,\mathscr{E})$ a measurable space.
* Let $X : T \times \Omega \to E$ be a *S.P.*
* Let us suppose in this section that T is an interval (unlimited or not) of $\mathbb{R}$ or a subset of $\mathbb{N}$. To fix our ideas, we can suppose that $T = [0, +\infty)$.

Definition 33 (Filtration)**.** Let us have $(\mathscr{F}_t)_{t\in T}$ a family of $\sigma - algebras$ of set of Ω. We say that this family is a filtration if

- for all $t \in T$, we have that $\mathscr{F}_t \subset \mathscr{F}$,
- for all $s < t$ that are elements of T, we have that $\mathscr{F}_s \subset \mathscr{F}_t$.

Definition 34 ($\mathscr{F}_\infty$)**.** Given a filtration as the one above, we define

$$\mathscr{F}_\infty := \bigvee_{t\in T} \mathscr{F}_t, \tag{12}$$

that is the smallest $\sigma - algebra$ that contains every $\mathscr{F}_t$.

Definition 35 (Adapted)**.** Given a filtration as the one above, we say that X is adapted if for all $t \in T$, we have that X_t is $\mathscr{F}_t - measurable$.

Definition 36 (Progeressively Measurable)**.** The *S.P.* X is *Progressively Measurable* if for every $t \geq 0$, we have that

$$X|_{[0,t]\times\Omega} : ([0,t] \times \Omega,\ \ \mathfrak{B}([0,t]) \otimes \mathscr{F}_t) \to (E,\mathscr{E})$$

is $\mathfrak{B}([0,t]) \otimes \mathscr{F}_t - measurable$.

Definition 37 (Filtration right-C^0)**.** A filtration $(\mathscr{F}_t)_{t\in T}$ is right-continuous if for all $t \in T$, we have that

$$\mathscr{F}_t = \bigcap_{\epsilon>0} \mathscr{F}_{t+\epsilon}.$$

Now, let $\tau : \Omega \to [0, +\infty]$ be a r.r.v.

Definition 38 (Stopping Time)**.** τ is a stopping time if for all $t \geq 0$, we have $\{\tau \leq t\} \in \mathscr{F}_t$.

In the future, we denote a stopping time with *S.T.*

Definition 39 (σ-algebra associated to a S.T.)**.** If the function τ above is a stopping time, we define

$$\mathscr{F}_\tau := \{A \in \mathscr{F}_\infty \mid \forall t \in T,\ \ A \cap \{\tau \leq t\} \in \mathscr{F}_t\} \tag{13}$$

as the $\sigma - algebra$ associated to the stopping time.

Definition 40.

8.1 Null Sets

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probabilistic space.

Definition 41 (Negligible sets)**.** Let us have $N \subseteq \Omega$. We say that N is negligible if

$$\inf\{P(B) : \ B \in \mathscr{F}, \ N \subseteq B\} = 0.$$

that is there exists $C \in \mathscr{F}$ such that $N \subseteq C$ and $\mathbb{P}(C) = 0$.

Definition 42 (Set of Negligible Sets)**.** We define

$$\mathcal{N} := \{N \subseteq \Omega : \ N \ \textit{is negligible}\}.$$

If we want to emphasis the $\sigma - algebra$ and the probability, we can call $\mathcal{N}$ as $\mathcal{N}_{(\mathscr{F},\mathbb{P})}$.

Definition 43 (Complete $\sigma - algebra$)**.** We say that $\mathscr{F}$ is $\mathbb{P}$-complete if $\mathcal{N}_{(\mathscr{F},\mathbb{P})} \subseteq \mathscr{F}$.

Now, let us have $(\mathscr{F}_t)_{t\geq 0}$ a filtration with respect to $\mathscr{F}$.

Definition 44. We say that $(\mathscr{F}_t)_{t\geq 0}$ is complete if for all $t \geq 0$,

$$\mathcal{N}_{(\mathscr{F},\mathbb{P})} \subseteq \mathscr{F}_t.$$

Remark 23. In definition above, we just need that $\mathcal{N}_{(\mathscr{F},\mathbb{P})} \subseteq \mathscr{F}_0$.
Remark 24. It is important to note that $\mathscr{F}_t$ have to contain $\mathcal{N}_{(\mathscr{F},\mathbb{P})}$ and not $\mathcal{N}_{(\mathscr{F}_t,\mathbb{P}|_{\mathscr{F}_t})}$.

8.1.1 Property of the Null Sets

Let us have $(\omega, \mathscr{F}, \mathbb{P})$ a probabilistic space. Let $\mathcal{N}$ be the set of negligible sets.

Proposition 8.1. *The following statements hold true.*

1. $N \in \mathcal{N}$ *and* $A \in \mathscr{F} \implies N \cap A \in \mathcal{N}$.
2. $N \in \mathcal{N}$ *and* $M \in \mathcal{N} \implies N \cup M \in \mathcal{N}$ *and* $N \cap M \in \mathcal{N}$.

Proof. Immediate. □

8.1.2 Completion of a Sigma-Algebra

Given our probabilistic space $(\Omega, \mathscr{F}, \mathbb{P})$, if the sigma algebra $\mathscr{F}$ is just complete we do nothing. Otherwise, we want to built a complete filtration that is complete with respect to $\mathbb{P}$ (we just say that we want a $\mathbb{P} - complete\ filtration$).

We proceed in this way.

- We define $\mathscr{F}^P := \sigma(\mathscr{F} \cup \mathcal{N})$, with $\mathcal{N} := \mathcal{N}_{(\mathscr{F},\mathbb{P})}$.
- We prove the following theorem.

  Theorem 8.2. *Let us have* $A \subseteq \Omega$. *We have that*

  $$A \in \mathscr{F}^P \iff \exists B, C \in \mathscr{F} : \ B \subseteq A \subseteq C \ \textit{and} \ \mathbb{P}(B) = \mathbb{P}(C).$$

- We can extend $\mathbb{P}$ at $\mathbb{P}^* : \mathscr{F}^P \to [0,1]$ such that if we have A and B and C as above, then $\mathbb{P}^*(A) := \mathbb{P}(B) = \mathbb{P}(C)$.
- $\mathbb{P}^*$ is well defined (it is independent of B and C), and it is a probability on $\mathscr{F}^P$.
- Now we prove the following theorem

  Theorem 8.3. *Given $\mathscr{F}$, and $\mathscr{F}^P$, and $\mathbb{P}$ and $\mathbb{P}^*$ as above, we have that*

  $$\mathcal{N}_{(\mathscr{F},\mathbb{P})} = \mathcal{N}_{(\mathscr{F}^P,\mathbb{P}^*)}.$$

- We observe in the end that given $A \in \mathscr{F}^P$ such that $\mathbb{P}^*(A) = 0$, then for all $B \subseteq A$, we have that $B \in \mathscr{F}^P$, and our sigma algebra $\mathscr{F}^P$ is complete.
-

  Remark 25. We have that $\mathscr{F} \cup \mathcal{N}$ is a $pi - system$ for $\mathscr{F}^P$. This is immediate because

  - by definition, it generate. In fact $\sigma(\mathscr{F} \cup \mathcal{N}) = \mathscr{F}^P$.
  - It is closed by intersection. In fact for all $A_1, A_2 \in \mathscr{F}$ and $N_1, N_2 \in \mathcal{N}$, we have

  $$(A_1 \cup N_1) \cap (A_2 \cup N_2) = \underbrace{(A_1 \cap A_2)}_{\in \mathscr{F}} \cup \underbrace{(A_1 \cap N_2) \cup (N_1 \cap A_2) \cup (N_1 \cap N_2)}_{\mathcal{N}}.$$

  and this conclude.

8.2 Complete and Right Continuous Filtration

- Let us have $(\Omega, \mathscr{F}, \mathbb{P})$ a probabilistic space.
- Given a filtration $(\mathscr{F}_t)_{t\geq 0}$ that is indexed in $[0, +\infty)$, we want to build a $right - C^0$ $filtration$.
- Let us have $(\mathscr{G}_t)_{t\geq 0}$ a filtration with respect to $\mathscr{F}$.

Proposition 8.4. *Let us define for all $t \geq 0$,*

$$\mathscr{F}_t := \bigcap_{\epsilon>0} \mathscr{G}_{t+\epsilon}.$$

then

1. *$(\mathscr{F}_t)_{t\geq 0}$ is a filtration with respect to $\mathscr{F}$.*
2. *$(\mathscr{F}_t)_{t\geq}$ is $right - C^0$.*

Proof. The proof is simple.

- 1.

  Let us have $0 \leq t < t+k$, with $k > 0$ a real number. Then

$$\mathscr{F}_t := \bigcap_{\epsilon>0} \mathscr{G}_{t+\epsilon} = \bigcap_{0<\epsilon\leq k} \mathscr{G}_{t+\epsilon} \cap \bigcap_{\epsilon>k} \mathscr{G}_{t+\epsilon} \subseteq \bigcap_{\epsilon-k>0} \mathscr{G}_{t+\epsilon-k+k} \underbrace{=}_{\gamma=\epsilon-k} \bigcap_{\gamma>0} \mathscr{G}_{t+k+\gamma} = \mathscr{F}_{t+k}.$$

- 2.

  We just need to observe that, for all $t \geq 0$,

$$\bigcap_{\epsilon>0} \mathscr{F}_{t+\epsilon} = \bigcap_{\epsilon>0} \left(\bigcap_{\gamma>0} \mathscr{G}_{(t+\epsilon)+\gamma} \right) = \bigcap_{\epsilon>0,\gamma>0} \mathscr{G}_{t+\epsilon+\gamma} = \bigcap_{\delta>0} \mathscr{G}_{\delta} = \mathscr{F}_t.$$

□

8.2.1 We make a right-continuous and complete filtration

- Now, given a filtration $(\mathscr{G}_t)_{t\geq 0}$, we can have a *complete*, $righ - C^0$ filtration.
- Indeed, let us have $(\Omega, \mathscr{F}, \mathbb{P})$ a probabilistic space.
- Let us have $(\mathscr{G}_t)_{t\geq 0}$ a filtration with respect to $\mathscr{F}$.
- We define

$$\tilde{\mathscr{F}}_t := \sigma(\mathscr{G}_t \cup \mathcal{N}_{\mathscr{F}})$$

- Now $(\tilde{\mathscr{F}}_t)_{t\geq 0}$ is a complete filtration.
- Let us define now

$$\mathscr{F}_t := \bigcap_{\epsilon>0} \tilde{\mathscr{F}}_{t+\epsilon}$$

- Now we have that $(\mathscr{F}_t)_{t\geq 0}$ is $right - C^0$. It is immediate that it is still complete.

8.2.2 Filtration associated to a Process

- Let us set $T = [0, +\infty)$ (but with slightly changing another interval is ok, even a discrete set).
- Let us have $(\Omega, \mathscr{F}, \mathbb{P})$ a probability space, and let us have $(E, \mathscr{E})$ a measurable space.
- Let $X : T \times \Omega \to E$ be a $S.P.$

Definition 45 (Filtration associated to a $S.P.$)**.** We define the following filtrations,

1. We define as
$$\tilde{\mathscr{F}}_t^X := \sigma\left(\{X_s : s \in T,\ \ and\ \ s \leq t\}\right)$$
and we define $(\tilde{\mathscr{F}}_t^X)_{t\geq 0}$ the *filtration generated* by X.
2. We define as
$$\overline{\mathscr{F}}_t^X := \sigma\left(\tilde{\mathscr{F}}_t^X \cup \mathcal{N}\right),$$
and we define $(\overline{\mathscr{F}}_t^X)_{t\geq 0}$ as the completion of the above filtration.
3. We define as
$$\mathscr{F}_t^X := \bigcap_{\epsilon>0} \overline{\mathscr{F}}_{t+\epsilon}^X,$$
and we define $(\mathscr{F}_t^X)_{t\geq 0}$ the right continuous filtration that we obtain from the one above.

So, in the end, we have that $(\mathscr{F}_t)_{t\geq 0}$ is the filtration complete and right continuous associated to X.

Definition 46. Given the

We want to prove a theorem that we use in the section of *Martingales*. Let $(\Omega, \mathscr{F})$ and $(E, \mathscr{E})$ be two measurable spaces (domain and codomain of our *S.P.*), let $(\mathscr{F}_n)_{n\in\mathbb{N}}$ be a filtration.

Lemma 8.5. *Let $X = (X_n)_{n\in\mathbb{N}}$ be a S.P. adapted to the filtration above. Let $\tau : \Omega \to \mathbb{N}^{\geq 0}$ be a stopping time and we suppose that $\tau < +\infty$. Then X_τ is $\mathscr{F}_\tau$-measurable.*

Remark 26. $(X_\tau)(\omega) := X_{\tau(\omega)}(\omega)$.

Proof. Let us have $A \in \mathscr{E}$. We need to prove that $\{X_\tau \in A\} \in \mathscr{F}_\tau$. We observe that X_τ is well defined because $\tau < +\infty$. Let's begin.

- $\{X_\tau \in A\} \in \mathscr{F}_\infty$.
$$\{X_\tau \in A\} = \bigcup_{k=0}^{+\infty} \underbrace{\{X_k \in A\}}_{\in \mathscr{F}_k \subset \mathscr{F}_\infty} \cap \underbrace{\{\tau = k\}}_{\in \mathscr{F}_k \subset \mathscr{F}_\infty},$$
so the LHS belong to $\mathscr{F}_\infty$ because it is obtained by countable intersection and union.
- for all $n \geq 0$, $\{X_\tau \in A\} \cap \{\tau \leq n\} \in \mathscr{F}_n$.
$$\{X_\tau \in A\} \cap \{\tau \leq n\} \in \mathscr{F}_n = \bigcup_{k=0}^{n} \underbrace{\{X_k \in A\}}_{\in \mathscr{F}_k \subset \mathscr{F}_n} \cap \underbrace{\{\tau = k\}}_{\in \mathscr{F}_k \subset \mathscr{F}_n},$$
so the LHS belong to $\mathscr{F}_n$ likewise before.

□

9 Martingales

Let us have (Ω,$\mathscr{F}$,$\mathbb{P}$) a probabilistic space, let us have $(E, \mathscr{E}) = (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$, and let us have $(\mathscr{F}_t)_{t\geq 0}$ a filtration with respect to $(\Omega, \mathscr{F})$. Let $M := (M_t)_{t\geq 0}$ a $S.P.$

Definition 47 (Martingale)**.** We say that M is a *Martingale* (with respect to filtration $(\mathscr{F}_t)_{t\geq 0}$) if the following property hold true,

- M is adapted with respect to $(\mathscr{F}_t)_{t\geq 0}$,
- for all $t \geq 0$, $M_t \in L_1(\Omega, \mathscr{F}, \mathbb{P})$.
- for all $0 \leq s < t$ that are in T, we have that $M_s = E[M_t|\mathscr{F}_s]$.

if we have ($\leq$) in the third condition above, we have a $sub - martingale$, if we have ($\geq$) we have a $super - martingale$.

Remark 27. Let us have $\varphi(t) := \mathbb{E}[M_t]$, for all $t \geq 0$. We observe that

- M marti. Then we have for all $0 \leq s < t$ that

  $$\varphi(s) = \mathbb{E}[M_s] = \mathbb{E}[\ \mathbb{E}[M_t|\mathscr{F}_s]\ ] = \mathbb{E}[M_t] = \varphi(t),$$

  that is φ is a constant function.
- M sub-marti. Then φ is an increasing function.
- M super-marti. Then φ is a decreasing function.

In discrete time, that is we have $(M_n)_{n\in\mathbb{N}}$, third condition can be replaced by

$$\forall n \in \mathbb{N},\ \ \mathbb{E}[M_{n+1}|\mathscr{F}_n] = M_n \tag{14}$$

in fact, for example

$$\mathbb{E}[M_{n+2}|\mathscr{F}_n] \underbrace{=}_{tower} \mathbb{E}[\ \mathbb{E}[M_{n+2}|\mathscr{F}_{n+1}]\ \ |\mathscr{F}_n] \underbrace{=}_{(14)} \mathbb{E}[M_{n+1}|\mathscr{F}_n] \underbrace{=}_{(14)} M_n$$

Proposition 9.1 (Martingle and Convex Function)**.** *We have the following statements.*

1. *if*
   - $(M_t)_{t\geq 0}$ *is a martingle,*
   - $\varphi : \mathbb{R} \to \mathbb{R}$ *a convex function such that for all* $t \geq 0$, *we have* $\mathbb{E}[|\varphi(M_t)|] < +\infty$,

   then $(\varphi(M_t))_{t\geq 0}$ *is a sub-martingle.*
2. *If*
   - $(M_t)_{t\geq 0}$ *is a sub- martingle,*

- *$\varphi : \mathbb{R} \to \mathbb{R}$ a convex, increasing function such that for all $t \geq 0$, we have $\mathbb{E}[|\varphi(M_t)|] < +\infty$,*

then $(\varphi(M_t))_{t\geq 0}$ is a sub-martingle.

Remark 28. We observe that, if

- $(M_t)_{t\geq 0}$ is a sub-martingle

this does not imply that $|M_t|$ and M_t^2 is a sub-martingle. But, if

- $(M_t)_{t\geq 0}$ is a sub-martingle and for every $t \geq 0$, we have that $M_t \geq 0$ ($a.s.$?????),

then M_t^2 is a sub-martingle.

Lemma 9.2 (Stopped Process)**.** *Let us suppose that*

- *$(M_n)_{n\in\mathbb{N}}$ is a marti (sub,super),*
- *τ is a stopping time.*
  Then $(M_{n\wedge\tau})_{n\in\mathbb{N}}$ is a marti (sub,super).

Proof. The proof is a simple check given the following

- $$M_{n\wedge\tau} = M_0 + \sum_{k=1}^{n} I_{\{\tau\geq k\}}(M_k - M_{k-1})$$
- $\{\tau \geq k\} = \{\tau \leq k-1\}^c \in \mathscr{F}_{k-1}$

□

Theorem 9.3 (Optional Stopping Theorem)**.** *Let M =$(M_n)_{n\in\mathbb{N}}$be a process and τ be a stopping time that can have one of the following properties*

a) τ is bounded by an integer constant $N \geq 1$,

b) τ is finite, and M is bounded.

Then we can state

1. *M marti, a) or b) hold. Then M_τ is integrable, and $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$.*
2. *M super-marti, a) or b) hold. Then $\mathbb{E}[M_\tau] \leq \mathbb{E}[M_0]$.*
3. *M super-marti, $M \geq 0$, $\tau < +\infty$ a.s. Then $\mathbb{E}[M_\tau] \leq \mathbb{E}[M_0]$.*
4. *M sub-marti, a) holds. Then $\mathbb{E}[M_\tau] \leq \mathbb{E}[M_N]$.*

Proof. Let us prove the above statements following the order.

1. Let us assume a).
   - $|M_\tau| \le \sum_{k=0}^N |M_k| \implies M_\tau \in L^1$ because for all k, $M_k \in L^1$.
   - $M_\tau \underbrace{=}_{\tau\le N} M_{\tau\wedge N} \implies \mathbb{E}[M_\tau] = \mathbb{E}[M_{\tau\wedge N}] \underbrace{=}_{(*)} \mathbb{E}[M_{\tau\wedge 0}] = \mathbb{E}[M_0]$. In $(*)$ we have used that $(M_{\tau\wedge n})_n$ is a martingale because of *Lemma* (9.2) and *Remark* (27).

   Let us assume b).
   - M_τ is integrable because M is bounded.
   - $\forall n \in \mathbb{N}$, $\mathbb{E}[M_{\tau\wedge n}] \equiv \mathbb{E}[M_0]$ as above.
   - $\lim_{n\to+\infty} M_{\tau\wedge n} = M_\tau$ a.s. because $0 \le \tau < +\infty$ a.s.
   - $\mathbb{E}[M_0] = \lim_n \mathbb{E}[\ M_{\tau\wedge n}\ ] = \mathbb{E}[\ \lim_n M_{\tau\wedge n}\ ] = \mathbb{E}[M_\tau]$.
2. The proof is very similar to the one above.
3. We just need to use *Fatou*, that is (3.4).
4. By our hypothesis, we have $M_n \le \mathbb{E}[M_N|\mathscr{F}_n]$ if $n \le N$, and $\tau \le N$ a.s. So we have
   - $$\mathbb{E}[M_\tau] = \sum_{n=0}^N \mathbb{E}[I_{\{\tau=n\}} M_\tau] = \sum_{n=0}^N \mathbb{E}[I_{\{\tau=n\}} M_n] \underbrace{\le}_{(A)}$$
   - Now we observe that
     $I_{\{\tau=n\}} M_n \le \mathbb{E}[I_{\{\tau=n\}} M_N|\mathscr{F}_n] \implies \mathbb{E}[I_{\{\tau=n\}} M_n] \le \mathbb{E}[I_{\{\tau=n\}} M_N]$.
   - Now using the inequality above in the sum we obtain
     $$\underbrace{\le}_{(A)} \sum_{n=0}^N \mathbb{E}[I_{\{\tau=n\}} M_N] = \mathbb{E}[M_N].$$

□

Corollary 9.4. *Let $M = (M_n)_{n\ge 0}$ be a S.P., let τ_1 and τ_2 be two stopping times. Then we have the following statement.*

1. *M is a marti, $\tau_1 \le \tau_2$, condition a) or b) in (9.3) hold true for τ_2.*
   Then $\mathbb{E}[M_{\tau_2}|\mathscr{F}_{\tau_1}] = M_{\tau_1}$.
2. *M is a sub-marti, τ_1 is bounded by an integer constant N. Then $\mathbb{E}[M_N|\mathscr{F}_{\tau_1}] \ge M_{\tau_1}$.*

Proof. We prove the above statement following the order.

1. - We have that $\tau_1 < +\infty$ in either cases, $a)$ or $b)$, so by (8.5), we have that M_{τ_1} is $\mathscr{F}_{\tau_1} - measurable$.
   - If we prove that for all $A \in \mathscr{F}_{\tau_1}$, we have that $\mathbb{E}[M_{\tau_2} I_A] = \mathbb{E}[M_{\tau_1} I_A]$, then we can conclude by uniqueness of the conditional mean.
   - Let us set

$$\tau_3 = I_A \cdot \tau_1 + I_{A^C} \cdot \tau_2.$$

   Clearly, if $a)$ or $b)$ holds for τ_2, then it holds for τ_3. Moreover, it is easy to show that τ_3 is a stopping time.

   - So, M marti and τ_2 and τ_3 are S.T and a) or b) holds for both $\underbrace{\implies}_{(9.3)} \mathbb{E}[M_{\tau_3}] =$ $\mathbb{E}[M_0] = \mathbb{E}[M_{\tau_2}]$.
   - Now we observe that

$$M_{\tau_3} = I_A \cdot M_{\tau_1} + I_{A^C} \cdot M_{\tau_2}$$

   - So if we take expectation both parts we obtain

$$\mathbb{E}[M_{\tau_3}] = \mathbb{E}[I_A \cdot M_{\tau_1} + I_{A^C} \cdot M_{\tau_2}] = \mathbb{E}[M_0] = \mathbb{E}[M_{\tau_2}] =$$
$$= \mathbb{E}[I_A \cdot M_{\tau_2} + I_{A^C} \cdot M_{\tau_2}] \implies \mathbb{E}[M_{\tau_1} I_A] = \mathbb{E}[M_{\tau_2} I_A] \implies Thesis.$$

2. - We just need to prove that for all $A \in \mathscr{F}_{\tau_1}$, we have $\mathbb{E}[M_N I_A] \geq \mathbb{E}[M_{\tau_1} I_A]$.
   - Likewise before, by (9.3) we prove that $\mathbb{E}[M_{\tau_3}] \leq \mathbb{E}[M_N]$.
   - If we set $\tau_2 \equiv N$, we have the following identity

$$M_{\tau_3} = I_A \cdot M_{\tau_1} + I_{A^C} \cdot M_N$$

   and from here the thesis is straightforward.

□

Now we prove a very important inequalities that hold for (discrete) martingles.

Theorem 9.5 (Doob's Maximal Inequality)**.** *Let us have one of the following,*

1. $(M_n)_n$ *a martingle.*
2. $(M_n)_n$ *a positive* $sub - martingle$.

Then for every $N \geq 1$ *integer and* $\lambda > 0$*, we have*

$$\lambda \mathbb{P}[\max_{1 \leq n \leq N} |M_n| \geq \lambda] \leq \mathbb{E}[|M_N| I_{\{\max_{1 \leq n \leq N} |M_n| \geq \lambda\}}] \leq \mathbb{E}[|M_N|].$$

Proof. We use the stopping theorem.

1. Let us have M a (discrete) martingle.

- Let us set $M^* := \underset{1\le n\le N}{max}|M_n|$.
- Let us set $M^N_n := M_{n\wedge N}$. We have that this is a martingle ($\tau \equiv N$ is $S.T.$ plus Lemma (9.2)), and $M* = \underset{1\le n\le N}{max}|M^N_n|$.
- Let us set $A(\omega) := \{n : n \le N \;\; and \;\; |M^N_n(\omega)| \ge \lambda\}$. Let us define

$$\tau(\omega)\begin{cases} min(A(\omega)) \;\; if \;\; A \neq \emptyset, \\ N+1 \;\; if \;\; A = \emptyset. \end{cases}$$

  It is easy to show that τ is a stopping time.
- It is immediate that $\{M^* \ge \lambda\} = \{\tau \le N\} \in \mathscr{F}_\tau$.
- Now, we have the following implication,

$$(M^N_n)_n \;\; marti \implies (|M^N_n|)_n \; sub-marti$$

  thanks to *Lemma* (9.1) and the fact that $|\cdot|$ is a convex function.
- Now, it is always true that

$$\lambda I_{\{M^*\ge\lambda\}} \le |M^N_\tau| I_{\{M^*\ge\lambda\}}$$

- Now, $|M^N|$ sub-marti and $\tau \le N+1$ implies thanks to Lemma (9.4) that

$$|M^N_\tau| \le \mathbb{E}[\underbrace{|M^N_{N+1}|}_{=|M^N_N|=|M_N|} \;\; |\mathscr{F}_\tau] \implies I_{\{M^*\ge\lambda\}}|M^N_\tau| \le \mathbb{E}[I_{\{M^*\ge\lambda\}}|M_N||\mathscr{F}_\tau]$$

  since what we have said about $\{M^* \ge \lambda\}$ some point above.
- Now if we put together the inequalities that we have obtained and we take the expectation both parts we obtain

$$\lambda\mathbb{P}(M^* \ge \lambda) \le \mathbb{E}[|M_N| I_{\{M^*\ge\lambda\}}] \le \mathbb{E}[|M_N|].$$

2. Now let us have a (*discrete*) *positive* $sub-martingle$. The prove is (almost) exactly the same, because we can put the modulus function on M^N_N for every n.

□

Now we give an easy lemma.

Lemma 9.6. *Let X be a non-negative r.r.v. Then*

$$\mathbb{E}[X^p] = \int_0^{+\infty} pu^{p-1}\mathbb{P}[X \ge u]du.$$

Now we need a trivial estimate, but that is fundamental to understand the inequalities in the next section.

Proposition 9.7 (Trivial Estimate)**.** *Let us have $M = (M_n)_{n\ge0}$ a S.P. that can be*

1. a martingale,

2. a positive sub-martingale.

Let us set

$$M_n^* := \sup_{0 \le m \le n} |M_m|.$$

Then for every $p \in (1, +\infty)$ and for every $n_0 \in \mathbb{N}$, we have that

$$\mathbb{E}[(M_{n_0}^*)^p] \le (n+1)\mathbb{E}[|M_{n_0}|^p],$$

that is $M_{n_0}^* \in L^p \iff M_{n_0} \in L^p$.

Proof. We prove the proposition for martingles, then with slightly changes we can prove it even for sub-martingales.

- Let us suppose that M is a martingale, and let us fix $n_0 \in \mathbb{N}$.
- If $\mathbb{E}[|M_{n_0}|^p] = +\infty$ there is nothing to prove, so let us suppose that it is $< +\infty$.
- let us set $\varphi(x) := |x|^p$, for every $x \in \mathbb{R}$. We have that φ is convex because it is C^1 in $\mathbb{R}$, and it is increasing.
- Thanks to Proposition (9.1), we have that $(\varphi(M_n))_n = (|M_n|^p)_n$ is a sub-martingale.
- Since $(|M_n|^p)_n$ is a $sub - martingale$, its mean-function is increasing, so for every $0 \le m \le n_0$ we have

$$\mathbb{E}[|M_m|^p] \le \mathbb{E}[|M_{n_0}|^p].$$

- Then we can conclude with the following chain of inequalities,

$$(M_{n_0}^*)^p \le \sum_{i=0}^{n_0} |M_i|^p \implies \mathbb{E}[(M_{n_0}^*)^p] \le \mathbb{E}\left[\sum_{i=0}^{n_0} |M_i|^p\right] \le (n_0+1)\mathbb{E}[|M_{n_0}|^p],$$

  and this is the thesis.

□

Now we prove the Doob's Inequality. We keep the notation that we used in $Theorem$(9.5).

Theorem 9.8. *Let us have $M = (M_n)_{n \in \mathbb{N}}$ that can be*

1. a martingle.

2. a positive sub-martingle.

Then, for every $p > 1$ and for every $N \ge 1$ integer, we have

$$\mathbb{E}[(M_N^*)^p] \le \left(\frac{p}{p-1}\right)^p \mathbb{E}[|M_N|^p].$$

Remark 29. We observe briefly that

$$(M_N^*)^p = (\underset{1\le n\le N}{max} |M_n|)^p = \underset{1\le n\le N}{max} (|M_n|^p).$$

Proof. We use the Trivial Estimate (9.7) and the trick (9.6).

- Firstly, if $\mathbb{E}[|M_N|^p] = +\infty$ there is nothing to prove.
- Let us suppose then that $\mathbb{E}[|M_N|^p] < +\infty$. Thanks to (9.7), we have that $\mathbb{E}[(M_N^*)^p] < +\infty$. We can suppose also that $\mathbb{E}[(M_N^*)^p] > 0$, otherwise the inequality is trivially true.
- Now, using the *trick* (9.7) we obtain

$$\begin{aligned}
\mathbb{E}[(M_N^*)^p] &= \int_0^{+\infty} pu^{p-1}\mathbb{P}(M_N^* \ge u)du \quad \le \\
&\le \int_0^{+\infty} pu^{p-2}\mathbb{E}[|M_N| I_{\{M_N^* \ge u\}}]du = \\
&= \int_\Omega |M_N| \left(\int_0^{+\infty} pu^{p-2} I_{\{u \le M_N^*\}} du \right) d\mathbb{P} = \\
&= \int_\Omega |M_N| \left(\int_0^{M_N^*} pu^{p-2} du \right) d\mathbb{P} = \\
&= \frac{p}{p-1} \int_\Omega |M_N| (M_N^*)^{p-1} d\mathbb{P} = \\
&= \frac{p}{p-1} \mathbb{E}[|M_N|(M_N^*)^{p-1}] \le \\
&\le \frac{p}{p-1} \mathbb{E}[|M_N|^p]^{\frac{1}{p}} \cdot \mathbb{E}[|(M_N^*)^{(p-1)\cdot\frac{p}{p-1}}|]^{\frac{p-1}{p}},
\end{aligned}$$

  where the last inequality follow from *Holder* ($1 = \frac{1}{p} + \frac{p-1}{p}$). Now if we divide for $\mathbb{E}[(M_N^*)^p]^{1-\frac{1}{p}}$, we obtain the thesis.

□

9.1 Result in Continuous Time

Notation 1. *We use the following notations,*

* $M_T^* = sup_{t\in[0,T]}|M_t|$.

* $M_k^{(n,T)} := M_{\frac{kT}{2^n}}$.

* $M^{(n,T),*} = \max_{k=0,..,2^n} |M_k^{(n,T)}|$.

Now, we describe the setting.

- Let $M = (M_t)_{t\ge 0}$ be a martingale.

- Let $T > 0$ fixed.
- We observe that $(M_k^{(n,T)})_{k\geq 0}$ is a discrete martingale, so we can use the theorem for discrete martingale, in particular Doob.
- We need the following properties

$$M_T^* = \lim_n (M^{(n,T),*}) \tag{15}$$

everywhere on Ω (or a.c. if the filtration is complete).

Theorem 9.9 (Maximal Inequality)**.** *Let us have*

1. *M a martingale,*
2. *M a positive sub-martingale.*

Let us suppose that condition (15) *holds.*
Then for every $T > 0$ and $\lambda > 0$, we have that

$$\mathbb{P}(M_T^* \geq \lambda) \leq \frac{1}{\lambda}\mathbb{E}[|M_T|].$$

Remark 30. We remember (briefly) the theorem in the discrete time.

$$M = (M_n)_{n\in\mathbb{N}} \ \ marti/positive \ \ sub-marti \implies$$
$$\forall N \geq 1 \ \ and \ \ \lambda > 0, \ \ \mathbb{P}(M_N^* \geq \lambda) \leq \mathbb{E}[|M_N|]\frac{1}{\lambda}.$$

Proof. The proof follows the following steps.

- Let us have $T > 0$ a positive real number.
- Let $\lambda > 0$ be a positive real number.
- Let $0 < \epsilon < \lambda$ be a real number. Let us set $\lambda_1 := \lambda - \epsilon$.
- $\forall n \geq 1$ integer, let us set $A_n := \{M^{(n,T),*} \geq \lambda_1\}$.
- If $N = 2^n$, thanks to Doob Inequality we have

$$\mathbb{P}(A_n) \leq \mathbb{E}[|\underbrace{M_{2^n}^{(n,T)}}_{M_T}|]\frac{1}{\lambda_1}.$$

- We observe that $A_n \subseteq A_{n+1}$ for every n, so we have that $\mathbb{P}(\cup_n A_n) = \lim_n \mathbb{P}(A_n)$.
- We claim that

$$\{M_T^* \geq \lambda_1 + \epsilon\} \subseteq \cup_n A_n.$$

in fact, let us take ω on the LHS. Then

$$M_T^*(\omega) \underbrace{=}_{(15)} \lim_n M^{(n,T),*}(\omega) \geq \lambda_1 + \epsilon \implies$$
$$\forall \gamma > 0, \exists n_0 : \; n \geq n_0 \implies M^{(n,T),*}(\omega) \geq \lambda_1 + \epsilon - \gamma.$$

So, if $\gamma = \epsilon$, we obtain that for such n_0 we have

$$M^{(n_0,T),*}(\omega) \geq \lambda_1,$$

so we have that $\{M_T^* \geq \lambda_1 + \epsilon\} \subseteq A_{n_0} \subseteq \cup_n A_n$, as we wanted.

- So, we have the following inequalities,

$$\mathbb{P}(\{M_T^* \geq \lambda_1 + \epsilon\}) \leq \mathbb{P}(\cup_n A_n) = \lim_n \mathbb{P}(A_n) \leq \mathbb{E}[|M_T|] \frac{1}{\lambda_1}.$$

- If we substitute, we obtain that for every $0 < \epsilon < \lambda$

$$\mathbb{P}(M_T^* \geq \underbrace{(\lambda - \epsilon) + \epsilon}_{\lambda}) \geq \mathbb{E}[|M_T|] \frac{1}{\lambda - \epsilon}.$$

So, by taking the limit $\epsilon \to 0^+$, we obtain the thesis.

□

Theorem 9.10 (Doob Maximal Inequality)**.** *Let M be*

1. *a martingale,*
2. *a positive sub-martingale.*

Let us suppose that

$$\exists p > 1 \;\; s.t. \;\; \forall t \geq 0, \;\; \mathbb{E}[|M_t|^p] < +\infty.$$

and that Condition 15 hold. Then for every $T > 0$, we have that $M_T^* \in L^p$, and the following inequality hold

$$\mathbb{E}[(M_T^*)^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|M_T|^p].$$

Proof. We write the proof just for the martingale, the case of a positive sub-martingale is analogous.

- Let us define

$$X_n := \left(M^{(n,T),*}\right)^p = \left(\max_{k=0,\dots,2^n} |M_{\frac{kT}{2^n}}|\right)^p = \max_{k=0,\dots,2^n} \left(|M_{\frac{kT}{2^n}}|^p\right).$$

- Thanks to Doob maximal inequality (9.8), since $(M_k^{(n,T)})_{k\geq 0}$ is a martingale (or positive sub-marti), we have that

$$\mathbb{E}[X_n] \leq q \cdot \mathbb{E}[|M_T|^p], with \ q = \left(\frac{p}{p-1}\right)^p$$

where we have use as parameter $N = 2^n$. We observe that $RHS < +\infty$ because of our hp.

- Now, we have that
  - $0 \leq X_n \leq X_{n+1}$ for every n (it is immediate to see),
  - $\lim_n X_n = \left(\lim_n M^{(n,T),*}\right)^p = \left(M_T^*\right)^p$ thanks to Condition 15.

  So, by monotone convergence we have that

$$\mathbb{E}[(M_T^*)^p] = \lim_n \mathbb{E}[X_n].$$

- So, from the point above we deduce that

$$\mathbb{E}[(M_T^*)^p] \leq q \cdot \mathbb{E}[|M_T|^p],$$

  and this is the thesis.

□

Theorem 9.11 (Stopping Theorem Continuous Time). *Let us have $M = (M_t)_{t\geq 0}$ a S.P.*

- *$\exists\ p > 1$ such that $\forall t \geq 0$ we have $\mathbb{E}[|M_t|^p] < +\infty$.*
- *M is a $right - C^0$ martingale.*
- *τ is a bounded Stopping Time.*

Then M_τ is integrable, and $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$.

Proof. The proof uses (9.10) and it is made by approximation of τ.

- We observe preliminary that M $right - C^0$ implies that (15) holds true, so we can apply Doob's theorem above in continuous time.
- Let us have $\tau \leq N - 1$ a.c. in Ω.

  Now we can deduce that $M_\tau \in \mathcal{L}^1$ because $M_\tau \in \mathcal{L}^p$. In fact,

$$\mathbb{E}[|M_\tau|^p] \leq \mathbb{E}[(\sup_{t\in[0,N-1]} |M_t|)^p] = \mathbb{E}[(M_{N-1}^*)^p] \leq Cost \cdot \mathbb{E}[|M_{N-1}|^p] < +\infty$$

  where we have used (9.10).

- Let us define

$$\tau_n(\omega) := \sum_{k=0}^{+\infty} I_{\{\frac{k}{2^n} < \tau \leq \frac{k+1}{2^n}\}} \frac{k+1}{2^n}$$

- It is immediate that τ_n is a stopping time for every n.
- We have moreover that $\tau_n \to \tau$ from the right because for every ω, we have

$$0 \leq \tau_n(\omega) - \tau(\omega) \leq \frac{k+1}{2^n} - \frac{k}{2^n} = \frac{1}{2^n}.$$

- In addition, we obtain also that

$$\tau_n(\omega) = \tau_n(\omega) - \tau(\omega) + \tau(\omega) \leq \frac{1}{2^n} + (N-1) \leq 1 + N - 1 = N.$$

- So, $(M_{\frac{k}{2^n}})_{k\geq 0}$ marti and τ_n limited stopping time $\overbrace{\Longrightarrow}^{(9.3)} \mathbb{E}[M_{\tau_n}] = \mathbb{E}[M_0]$, for every n.
- Then, we have the following estimate that hold for every n,

$$|M_{\tau_n}| \leq |M_N^*|.$$

  Since $M_N^* \in L^p \subseteq L^1$ thanks to (9.10), we have that the *r.v.* M_{τ_n} are dominated in L^1.
- Then, since M is right continuous, we have that $M_{\tau_n} \to M_\tau$ if $n \to +\infty$.
- We can the conclude by dominated convergence, because we have

$$\mathbb{E}[M_0] = \mathbb{E}[M_{\tau_n}] = \lim_n \mathbb{E}[M_{\tau_n}] = \mathbb{E}[\lim_n M_{\tau_n}] = \mathbb{E}[M_\tau],$$

  and this is the thesis.

□

9.2 Doob Decomposition

Now we state and prove a theorem of decomposition for discrete sub-martingale.

Let us begin with a definition.

Definition 48. Let $A = (A_n)_{n\geq 0}$ be a *S.P.* in discrete time. We say that A is predictable (wrt a filtration $(\mathscr{F}_n)_{n\geq 0}$) if

$$\forall n \in \mathbb{N},\ \ A_n\ is\ \ -\mathscr{F}_{n-1} measurable.$$

Remark 31. There is a definition of predictable for *S.P.* in continuous time but it is more complicated and we omit it.

Theorem 9.12 (Doob Decomposition)**.** *Let X be a sub-martingale in discrete time wrt a filtration $(\mathscr{F}_n)_{n\geq 0}$. Then there exist*

- $M = (M_n)_{n\geq 0}$ *a martingale, with* $M_0 = 0$.
- $A = (A_n)_{n\geq 0}$ *an increasing, predictable process (wrt* $(\mathscr{F}_n)_{n\geq 0}$*), with* $A_0 = 0$.

such that for every $n \in \mathbb{N}$, we have $X_n = X_0 + M_n + A_n$. Moreover, the decomposition is unique, that is if we can find others M' and A' with the properties above, then M and M‘ are indistinguishable, and the same hold for A and A‘.

Proof. The proof is simple, we just need to find a good decomposition.

- Let us define the following sequence of process.
  - $M_0 = 0$,
  - $M_1 = X_1 - \mathbb{E}[X_1|\mathscr{F}_0]$,
  - $M_{n+1} = M_n + X_{n+1} - \mathbb{E}[X_{n+1}|\mathscr{F}_n]$.

  Then, we are forced to define $A_n \coloneqq X_n - X_0 - M_n$. Now, let us check that M and A have the properties that we seek.

- M is a martingale. It is really simple.
  - Adaptness is trivial because we have sum of $\mathscr{F}_n$ measurable functions.
  - Integrability is trivial for we have sum of integrable functions.
  - Martingale property. We just need to write

  $$\mathbb{E}[M_{n+1} - M_n|\mathscr{F}_n] = \mathbb{E}[X_{n+1} - \mathbb{E}[X_{n+1}|\mathscr{F}_n]|\mathscr{F}_n] = 0.$$

  and this conclude.

- A have the properties.
  - *predictability* wrt $(\mathscr{F}_n)_{n\geq 0}$. We have

  $$A_n = X_n - M_n = M_{n-1} - \mathbb{E}[X_n|\mathscr{F}_{n-1}],$$

  so it is $\mathscr{F}_{n-1}$ measurable because it is sum of $\mathscr{F}_{n-1} - measurable$ function.
  - *increasingness*. For every n, we have

  $$A_{n+1} = X_{n+1} - M_{n+1} - X_0 = \mathbb{E}[X_{n+1}|\mathscr{F}_n] - M_n - X_0 \geq X_n - M_n - X_0 = A_n.$$

  where in the inequality we used the $sub - martingale$ property.

  So we have the thesis.

- We observe that the A is increasing in the set where the sub-martingale property holds with probability 1.

□

Now we can improve $Corollary$ 9.4 for discrete sub-martingale.

Corollary 9.13. *Let us have the following setting,*

- *Let $X = (X_n)_{n\geq 0}$ be a sub-martingale in discrete time (wrt a filtration $(\mathscr{F}_n)_{n\geq 0}$).*

- *Let τ_1 and τ_2 be bounded stopping time (wrt the same filtration for X) such that $\tau_1 \leq \tau_2$.*

  Then

$$\mathbb{E}[X_{\tau_2}|\mathscr{F}_{\tau_1}] \geq X_{\tau_1},$$

and in particular we obtain

$$\mathbb{E}[X_{\tau_2}] \geq \mathbb{E}[X_{\tau_1}].$$

Proof. We use Theorem (9.12) that we have just discover.

- Since X is a sub-martingale, thanks to Theorem (9.12) we have that $X = M + A$, with M a martingale and A a suitable increasing predictable process.

- We observe preliminary that $\tau_1 \leq \tau_2$ implies that $\mathscr{F}_{\tau_1} \subseteq \mathscr{F}_{\tau_2}$ (easy to show,) so since M_{τ_2} is $\mathscr{F}_{\tau_2} - measurable$, it makes sense to compute $\mathbb{E}[M_{\tau_2}|\mathscr{F}_{\tau_1}]$. Idem for A_{τ_2} and X_{τ_2}.

- So, we have that

$$\mathbb{E}[X_{\tau_2}|\mathscr{F}_{\tau_1}] = \mathbb{E}[M_{\tau_2}|\mathscr{F}_{\tau_1}] + \mathbb{E}[A_{\tau_2}|\mathscr{F}_{\tau_1}]$$

- Now, we have to estimate both the terms above. We have

  - M is a martingale,
  - $\tau_1 \leq \tau_2$ stopping times,
  - τ_2 is bounded.

  So we can apply Corollary (9.4), and we have that $\mathbb{E}[M_{\tau_2}|\mathscr{F}_{\tau_1}] = M_{\tau_1}$.

- Moreover, we have that $\mathcal{A}_{\tau_2} \geq \mathcal{A}_{\tau_1}$ because $\tau_2 \geq \tau_1$ *a.c.* Furthermore, we have that A_{τ_1} is $\mathscr{F}_{\tau_1} - measurable$.

- Then we can conclude that

$$\mathbb{E}[X_{\tau_2}|\mathscr{F}_{\tau_1}] \geq M_{\tau_1} + \underbrace{\mathbb{E}[A_{\tau_1}|\mathscr{F}_{\tau_1}]}_{A_{\tau_1}} = X_{\tau_1},$$

  and this is the thesis. If we take expectation both part we obtain even the inequality that we seek.

□

9.3 Convergence for Sub-Martingale

9.3.1 Criterion of Convergence

- Let us have $(x_n)_n$ a sequence of real number.
- Let us define
  - $\sigma_0 = \tau_0 = 0$.
  - For $i \geq 0$, let us define

$$\sigma_{i+1} \coloneqq \inf\{n > \tau_i : x_n \leq a\} \quad \tau_{i+1} \coloneqq \inf\{n > \sigma_{i+1} : x_n \geq b\}$$

Remark 32. $inf(\emptyset) = +\infty$.

- We say that there is an upcrossing in $[a, b]$ between σ_i and τ_i if $\tau_i < +\infty$.
- We define $\gamma_{a,b} \coloneqq Card\{i \geq 1 \ : \ \tau_i < +\infty\}$.

The following is a simple result of analisys.

Lemma 9.14. *The following fact are equivalent,*

1. $(x_n)_n$ *is convergent (to a limit* $l \in \mathbb{R} \cup \{\pm\infty\}$*).*
2. *For every* $a < b$ *real, we have that* $\gamma_{a,b} < +\infty$*.*
3. *For every* $a < b$ *rational, we have that* $\gamma_{a,b} < +\infty$

Remark 33. We observe that if $[a', b'] \subseteq [a, b]$, then $\gamma_{a,b} \leq \gamma_{a',b'}$.

Proof. Exercise. □

9.3.2 Doob upcrossing lemma

Now, let us turn to processes. Let $(M_n)_n$ be a sub martingale in discrete time.

- Let us define
  - $\sigma_0(\omega) = \tau_0(\omega) = 0$. for every $\omega \in \Omega$.
  - For $i \geq 0$, let us define

$$\sigma_{i+1}(\omega) \coloneqq \inf\{n > \tau_i(\omega) : M_n(\omega) \leq a\} \quad \tau_{i+1}(\omega) \coloneqq \inf\{n > \sigma_{i+1}(\omega) : M_n(\omega) \geq b\}$$

- The *r.r.v.* above are *all stopping times*. The proof is a simple induction that lays on the following lemma. Let us have $(\Omega, \mathscr{F}, \mathbb{P}, (\mathscr{F}_n)_{n \geq 0})$ a filtered probabilistic space and $(E, \mathfrak{B}(E))$ a metric space. Let us have also a *S.P.* (M_n) adapted wrt $(\mathscr{F}_n)_n$ and $E - valued$.

  Lemma 9.15. *Let us have* $\tau : \Omega \to [0, +\infty]$ *a stopping time. Then*

$$\sigma(\omega) \coloneqq \inf\{n > \tau(\omega) : M_n(\omega) \in B\}$$

  with $B \in \mathfrak{B}(E)$ *is a stopping time.*

Proof. We just need to observe that for $k \in \mathbb{N}$,

$$\{\sigma \leq k\} = \bigcup_{i=0}^{k-1} \left(\{\tau = i\} \bigcap \left(\bigcup_{h=i+1}^{k} \{M_h \in B\} \right) \right).$$

□

- Now, let us consider the random number of upcrossing.

$$\gamma_{a,b}(\omega) := Card\{i \geq 1 : \tau_i(\omega) < +\infty\}.$$

For what we have said in the subsubsection above, if $\gamma_a, b(\omega)$ is bounded for every $a \in \mathbb{Q}$ and $b \in \mathbb{Q}$ with $a < b$, then we have that the sequence $(M_n(\omega))_n$ converges to a limit finite or infinite.

-

Remark 34. We observe that

$$\sigma_0 \leq \tau_0 \leq \sigma_1 \leq \tau_1 \leq \sigma_2 \leq ...\sigma_i \leq \tau_i \leq \sigma_{i+1}...$$

and we can not have that σ_i and τ_i are bounded from a constant independent of ω for some $i \geq 1$. Indeed, if it was true we would have

$$\mathbb{E}[M_{\tau_i}] \leq \mathbb{E}[M_{\sigma_{i+1}}].$$

since Theorem (9.13) holds true ($(M_n)_n$ is a sub-martingale and $\tau_i \leq \sigma_{i+1}$ are bounded stopping times).

But on the same time, we would have

$$M_{\sigma_{i+1}} \leq a < b \leq M_{\tau_i},$$

so if we take expectation both parts we obtain a contradiction.

- Now, let us define

$$\gamma_{a,b}^N(\omega) := Card\{i \geq 1 \ : \ \tau_i(\omega) \leq N\},$$

that is the number of upcrossing before the instant N.

- Now, let us set $\varphi(x) := (x - a)_+ + a$. This is a convex increasing function. Moreover, we also have that

$$|\varphi(x)| \leq |a| + |x|,$$

so it is immediate that $(\varphi(M_n))_n$ is again a sub-martingale since Proposition (9.1) holds true.

- Now we ask ourselves which are the relation between the upcrossing starting and final times that we have if we consider the *r.v.* $(M_n)_n$ and $(\varphi(M_n))_n$. We discover that they are the same, because it is immediate to show that

$$M_n \leq a \iff \varphi(M_n) \leq a,$$
$$M_n \geq b \iff \varphi(M_n) \geq b.$$

Since they are the same, we do not distinguish between the upcrossing of M_n and $\varphi(M_n)$.

Now we are ready to state the *Doob′s upcrossing lemma.*

Lemma 9.16 (Doob Upcrossing Lemma)**.** *Let us have*

- $(M_n)_{n\geq 0}$ *a sub-martingale.*
- $a < b$ *two real numbers.*
- $N \geq 1$ *an integer.*

Then we have the estimate

$$(b-a)\mathbb{E}[\gamma_{a,b}^N] \leq \mathbb{E}[(M_N - a)_+]$$

Proof. The proof is divided by steps that simplify the dissertation.

1. First Step.
   - We omit the random element ω when it is not necessary.
   - We set for simplicity $\tilde{M}_n \coloneqq \varphi(M_n)$.
   - If $1 \leq i \leq \gamma_{a,b}^N$, (we suppose $\gamma_{a,b}^N(\omega)$ strictly positive) then by definition we have $\tilde{M}_{\tau_i} - \tilde{M}_{\sigma_i} \geq (b-a)$.
   - So, if we sum we obtain

   $$(A) \coloneqq \sum_{i=1}^{\gamma_{a,b}^N} (\underbrace{\tilde{M}_{\tau_i}}_{\tilde{M}_{\tau_i \wedge N}} - \underbrace{\tilde{M}_{\sigma_i}}_{\tilde{M}_{\sigma_i \wedge N}}) \geq \gamma_{a,b}^N (b-a).$$

   We remember that $i \leq \gamma_{a,b}^N$ implies that $\sigma_i \leq \tau_i \leq N$, and $c \wedge d = min\{c, d\}$. This is a bad estimate because the sum depends upon the random variable $\gamma_{a,b}^N$. We want to put something deterministic, like N.
   - Let us set for the sake of simplicity $k = \gamma_{a,b}^N$. Surely we have that $\tau_k \leq N$ and $\tau_{k+1} > N$ by definition, and we remember that

   $$\sigma_0 \leq \tau_0 \leq \sigma_1 \leq \tau_1 \leq ... \leq \sigma_k \leq \tau_k \leq N.$$

   - We ask ourselves what σ_{k+1} can do. We can have

- $\sigma_{k+1} > N$. Then for every $i > k$, we have $N < \sigma_i < \tau_i$. Then we have

$$(A) = (A) + \sum_{i=k+1}^{N} \underbrace{\tilde{M}_{\tau_i \wedge N} - \tilde{M}_{\sigma_i \wedge N}}_{=0} = \sum_{i=1}^{N} \tilde{M}_{\tau_i \wedge N} - \tilde{M}_{\sigma_i \wedge N}.$$

- $\sigma_{k+1} \leq N$. Then $\tau_k < \sigma_{k+1} \leq N < \tau_{k+1}$, and

$$\sum_{i=1}^{N} \tilde{M}_{\tau_i \wedge N} - \tilde{M}_{\sigma_i \wedge N} = \sum_{i=1}^{\gamma_{a,b}^N} \tilde{M}_{\tau_i \wedge N} - \tilde{M}_{\sigma_i \wedge N} + \underbrace{\tilde{M}_{\tau_{k+1} \wedge N}}_{\tilde{M}_N} - \underbrace{\tilde{M}_{\sigma_{k+1} \wedge N}}_{\tilde{M}_{\sigma_{k+1}}} + 0 = (B).$$

  Now, we observe that for every n we have $\tilde{M}_n \geq a$, so

$$\tilde{M}_N - \underbrace{\tilde{M}_{\sigma_{k+1}}}_{=a} \geq a - a \geq 0$$

  and in conclusion $(B) \geq (A) \geq \gamma_{a,b}^N (b-a)$.

- We are happy because we have obtained an estimate independent of the random variable $\gamma_{a,b}^N$, but that is just dependent upon N, that is a fixed constant.

2. Second Step.

   - We simply deduce the following inequality that follows from the telescopic series,

$$\begin{aligned}
\tilde{M}_N - \tilde{M}_{\sigma_1 \wedge N} &= \sum_{i=1}^{N} \left(\tilde{M}_{\sigma_{i+1} \wedge N} - \tilde{M}_{\sigma_i \wedge N} \right) \pm \left(\tilde{M}_{\tau_i \wedge N} \right) = \\
&= \sum_{i=1}^{N} \left(\tilde{M}_{\sigma_{i+1} \wedge N} - \tilde{M}_{\tau_i \wedge N} \right) + \sum_{i=1}^{N} \left(\tilde{M}_{\tau_i \wedge N} - \tilde{M}_{\sigma_i \wedge N} \right) \geq \\
&\geq \underbrace{\sum_{i=1}^{N} \left(\tilde{M}_{\sigma_{i+1} \wedge N} - \tilde{M}_{\tau_i \wedge N} \right) + \gamma_{a,b}^N (b-a)}_{:=(C)}.
\end{aligned}$$

   In fact we observe that surely $N \leq \sigma_{N+1}$ since $k \leq N$ and $N \leq \tau_{k+1} \leq \sigma_{k+1} \leq \sigma_{N+1}$.

3. Third Step.

   - We firstly observe that
     - $(\tilde{M}_n)_{n \geq 0}$ is a $sub-martingale$,
     - $\tau_i \wedge N \leq \sigma_{i+1} \wedge N \leq N$ are two bounded stopping times,

     implies that $\mathbb{E}[\tilde{M}_{\sigma_{i+1} \wedge N} - \tilde{M}_{\tau_i \wedge N}] \geq 0$, since Lemma (9.13) holds true.
   - We can also say that

$$\tilde{M}_N - \tilde{M}_{\sigma_1 \wedge N} = (M_N - a)_+ - (M_{\sigma_1 \wedge N})_+ \leq (M_N - a)_+.$$

- We remember that we have supposed always that $\gamma_{a,b}^N(\omega) \geq 1$, that is $\tau_1(\omega) < +\infty$ (ω was fixed). Moreover, if $\gamma_{a,b}^N(\omega) = 0$, that is $\tau_1(\omega) = +\infty$, we can save ourselves anyway, because an immediate count show us that

  $$(C)(\omega) = 0 \leq (M_N(\omega) - a)_+,$$

  So we have that $(C)(\omega) \leq (M_N(\omega) - a)_+$ everywhere in Ω.
- Now, if we take expectation in what we had obtained in Step 2 we deduce immediately by using what we said above that

  $$(b - a)\mathbb{E}[\gamma_{a,b}^N] \leq \mathbb{E}[(M_N - a)_+].$$

  and this is the thesis.

□

Now we are ready to give sufficient condition to have *a.c.* convergence of the sub-martingale.

Theorem 9.17 (Doob Convergence Theorem)**.** *Let us have $(M_n)_n$ a sub-martingale. Let us suppose that*

$$\sup_{n\in\mathbb{N}}\{\mathbb{E}[(M_n)_+]\} < +\infty.$$

Then we have that there exists $M_\infty : (\Omega, \mathscr{F}) \to (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ measurable such that the sequence of r.r.v. $(M_n)_n$ converges a.c. to M_∞, and the limit is integrable, that is $M_\infty \in \mathcal{L}^1(\Omega, \mathscr{F}, \mathbb{P})$.

Proof. The proof follows from the above one.

- We firstly observe that, since M is a sub-martingale, we have that

  $$\sup_{n\in\mathbb{N}}\{\ \mathbb{E}[(M_n)_+]\ \} < +\infty \iff \sup_{n\in\mathbb{N}}\{\ \mathbb{E}[|M_n|]\ \} < +\infty.$$

  In fact one arrow is obvious since $(M)_+ \leq |M|$. The other is easy since we have the equality $|x| = 2(x)_+ - x$, so if we substitute M_n to x we obtain

  $$\mathbb{E}[|M_n|] \leq 2\mathbb{E}[(M_n)_+] - \mathbb{E}[M_n] \leq 2\mathbb{E}[(M_n)_+] - \mathbb{E}[M_0],$$

  where in the last inequality we used that M is a sub-martingale.

- Now, let us enter in the core of the proof.

  - Let us fic $a \in \mathbb{Q}$ and $\in \mathbb{Q}$ with $a < b$.
  - Now, we have that for every $N \geq 1$ that

    $$\mathbb{E}[\gamma_{a,b}^N](b - a) \leq \mathbb{E}[(M_N - a)_+] \leq \mathbb{E}[(M_n)_+] + |a| < +\infty$$

    and $\gamma_{a,b}^N \uparrow_N \gamma_{a,b}$, so by monotone convergence we obtain

    $$\mathbb{E}[\gamma_{a,b}] < +\infty,$$

    and from this we deduce that $0 \leq \gamma_{a,b} < +\infty$ *a.c.*

– Now, let us set $N_{a,b} \coloneqq \{\omega : \ \gamma_{a,b}(\omega) < +\infty\}$, and let us set

$$N \coloneqq \bigcap_{a,b\in\mathbb{Q},\ a<b} N_{a,b}.$$

Since for every a and b we have $\mathbb{P}(N_{a,b}) = 1$, we have also that $\mathbb{P}(N) = 1$.

– Now we just need to observe that, since (9.14) hold true, we have that

$$N = \{\omega : \ \forall a < b \ rational, \ \gamma_{a,b}(\omega) < +\infty\} = \{\omega : \ \exists \ \lim_n M_n(\omega) \in \mathbb{R} \cup \{\pm\infty\}\},$$

so M_n converges $a.c.$ to the r.v. $M_\infty(\omega) = I_N(\omega)(\lim_n M_n(\omega))$.

– In a nutshell, it is convenient to see that $M_\infty = \underline{\lim}_n M_n$ (it is surely measurable), and that $M_n \to M_\infty$ $a.c.$

- $M_\infty \in \mathcal{L}^1(\Omega, \mathscr{F}, \mathbb{P})$. We simply use $Fatou$,

$$\mathbb{E}[\ |M_\infty|\ ] = \mathbb{E}[\ \underline{\lim}\, |M_n|\ ] \leq \underline{\lim}\, \mathbb{E}[\ |M_n|\ ] \leq \sup_n \mathbb{E}[\ |M_n|\ ] < +\infty,$$

So again $M_\infty \in \mathcal{L}^1$, and $|M_\infty| < +\infty$ $a.c.$

□

9.4 Characterization Of Convergence for Martingale

Now we would like to characterize the convergence in mean $\mathcal{L}^1$ for Martingale. We need some results and a definition.

Definition 49 (Uniformly Integrable)**.** Let $X = (X_i)_{i\in I}$ be a family of $r.v.$ We say that X is uniformly integrable (U.I.) if

$$\lim_{k\to+\infty} \left(\sup_{i\in I} \ \{\mathbb{E}[\ |X_i| I_{\{|X_i|\geq k\}}\ ]\} \right).$$

9.4.1 Family of U.I. r.r.v.

Let us have $(\Omega, \mathscr{F}, \mathbb{P})$ a probabilistic space. Let $X : \Omega \to \mathbb{R}$ be a $r.r.v.$ with $X \in \mathcal{L}^1$ and let us consider

$$\mathcal{S} \coloneqq (\ \mathbb{E}[X|\mathcal{G}] \ | \ \mathcal{G} \subseteq \mathscr{F} \ is \ a \ \sigma - field \ ).$$

Lemma 9.18 (Family U.I.)**.** *It holds true that $\mathcal{S}$ is U.I.*

Proof. We just give a sketch of the proof.

- Let us fix $\mathcal{G} \subseteq \mathscr{F}$ a $\sigma - field$ and let us fix k a natural number.
- We define as $Y_{\mathcal{G},X} \coloneqq \mathbb{E}[X|\mathcal{G}]$ (a version of) the conditional expectation of X wrt $\mathcal{G}$.

- Now, we have

$$
\begin{aligned}
0 &\le \mathbb{E}\left[\ |Y_{\mathcal{G},X}|I_{\{|Y_{\mathcal{G},X}|\ge k\}}\ \right] \le \mathbb{E}\left[|X|I_{\{|Y_{\mathcal{G},X}|\ge k\}}\right] \le \\
&\le \mathbb{E}[|X|I_{\{|Y_{\mathcal{G},X}|\ge k\}\cap\{|X|\ge\sqrt{k}\}}] + \mathbb{E}[|X|I_{\{|Y_{\mathcal{G},X}|\ge k\}\cap\{|X|<\sqrt{k}\}}] \le \\
&\le \mathbb{E}[|X|I_{\{|X|\ge\sqrt{k}\}}] + \sqrt{k}\mathbb{P}(|Y_{\mathcal{G},X}|\ge k) \le \\
&\le \mathbb{E}[|X|I_{\{|X|\ge\sqrt{k}\}}] + \frac{\sqrt{k}\mathbb{E}[|Y_{\mathcal{G},X}|]}{k} \le \\
&\le \mathbb{E}[|X|I_{\{|X|\ge\sqrt{k}\}}] + \frac{\mathbb{E}[|X|]}{\sqrt{k}}.
\end{aligned}
$$

  We have used the basic property of the conditional expectation and the Markov inequality.

- We conclude because the last term vanishes thanks to dominated convergence and because we have a sequence that goes to zero, and this is independent of the sigma-filed $\mathcal{G}$.

□

Theorem 9.19 (Vitali's Convergence Theorem)**.** *Let X_n be a sequence of r.v. that are integrable, so $X_n \in \mathcal{L}^1$ and $X \in \mathcal{L}^1$. Then the following statements are equivalents*

1. *$X_n \to X$ in $\mathcal{L}^1$,*
2. *The following two conditions hold true,*
   - *$X_n \to X$ in probability.*
   - *$(X_n)_n$ is U.I.*

Theorem 9.20 (Characterization of U.I. Martingale)**.** *Let us have $M = (M_n)_n$ a U.I. martingale. Then*

1. *There exists $M_\infty \in \mathcal{L}^1$ s.t. $M_n \to M_\infty$ a.c. and in $\mathcal{L}^1$.*
2. *For every $n \ge 1$, we have $M_n = \mathbb{E}[M_\infty|\mathscr{F}_n]$.*

Proof. The proof is straightforward given the above results.

1. - Let us have K_0 s.t. $k \ge K_0$ implies that $\sup_{n\ge 0}(\mathbb{E}[|M_n|I_{\{|M_n|\ge K\}}]) \le 1$.
   - So, we can write that

     $$\mathbb{E}[|X_n|] = \mathbb{E}[|...|I_{\{|X_n|\ge K_0\}}] + \mathbb{E}[|...|I_{\{|X_n|\le K_0\}}] \le 1 + K_0.$$

   - So we have that $\sup_{n\ge 0}\mathbb{E}[|X_n|] < +\infty$.
   - Now, M is a martingale (so it is a sub-martingale) and the sup-condition hold true, so by Theorem (9.17) we have that $M_n \to M_\infty$ a.c, and $M_\infty \in \mathcal{L}^1$.

- Now, $M_n \to M_\infty$ *a.c.* implies that $M_n \to M_\infty$ in probability, and with the fact that M is $U.I.$ we have by Theorem (9.19) that $M_n \to M_\infty$ in $\mathcal{L}^1$.

2. - We firstly observe that for every n, for every $A \in \mathscr{F}_n$, we have that

$$\mathbb{E}[M_m I_A] \to \mathbb{E}[M_\infty I_A].$$

   In fact, $|M_m - M_\infty| I_A \leq |M_m - M_\infty|$ that goes to zero in $\mathcal{L}^1$, so by comparison we obtain that $\mathbb{E}[M_m I_A] \to \mathbb{E}[M_\infty I_A]$ if $m \uparrow +\infty$.

   - Now, let us fix n. If we prove that for every $A \in \mathscr{F}_n$ we have that

$$\mathbb{E}[M_\infty I_A] = \mathbb{E}[M_n I_A].$$

   we obtain the thesis, that is $\mathbb{E}[M_\infty | \mathscr{F}_n] = M_n$.

   - This is simple because, since M is a martingale, for every $m \geq n$ we obtain that $M_n = \mathbb{E}[M_m | \mathscr{F}_n]$, so for every $A \in \mathscr{F}_n$ we have

$$\mathbb{E}[M_n I_A] = \mathbb{E}[M_m I_A] \to \mathbb{E}[M_\infty I_A] \ \ if \ \ m \uparrow +\infty,$$

   that is the thesis because the sequence become eventually constant.

□

Corollary 9.21 (Levi Corollary)**.** *Let us have the following setting.*

- *Let $X \in \mathcal{L}^1(\Omega, \mathscr{F}, \mathbb{P})$ be a r.v. and let $(\mathscr{F}_n)_n$ be a filtration.*
- *Let us define for every $n \geq 0$ the r.v. $M_n = \mathbb{E}[X | \mathscr{F}_n]$.*

Then we have that $M = (M_n)_n$ is a martingale, and $M_n \to \mathbb{E}[X | \mathscr{F}_\infty]$ a.c. and in $\mathcal{L}^1$, with $\mathscr{F}_\infty = \sigma(\cup_n \mathscr{F}_n)$.

Proof. The proof follows in part from Theorem (9.20) and later from Theorem (2.4).

- It is immediate to show that M is a martingale.
- It is an exercise to show that M is $U.I.$
- So by Theorem (9.20) we have that $M_n \to M_\infty$ a.c. and in $\mathcal{L}^1$, and $\mathbb{E}[M_\infty | \mathscr{F}_n] = M_n$.
- We have to show that $M_\infty = \mathbb{E}[X | \mathscr{F}_\infty]$.
- We firstly observe that M_∞ is $\mathscr{F}_\infty - measurable$ because it is defined in (9.17) as $\underline{\lim}$ of $\mathscr{F}_\infty - measurable$ function (in fact $\mathscr{F}_n \subseteq \mathscr{F}_\infty$ for every n).
- We observe that $F = \cup_n \mathscr{F}_n$ is a $\pi - system$ for $\sigma(F)$ because $\mathscr{F}_n \subseteq \mathscr{F}_{n+1}$ for every n (easy exercise).

- Now, let us have $A \in F$. Then we have that $A \in \mathscr{F}_{n_0}$ for some n_0 natural. So, if we have $n \geq n_0$, we have that $A \in \mathscr{F}_n$, and we can write

$$\mathbb{E}[XI_A] = \mathbb{E}[\underbrace{\mathbb{E}[X|\mathscr{F}_n]}_{M_n} I_A] = \mathbb{E}[M_n I_A] \to \mathbb{E}[M_\infty I_A] \ \ if \ \ n \uparrow +\infty$$

The convergence to $\mathbb{E}[M_\infty I_A]$ follows from the convergence in $\mathcal{L}^1$ of M_n to M_∞, so we have that $\mathbb{E}[M_\infty I_A] = \mathbb{E}[XI_A]$ because the sequence become eventually constant.

- Now, since F is a $\pi - system$ for $\sigma(F)$, we have that $\mathbb{E}[XI_A] = \mathbb{E}[M_\infty I_A]$ for every $A \in \mathscr{F}_\infty$ thanks to a corollary of Theorem (2.4), that is $\mathbb{E}[X|\mathscr{F}_\infty] = M_\infty$.

□

9.5 Quadratic Variation For Martingale Definition

Let's start with an observation

- Let us have $M = (M_n)_n$ a martingale.
- By (9.1), we have that $((M_n)^2)_n$ is a sub-martingale, so by (9.12) we have that we can find $\langle M \rangle = (\langle M \rangle_n)_n$ an increasing predictable process that bring us to the decompositon.
- **Definition 50.** We call the process $\langle M \rangle_n$ the quadratic variation of M.
- We can deduce an explicit formula for $\langle M \rangle$, that is $\langle M \rangle_0 = 0$, and for every $n \geq 1$

$$\langle M \rangle_n = \sum_{k=1}^{n} \mathbb{E}[(M_k)^2 - (M_{k-1}^2)|\mathscr{F}_{k-1}] = \sum_{k=1}^{n} \mathbb{E}[(M_k - M_{k-1})^2|\mathscr{F}_{k-1}],$$

where the last equality follows directly from a direct count.

9.6 Quadratic Variation and a.c. limit

- Let us have M a discrete time martingale that is square integrable and let us have $\langle M \rangle$ its quadratic variation.
- We have that $\langle M \rangle$ is an *a.c.* increasing process, so there exists the limit of $\langle M \rangle_n(\omega)$ a.c.
- In particular we have that $\langle M \rangle_n \uparrow \langle M \rangle_\infty$ a.c. and we have by monotone convergence that $\mathbb{E}[\langle M \rangle_n] \uparrow \mathbb{E}[\langle M \rangle_\infty]$.
- Since $(M_n)^2 = (M_0)^2 + N_n + \langle M \rangle_n$ for every $n \geq 0$, and N is a martingale, we have that $\mathbb{E}[(M_n)^2] = cost + \mathbb{E}[\langle M \rangle_n]$, so we obtain that

$$\sup_n \ \{ \ \mathbb{E}[(M_n)^2] \ \} < +\infty \iff \mathbb{E}[\ \langle M \rangle_\infty \ ] < +\infty.$$

9.7 Local Martingale

Definition 51 (Local Martingale)**.** We say that M is a *Local Martingale* with respect to $(\mathscr{F}_t)_{t\geq 0}$ if the following properties hold true,

- M is adapted with respect to $(\mathscr{F}_t)_{t\geq 0}$,
- there exists $(\tau_n)_{n\in\mathbb{N}}$ r.r.v such that
  - τ_n is a stopping time for all n,
  - $\tau_n \nearrow +\infty$,
  - for all $n \in \mathbb{N}$, $(M_{t\wedge\tau_n})_{t\geq 0}$ is a martingale.

9.8 Quadratic Variation for Martingales

We use the following notation.

- Let us have $T > 0$. We indicate as π a partition of the interval $[0, T]$, with π defined as

$$\pi = \{0 = t_0 < t_1 < t_2 < ... < t_n = T\}.$$

- The *size* of a partition is defined as

$$|\pi| = \max_{1\leq k\leq n} |t_k - t_{k-1}|.$$

- A sequence of partition $(\pi_n)_{n\in\mathbb{N}}$ is *nested* if for all $n \in \mathbb{N}$, we have

$$\pi_n \subseteq \pi_{n+1}.$$

- If we fix a partition π, we denote as $(\langle M\rangle_t^\pi)_{t\in[0,T]}$ the process such that

$$\langle M\rangle_t^\pi := \sum_{i:\ t_{i+1}\leq t} (M_{t_{i+1}} - M_{t_i})^2 + (M_t - M_{t_k})^2,$$

  with $k := max\{j : \ t_j \leq t\}$. The sum in this way is extended from 0 to $k-1$.

- Given $(X_t)_{t\geq 0}$ a *S.P.*, we define for all $T > 0$ the following

$$\|X\|_{\infty,T}(\omega) := \sup_{t\in[0,T]} \{|X_t(\omega)|\}.$$

  (The variable ω can be omitted if it is not ambiguous).

Theorem 9.22 (Mega Theorem)**.** *Let us have $(M_t)_{t\geq 0}$ a continuous martingale wrt a filtration $(\mathscr{F}_t)_{t\geq 0}$. Then*

1. *we can find $(A_t)_{t\geq 0}$ a continuous non-decreasing process, adapted wrt $(\mathscr{F}_t)_{t\geq 0}$ such that*

- *for every $T > 0$,*
- *for every sequence $(\pi_n)_{n\in\mathbb{N}}$ of nested partitions of $[0,T]$ with $|\pi_n| \to 0$,*

we have that for every $\epsilon > 0$,

$$\mathbb{P}(\ \left\|\langle M\rangle^{\pi_n} - A\right\|_{\infty,T} \geq \epsilon\ ) \to 0 \ \ if\ \ n \uparrow +\infty,$$

that is $\langle M\rangle^{\pi_n} \to A$ in probability, and this limit is independent of the partition.

2. *Moreover, we have the following two properties,*
   - *$(M_t^2 - A_t)_{t\geq 0}$ is a martingale.*
   - *If $(A'_t)_{t\geq 0}$ is another process such that*
     * $(A'_t)_{t\geq 0}$ *is* $(\mathscr{F}_t)_{t\geq 0}$ *− adapted,*
     * $(A'_t)_{t\geq 0}$ *is continuous,*
     * $(M_t^2 - A'_t)_{t\geq 0}$ *is a martingale,*
     * $A'_0 = 0$,

     then A and A' are indistinguishable.

Definition 52 (Quadratic Variation for Martingle)**.** The process A found above is called the *quadratic variation* of the martingle M, and it is denoted by $\langle M\rangle = (\langle M\rangle_t)_{t\geq 0}$.

9.9 Semi-Martingale

9.9.1 BV Function

Let us denote as Σ the set of all the partitions of $[a,b]$, that is

$$\pi \in \Sigma \implies \pi = \{a = t_0 < t_1 < t_2 < ... < t_{n-1} < t_n = b\}.$$

for some $n \in \mathbb{N}$.

Definition 53 (BV Function)**.** Let $f : [a,b] \to \mathbb{R}$ be a function. We say that f is of *Bounded Variation* (BV) if

$$\sup_{\pi\in\Sigma} \left\{ \sum_{t_k\in\pi} |f(t_{k+1}) - f(t_k)| \right\} < +\infty.$$

The following theorem hold true. Sooner or later we prove it.

Theorem 9.23. *Let $f : [a,b] \to \mathbb{R}$ be a function. Then the following statements are equivalent,*

1. *f is BV.*
2. *There exist $f_1 : [a,b] \to \mathbb{R}$ and $f_2 : [a,b] \to \mathbb{R}$ such that*
   - *f_1 and f_2 are non-decreasing.*

- $f = f_1 - f_2$.

Now we give a notion of BV for $S.P.$

- Let us have $(\Omega, \mathscr{F}, \mathbb{P})$ and a *probabilistic space*.
- Let us have $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$.
- Let $X : [0, +\infty) \times \Omega \to \mathbb{R}$ be a $S.P.$ We denote X as $(X_t)_{t\geq 0}$.

Definition 54 (BV for $S.P.$)**.** Let us define, for all $0 \leq a < b < +\infty$.

$$N := \{\omega \in \Omega : \ \forall \ [a, b] \subseteq [0, +\infty), \ X|_{[a,b]\times\{\omega\}} \ is \ BV\}.$$

We say that X is BV is $\mathbb{P}(N) = 1$, that is its trajectory are $a.c.$ BV functions.

Now we can give a new definition. Let us have a filtration $(\mathscr{F}_t)_{t\geq 0}$.

Definition 55 (Semi-Martingale)**.** We say that X is a $semi-martingale$ (wrt filtration $(\mathscr{F}_t)_{t\geq 0}$) if there exist

1. $(M_t)_{t\geq 0}$ a local martingale,
2. $(V_t)_{t\geq 0}$ a BV process,

such that $X_t = M_t + V_t$. We say that X is a $continuous\ semi-martingale$ if M and V are continuous.

Now we enunciate an important theorem that holds for $semi-martingles$.

Theorem 9.24. *Let $(X_t)_{t\geq 0}$ a continuous semi-martingale. The we have the following facts.*

1. *The decomposition*

$$X_t = X_0 + M_t + V_t$$

where

- *M_t is a continuous local martingale such that $M_0 = 0$,*
- *V_t is a continuous BV process such that $V_0 = 0$,*

is unique.

2. *The following is a statement on the existence of a limit (in probability).*

- *Let us fix $t > 0$.*
- *Let us consider $(\pi_k)_{k\in\mathbb{N}}$ a sequence of partitions of $[0, t]$ such that*
  * $\pi_k = \{0 = t_0 < t_1 < ... < t_{n_k} = t\}$,
  * $|\pi_k| = \sup\{t_{i+1} - t_i : \ i = 0, .., n_k - 1\} \to 0$ *when* $k \uparrow +\infty$.

Then the following limits exist in probability (they are independent of the partitions) and are equal,

$$\lim_{k\to+\infty}\sum_{i=0}^{n_k-1}|X_{t_{i+1}}-X_{t_i}|^2=\lim_{k\to+\infty}\sum_{i=0}^{n_k-1}|M_{t_{i+1}}-M_{t_i}|^2.$$

Definition 56. The limit quantity of the above theorem is denoted as

$$\langle X\rangle_t \coloneqq \lim_{k\to+\infty}\sum_{i=0}^{n_k-1}|X_{t_{i+1}}-X_{t_i}|^2,$$

and it is called the *Quadratic Variation* of X in $[0,t]$.

10 Brownian Motion

10.1 Gaussian Processes

Let T be an arbitrary index set.

Definition 57. Let $X=(X_t)_{t\in T}$ be a real valued *S.P.* We say that X is *Gaussian* if $\forall t_1,..,t_n\in T$, we have that $(X_{t_1},..,X_{t_n})$ is a *Gaussian Vector*.

Let $X=(X)_{t\in T}$ be a Gaussian Process.

Definition 58. We can define

- $E(X)(t)\coloneqq m(t)\coloneqq \mathbb{E}[X_t]$,
- $Cov(X)(t,s)\coloneqq C(t,s)\coloneqq Cov(X_t,X_s)$, with $s,t\in T$.

Remark 35. We can even highlight the dependence of m and C from the *S.P*, so we can call these function m_X and C_X.

Let $(\Omega,\mathscr{F},\mathbb{P})$ be a probabilistic space.

Proposition 10.1. *Let $X=(X_t)_{t\in T}$ and $(Y_t)_{t\in T}$ be two Gaussian S.P. If we have*

- *$m_X(t)=m_Y(t)$ for all $t\in T$,*
- *$C_X(t,s)=C_Y(t,s)$ for all $t,s\in T$.*

Then X and Y have the same law.

Remark 36. The law of a *S.P.* is defined in (24).

Proof. The proof is made in this way.

- Because of Proposition (7.3), we just need to check that X and Y have the same *finite dimensional distributions.*

- Let us have $\bar{t} = (t_1, .., t_n)$ with $t_i \in T$ for all i, and $t_i = t_j$ if, and only if $i = j$.
- Since we have that $X_{\bar{t}}$ and $Y_{\bar{t}}$ are $Gaussian\ Vectors$, we have that their law is uniquely determined by the vector of the mean and the matrix of covariance.
- It is a fast check to control that $\mathbb{E}[X_{\bar{t}}] = \mathbb{E}[Y_{\bar{t}}]$ and $Cov(X_{\bar{t}}) = Cov(Y_{\bar{t}})$, so $X_{\bar{t}}$ and $Y_{\bar{t}}$ have the same law, because they have the same characteristic function.

□

Let T be a set.

Definition 59. Let $C : T \times T \to \mathbb{R}$. We say that C is positive semi-definite if for every $n \geq 1$, for all $t_1, .., t_n \in T$, and for all $(\xi_1, .., \xi_n) \in \mathbb{R}^n$, we have that

$$\sum_{i,j=1}^{n} C(t_i, t_j)\xi_i\xi_j \geq 0.$$

Remark 37. In practice, let us have

- $t_1, .., t_n \in T$, with $(t_1, .., t_n)$.
- $(\xi_1, .., \xi_n) \in \mathbb{R}^n$.

If we set

- $C_{\bar{t}} = \big(C(t_i, t_j)\big)_{i,j=1,..,n}$.
- $\xi = (\xi_1, .., \xi_n)^T$.

we have

$$(\xi^T)C_{\bar{t}}(\xi) \geq 0.$$

Remark 38. We remember that C is symmetric if $C(s, t) = C(t, s)$ for all $t \in T$ and $s \in T$.

Let T be a set of index.

Proposition 10.2 (Existence Gaussian Process)**.** *Let us have*

- $m : T \to \mathbb{R}$*,*
- $C : T \times T \to \mathbb{R}$*, a symmetric, positive semi-definite function.*

Then there exists a Gaussian process X such that $E(X)(t) = m(t)$ and $Cov(X)(t, s) = C(t, s)$, with E and Cov defined in (58).

Proof. **ON WORK**. □

10.2 Definitions

Let us set $T = [0, +\infty)$. Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probabilistic space. Let $B : T \times \Omega \to (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ be a *S.P.* (a real one).

Definition 60 (Intrinsic Brownian Motion)**.** We say that B is a (*standard*) *Brownian Motion* (*B.m.* for friends) if

- $\mathbb{P}\big(\{\omega : B_0(\omega) = 0\}\big) = 1$.
- for all $0 \leq s < t$, we have that $B_t - B_s \sim N(0, t - s)$.
- for all $n \geq 1$, for all $0 \leq t_1 \leq t_2 \leq ... \leq t_n$, we have that

$$B_{t_1},\ B_{t_2} - B_{t_1},\ B_{t_3} - B_{t_2},\ ...,\ B_{t_n} - B_{t_{n-1}}.$$

  are independent random variables.
- B is *a.c.* continuous, as defined in (27).

Remark 39. We remember that $X \sim ..$ means *"The random variable X has law .."*, and the *law of X* is simply the probability $\mathbb{P}_X$.

Definition 61. A process that satisfy every condition but continuity of trajectory is called *Brownian Motion in Law.*

Definition 62 (Wiener measure)**.**

- Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probabilistic space.
- Let $B : T \times \Omega \to (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ be a *S.P.* (a real one).

Definition 63 (Bm with respect to a given Filtration.)**.** We say that B is a (*standard*) *Brownian Motion* (*B.m.* for friends), adapted with respect to filtration $(\mathscr{F}_t)_{t\geq 0}$ if

- $\mathbb{P}\big(\{\omega : B_0(\omega) = 0\}\big) = 1$.
- for all $0 \leq s < t$, we have that $B_t - B_s \sim N(0, t - s)$.
- for all $0 \leq s < t$, the r.v. $B_t - B_s$ is independent from $\mathscr{F}_s$.
- B is *a.c.* continuous, as defined in (27).

Remark 40. We observe that, since $B = (B_t)_{t\geq 0}$ is adapted, if we have $0 \leq s < t$ then we have

- B_t is $\mathscr{F}_t$ measurable,
- B_s is $\mathscr{F}_s$ measurable and $\mathscr{F}_s \subseteq \mathscr{F}_t \implies B_s$ is $\mathscr{F}_t$ measurable.

Then $B_t - B_s$ is $\mathscr{F}_t$ measurable.

Let us have the previous setting.

Proposition 10.3. *Let B be a S.P. that is a Bm according to Definition* (63)*, with respect to a given filtration* $(\mathscr{F}_t)_{t\geq 0}$*. Then B is a Bm according to* (60)*.*

Proof. The proof is not hard. We just need to prove that the increments are independent because the other properties remain the same.

- Let $0 \leq t_1 \leq t_2 \leq ... \leq t_n \in T$ be real numbers. We prove our thesis by induction.
- $n = 2$.
  - We have B_{t_1} and $B_{t_2} - B_{t_1}$.
  - We have that $B_{t_2} - B_{t_1}$ is independent of $\mathscr{F}_{t_1}$ by our hypothesis, but $\sigma(B_{t_1}) \subseteq \mathscr{F} t_1$ because B is adapted, so $B_{t_2} - B_{t_1}$ is independent of $\sigma(B_{t_1})$, that is $B_{t_2} - B_{t_1}$ are independent.
- $n > 2$
- $B_{t_n} - B_{t_{n-1}}$ is independent of $\mathscr{F}_{t_{n-1}}$ and $B_{t_{n-1}} - B_{t_{n-2}}$,..., $B_{t_2} - B_{t_1}, B_{t_1}$ are $\mathscr{F}_{t_{n-1}}$ measurable $\implies$ $B_{t_n} - B_{t_{n-1}}$ and $B_{t_{n-1}} - B_{t_{n-2}}$,..., $B_{t_2} - B_{t_1}, B_{t_1}$ are independents.
- by inductive hypothesis, we have that $B_{t_{n-1}} - B_{t_{n-2}}$,..., $B_{t_2} - B_{t_1}, B_{t_1}$ are independents. Thanks to Corollary (2.12), we easily conclude.

□

Now we want to prove that *Definition* (60) $\implies$ Definition(63) if we chose a proper filtration.

Remark 41. Let $X = (X_t)_{t\geq 0}$ be a *S.P.* We recall that in Section (8.2.2) we have introduced

- the filtration $(\tilde{\mathscr{F}}_t^X)_{t\geq 0}$, where $\tilde{\mathscr{F}}_t^X := \sigma(X_s : 0 \leq s \leq t)$,
- the filtration $(\overline{\mathscr{F}}_t^X)_{t\geq 0}$, that is the completion of the above filtration,
- the filtration $(\mathscr{F}_t^X)$ that is the right continuous filtration associated to the one above. This is the filtration $right - C^0$ generated by the process.

Proposition 10.4 (Bm with respect to generated filtration)**.** *Let us have*

- $B = (B_t)_{t\geq 0}$ *a Bm according to Definition* (60).

Then B is a Bm with respect to $(\tilde{\mathscr{F}}_t^B)_{t\geq 0}$, according to Definition (63).

Proof. The proof is this.

- B is adapted with respect to $(\tilde{\mathscr{F}}_t^B)_{t\geq 0}$ by definition.
- The only non trivial property is the one on the independence. Let's see.

- Let us define

$$\mathcal{A}_n := \left\{ \bigcap_{i=1}^{n} \{B_{s_i} \in A_i\} \ : \ 0 \le s_1 < s_2 ... < s_n \le s \ \textit{and} \ A_i \in \mathfrak{B}(\mathbb{R}) \right\},$$

$$\mathcal{A} := \bigcup_{n=1}^{+\infty} \mathcal{A}_n.$$

We say that $\mathcal{A}$ is a $\pi - system$ for $\tilde{\mathscr{F}}_s^B$.

  - It is trivially closed by intersection and it contains Ω.
  - It generate, that is $\sigma(\mathcal{A}) = \tilde{\mathscr{F}}_s^B$. In fact, we have

$$\forall \ s_1 : \ 0 \le s_1 \le s, \ \sigma(B_{s_1}) \subseteq \mathcal{A}_1 \subseteq \mathcal{A} \implies$$
$$\tilde{\mathscr{F}}_s^B = \sigma \left(\bigcup_{0 \le s_1 \le s} \sigma(B_{s_1}) \right) \subseteq \sigma(\mathcal{A}) \subseteq \tilde{\mathscr{F}}_s^B.$$

- Now, thanks to Corollary (2.9), since
  - $\mathcal{A}$ is a $\pi - system$ for $\tilde{\mathscr{F}}_s^B$,
  - $\sigma(B_t - B_s)$ is a $\pi - system$ for itself.

If we prove that

$$\forall A_1 \in \sigma(B_t - B_s) \ \textit{and} \ \forall A_2 \in \mathcal{A} \ , \textit{then} \ \mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2) \tag{16}$$

then we have that $B_t - B_s$ and $\tilde{\mathscr{F}}_s^B$ are independent, as we wanted.

- We proceed in this way.
  - Let us fix $A \in \mathcal{A}$.
    So we have that $A = \bigcap_{i=1}^{n} \{B_{s_i} \in A_i\}$, with $0 \le s_1 < s_2 < ... < s_n \le s$ and $A_i \in \mathfrak{B}(\mathbb{R})$ for all i.
  - Now, we have that $B_{s_1}, B_{s_2} - B_{s_1}, .., B_{s_n} - B_{s_{n-1}}, B_t - B_s$ are independent since we have assumed Definition (60).
  - Thanks to Corollary (2.12), we have that (if we indicate as $B_{s_0} = 0$),

$$\begin{aligned} \sigma(B_t - B_s) \ \perp \ & \bigvee_{i=1}^{n} \sigma(B_{s_i} - s_{i-1}) = \\ & = \sigma(B_{s_1}, B_{s_2} - B_{s_1}, .., B_{s_n} - B_{s_{n-1}}) \underbrace{=}_{(*)} \\ & = \sigma(B_{s_1}, B_{s_2}, .., B_{s_n}), \end{aligned}$$

    where in $(*)$ we have used what we have discover with Corollary (5.6) (it is easy to fix the detail, given the Corollary).

– Now, $A \in \sigma(B_{s_1}, B_{s_2}, .., B_{s_n})$, so for all $B \in \sigma(B_t - B_s)$ we have that

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B),$$

and this is what we wanted to obtain the thesis, because we have obtained (16).

□

Corollary 10.5 (Bm with respect to the completion of the filtration above.)**.** *Let us have*

- $B = (B_t)_{t\geq 0}$ *a Bm according to Definition* (60).

Then B is a Bm with respect to $(\overline{\mathscr{F}}_t^B)_{t\geq 0}$, according to Definition (63).

Proof. The proof uses Proposition (2.16). As before we just need to check the independence condition.

- Let us have $s < t$, and and let us consider $\sigma(B_t - B_s)$ and $\overline{\mathscr{F}}_s^B$. We want to show that they are independent.
- we remember that $\overline{\mathscr{F}}_s^B = \sigma(\tilde{\mathscr{F}}_s^B \cup \mathcal{N})$, with $\mathcal{N} = \mathcal{N}_{(\mathscr{F},\mathbb{P})}$ the $\mathbb{P}-$*null sets* with respect to $\mathscr{F}$.
- Thanks to Remark (25), we know that $\tilde{\mathscr{F}}_s^B \cup \mathcal{N}$ is a $\pi -$ *system* for $\overline{\mathscr{F}}_s^B$.
- Now we want to use (2.16) with $\mathcal{I} = \tilde{\mathscr{F}}_s^B \cup \mathcal{N}$ and $X = B_t - B_s$.
- So, let us have $A \in \tilde{\mathscr{F}}_s^B$ and $N \in \mathcal{N}$. Then, for all $\varphi \in C_B^0(\mathbb{R})$ we have

$$\mathbb{E}[\varphi(B_t - B_s)I_{A\cup N}] = \mathbb{E}[\varphi(B_t - B_s)I_A] + \mathbb{E}[\varphi(B_t - B_s)I_{N\setminus A}] =$$
$$\mathbb{E}[\varphi(B_t - B_s)I_A] = \mathbb{E}[\varphi(B_t - B_s)]\mathbb{P}(A) = \mathbb{E}[\varphi(B_t - B_s)]\mathbb{P}(A \cup N).$$

  We have used that $\sigma(B_t - B_s)$ and $\tilde{\mathscr{F}}_s$ are independents thanks to the Proposition above, and $N \setminus A \in \mathcal{N}$ (so the integral on it is equal to zero) and $\mathbb{P}(A) = \mathbb{P}(A \cup N)$.

- We have concluded because we have the hypothesis of *Proposition* (2.16).

□

Corollary 10.6 (Bm with respect to the completion Right-C^0 filtration above.)**.** *Let us have*

- $B = (B_t)_{t\geq 0}$ *a Bm according to Definition* (60).

Then B is a Bm with respect to $(\mathscr{F}_t^B)_{t\geq 0}$, according to Definition (63).

Proof. The proof is not hard and it follows from the previous one by using an argument of continuity.

- As before, we just need to check the independence condition.
- So, let us have $s < t$. We want to show that $B_t - B_s \perp \mathscr{F}_s^B$, and we remember that $\mathscr{F}_s^B = \cap_{\epsilon>0}\overline{\mathscr{F}}_{s+\epsilon}^B$.

- We want to use Proposition (2.16) with $\mathcal{I} = \mathscr{F}_s$ and $X = B_t - B_s$, that is we want to prove that

$$\forall A \in \mathscr{F}_s^B,\ \forall \varphi \in C_B^0(\mathbb{R}),\ \mathbb{E}[\varphi(B_t - B_s)I_A] = \mathbb{E}[\varphi(B_t - B_s)]\mathbb{P}(A).$$

- So, let us have $A \in \mathscr{F}_s^B$ and $\varphi \in C_B^0(\mathbb{R})$.
- We have for all $\epsilon > 0$ such that $s + \epsilon < t$, that $A \in \overline{\mathscr{F}}_{s+\epsilon}^B$, so thanks to Corollary above we have that $(B_t - B_{s+\epsilon}) \perp \overline{\mathscr{F}}_{s+\epsilon}^B$, and this implies that

$$\mathbb{E}[\varphi(B_t - B_{s+\epsilon})I_A] = \mathbb{E}[\varphi(B_t - B_{s+\epsilon})]\mathbb{P}(A).$$

- Now, we have
  - $\varphi(B_t - B_{s+\epsilon})I_A \overset{\epsilon \to 0^+}{\to} \varphi(B_t - B_s)I_A$ a.s. $\omega \in \Omega$, because B is a continuous process and φ is a continuous function.
  - for all $\epsilon > 0$, we have that $\varphi(B_t - B_{s+\epsilon})I_A$ is dominated by a constant.

  Analogously, we have that

  - $\varphi(B_t - B_{s+\epsilon}) \overset{\epsilon \to 0^+}{\to} \varphi(B_t - B_s)$ a.s. $\omega \in \Omega$, because B is a continuous process and φ is a continuous function.
  - for all $\epsilon > 0$, we have that $\varphi(B_t - B_{s+\epsilon})$ is dominated by a constant.

  So thanks to dominated convergence, we have that

$$\begin{array}{ccc} \mathbb{E}[\varphi(B_t - B_{s+\epsilon})I_A] & \overset{\epsilon \to 0^+}{\to} & \mathbb{E}[\varphi(B_t - B_s)I_A] \\ || & & \\ \mathbb{E}[\varphi(B_t - B_{s+\epsilon})]\mathbb{P}(A) & \overset{\epsilon \to 0^+}{\to} & \mathbb{E}[\varphi(B_t - B_s)]\mathbb{P}(A), \end{array}$$

  so the limit have to be equal, that is what we wanted to prove.

$\square$

Remark 42 (Property of Markov Process)*.* We have that the Bm ha centered and independents increments so it is a Martingle and a Markov Process.

11 Stochastic Integral

We would like to generalize the notion of integrals using as integrator the Brownian motion. I don't know why we are doing this yet, but when I discover it I write something more in this introuduction.

11.1 Why it is difficult to define SI

It is hard to define the Stochastic Integrals (S.I.) with respect to Bm because it is not stable by approximation. In fact one can imagine to define $S.I.$ starting by the Rieman Integrals, that is

- We have a function $f : [a, b] \to \mathbb{R}$ that we want to integrate.
- We have a function $g : [a, b] \to \mathbb{R}$ that we want to use as integrator, that is g has the role of identity in Rieman Integral.
- So, let us have $\pi_n := \{t_0 < ... < t_n\}$ a partition of $[a, b]$, so we have that $t_0 = a$ and $t_n = b$.
- we can define

$$X_{\pi_n} := \sum_{i=0}^{n-1} f(\tilde{t}_i)(g(t_{i+1}) - g(t_i)), \;\; \tilde{t}_i \in [t_i, t_{i+1}] \;\; a \;\; generic \;\; point.$$

- Now we take the sup when we vary the partition and the points.
- We hope that this limit , if the partition is dense enough, does not depend upon the partition itself and the point that we choose.
- unfortunately, if g is a B_m, we have many problem.

11.1.1 Proof that our definition does not work with Bm

- Let us have $B = (\Omega, \mathscr{F}, \mathbb{P}, (\mathscr{F}_t)_{t \geq 0})$ a *standard Bm* (that is according to Definition (63)).
- Let us consider the dyadic partition

$$\pi_n := \{t_0^n, ..., t_{2^n}^n\},$$

and for all $k \in \{0, 1, ..., 2^n\}$, we define $t_k^n = \frac{k}{2^n}$.

- Let us consider the three following approximation,

$$X_n = \sum_{k=0}^{2^n-1} B_{\frac{k}{2^n}}(B_{\frac{k+1}{2^n}} - B_{\frac{k}{2^n}})$$
$$Y_n = \sum_{k=0}^{2^n-1} B_{\frac{k+1}{2^n}}(B_{\frac{k+1}{2^n}} - B_{\frac{k}{2^n}})$$
$$Z_n = \frac{1}{2}\sum_{k=0}^{2^n-1} (B_{\frac{k}{2^n}} + B_{\frac{k+1}{2^n}})(B_{\frac{k+1}{2^n}} - B_{\frac{k}{2^n}}) = \frac{X_n + Y_n}{2}.$$

Now we prove the following claim.

Proposition 11.1. *The sequence $(X_n)_{n\geq 0}$ converges in $L^2(\Omega)$ to a r.v. X_∞, and we have that $\mathbb{E}[X_\infty] = 0$.*

Proof. Let us define $\Delta_h^m := B_{\frac{h+1}{2^m}} - B_{\frac{h}{2^m}}$.

- Now we write in a proper way $X_{n+1} - X_n$. We have

$$X_{n+1} - X_n = \sum_{k=0}^{2^{n+1}\cdot 2-1} B_{\frac{k}{2^{n+1}}} \underbrace{(B_{\frac{k+1}{2^{n+1}}} - B_{\frac{k}{2^{n+1}}})}_{\Delta_k^{n+1}} - \sum_{k=0}^{2^n-1} B_{\frac{k}{2^n}} \underbrace{(B_{\frac{k+1}{2^n}} - B_{\frac{k}{2^n}})}_{\Delta_k^n}.$$

- Now we observe that

$$\begin{aligned} X_{n+1} &= \sum_{k=0}^{2^n\cdot 2-1} B_{\frac{k}{2^{n+1}}} \Delta_k^{n+1} = \\ &= \sum_{k=0}^{2^n-1} B_{\frac{2k}{2^{n+1}}} \Delta_{2k}^{n+1} + \sum_{k=0}^{2^n-1} B_{\frac{2k+1}{2^{n+1}}} \Delta_{2k+1}^{n+1} \end{aligned}$$

- Now we substitute and with a little algebra we obtain

$$X_{n+1} - X_n = \sum_{k=0}^{2^n-1} B_{\frac{k}{2^n}} (\Delta_{2k}^{n+1} - \Delta_k^n) + B_{\frac{2k+1}{2^{n+1}}} \Delta_{2k+1}^{n+1}$$

- Now we observe that

$$\Delta_{2k}^{n+1} - \Delta_k^n = (B_{\frac{2k+1}{2^{n+1}}} - B_{\frac{2k}{2^{n+1}}}) - (B_{\frac{k+1}{2^n}} - B_{\frac{k}{2^n}}) = -\Delta_{2k+1}^{n+1}$$

- If we substitute again, we obtain

$$X_{n+1} - X_n = \sum_{k=0}^{2^n-1} \Delta_{2k+1}^{n+1} \underbrace{(B_{\frac{2k+1}{2^{n+1}}} - B_{\frac{2k}{2^{n+1}}})}_{\Delta_{2k}^{n+1}} = \sum_{k=0}^{2^n-1} \Delta_{2k+1}^{n+1} \Delta_{2k}^{n+1}.$$

- It is immediate to see that $\{\Delta_{2k}^{n+1}\Delta_{2k+1}^{n+1}\}_{k=0,..,2^n-1}$ are orthogonal, so we have that

$$\mathbb{E}[(X_{n+1} - X_n)^2] = \sum_{k=0}^{2^n-1} = \mathbb{E}[(\Delta_{2k+1}^{n+1})^2 (\Delta_{2k}^{n+1})^2] = 4 \cdot \frac{1}{2^n}$$

  We have used that $\Delta_{2k}^{n+1} \perp \Delta_{2k+1}^{n+1}$, and they are $N(0, \frac{1}{2^{n+1}})$.

- Now it is an easy exercise to show that $(X_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and thus it converges in $L^2(\Omega)$ to $X_\infty \in L^2(\Omega)$.

$\square$

Now we make the math with the sequence $(Y_n)_{n\geq 0}$.

Proposition 11.2. *The sequence $(Y_n)_{n\geq 0}$ converges in $L^2(\Omega)$ to a r.v. Y_∞ such that $X_\infty \neq Y_\infty = B_1^2$.*

Proof. The proof lays on the above Proposition. We use the same notations that we have used in that proposition.

- We firstly observe that for all $m \in \mathbb{N}$, we have

$$\begin{aligned} Y_m = &\sum_{k=0}^{2^m-1} B_{\frac{k+1}{2^m}} \underbrace{(B_{\frac{k+1}{2^m}} - B_{\frac{k}{2^m}})}_{\Delta_k^m} = \\ = &\sum_{k=0}^{2^m-1} B_{\frac{k}{2^n}} \Delta_k^m + \sum_{k=0}^{2^m-1} \underbrace{(B_{\frac{k+1}{2^m}} - B_{\frac{k}{2^m}})}_{\Delta_k^m} \Delta_k^m = \\ = &X_m + \sum_{k=0}^{2^m-1} (\Delta_k^m)^2. \end{aligned}$$

- So, we can write

$$\begin{aligned} Y_n &= X_n + \sum_{k=0}^{2^n-1} (\Delta_k^n)^2, \\ Y_{n+1} &= X_{n+1} + \sum_{k=0}^{2^{n+1}-1} (\Delta_k^{n+1})^2 \end{aligned}$$

  and this implies that

$$Y_{n+1} - Y_n = X_{n+1} - X_n + \sum_{k=0}^{2^{n+1}-1} (\Delta_k^{n+1})^2 - \sum_{k=0}^{2^n-1} (\Delta_k^n)^2,$$

  and now it is easy to show that

$$\sum_{k=0}^{2^{n+1}-1} (\Delta_k^{n+1})^2 - \sum_{k=0}^{2^n-1} (\Delta_k^n)^2 = 2(X_n - X_{n+1}),$$

  so we conclude that

$$Y_{n+1} - Y_n = X_n - X_{n+1},$$

  for all $n \in \mathbb{N}$.

- From the last equality above, we have that $(Y_n)_{n\geq 0}$ is a Cauchy sequence in $L^2(\Omega)$, because $\|Y_{n+1} - Y_n\| = \|X_{n+1} - X_n\|$, so we can find $Y_\infty \in L^2(\Omega)$ that is the limit of the sequence.

- Now, we have that for all $n \in \mathbb{N}^{\geq 0}$,

$$Y_n = Y_0 + \sum_{j=1}^{n}(Y_j - Y_{j-1}) = Y_0 - \sum_{j=1}^{n}(X_j - X_{j-1}) = Y_0 + X_0 - X_n.$$

  Now, if we take the limit as $n \uparrow \infty$, we obtain that

$$X_\infty + Y_\infty = X_0 + Y_0 = B_1^2,$$

  and the last equality holds true by trivial substitution.

□

Remark 43. We have obtained in particular that $X_\infty \neq Y_\infty$.

Now let's see what is the behaviour of Z.

Proposition 11.3. *The sequence $(Z_n)_n$ converges to $\frac{1}{2}B_1^2$.*

Proof. We just need to observe that for every n,

$$Z_n = \frac{Y_n + X_n}{2} = \frac{X_0 + Y_0}{2} = \frac{1}{2}B_1^2,$$

so it trivially converges to the limit that we claimed. □

11.2 Definition of Stochastic Integral (for E.P.)

We want the uniqueness of the limit, so we have to operate some "choice" in the definition of the integral wrt (with respect to) a Bm.

Setting 2. We put ourselves in the following setting.

* Let us have $B = (\Omega, \mathscr{F}, \mathbb{P}, (\mathscr{F}_t)_{t\geq 0})$ a Bm, with $(\mathscr{F}_t)_{t\geq 0}$ a generic filtration.
* Let $X : (\Omega, \mathscr{F}) \to (\mathbb{R}^{[0,+\infty)}, \mathfrak{B}(\mathbb{R})^{\bigotimes[0,+\infty)})$ be a $S.P.$

Definition 64 (Elementary Process)**.** We say that $X = (X_t)_{t\geq 0}$ is an Elementary Process ($E.P.$) if

- X is adapted with respect to $(\mathscr{F}_t)_{t\geq 0}$,
- There exist
  - $0 = t_1 < t_2 < ... < t_n$ a sequence of real numbers, with $n \geq 2$,
  - $X_{t_1}, ..., X_{t_n}$ a sequence of $r.v.$ with $X_{t_i} \in \mathfrak{M}((\Omega, \mathscr{F}_{t_i}), (\mathbb{R}, \mathfrak{B}(\mathbb{R})),$

such that for all $t \geq 0$, we have

$$X_t = \sum_{i=1}^{n-1} X_{t_i} I_{[t_i, t_{i+1})}(t).$$

Remark 44. We observe that the sequence of real numbers have to be fixed for *every* $\omega \in \Omega$, that is this sequence is independent of ω.

Definition 65 (Square Integrable E.P.)**.** Given X a $E.P.$ according to Definition (64), we say that X is square integrable if for all $i = 1, ..., n$, we have that

$$\mathbb{E}[X_{t_i}^2] < +\infty$$

Remark 45. Let us have $a \in \mathbb{R}$, and X an $E.P.$ Starting by this, we can represent X by adding the point a. In fact,

- if $a \in \{t_1, .., t_n\}$ there is nothing to do.
- Otherwise, we have that $a \in (t_k, t_{k+1})$, for some index k. In this case we can write

  $$X_t = \sum_{i=1}^{n-1} X_{t_i} I_{[t_i, t_{i+1})}(t) = \sum_{i \neq k} X_{t_i} I_{[t_i, t_{i+1})}(t) + \underbrace{X_{t_k} I_{[t_k, t_{k+1})}}_{(A)}.$$

  Now, we can break (A) in the following way,

  $$(A) = X_{t_k} I_{[t_k, a)} + X_{t_k} I_{[a, t_{k+1})}.$$

  Since $\mathscr{F}_{t_k} \subseteq \mathscr{F}_a$, we have that X_{t_k} is $\mathscr{F}_a - measurable$, so if we reorganize the point and the $r.v.$ we obtain a new representation of X with the set of points $\{t_1, .., t_n\} \cup \{a\}$. We remark that the $r.v.$ that correspond to a is X_{t_k}.
- if $a > t_n$, we simply observe that

  $$X_t = X_t + 0 \cdot I_{[t_n, a)}.$$

Definition 66 (Stochastic Integral for E.P.)**.** Let us have

- X an elementary process as in Definition (64).
- $0 \leq a \leq b$ two real numbers.

We define the Stochastic Integral ($S.I.$) of X (with respect to the fixed Bm of Setting (2)) as

$$\int_a^b X_s dB_s := \sum_{i=1}^{n-1} X_{t_i} (B_{(a \vee t_{i+1}) \wedge b} - B_{(a \vee t_i) \wedge b}). \tag{17}$$

11.2.1 Good Definition of SI (To improve the notation of this subsubsection)

In this moment, we don't know if we have given a good definition for our $S.I.$ of $E.P.$ because different representation could give different values of Formula (17). Given X an $E.P.$ such that

$$X_t = \sum_{i=1}^{n-1} X_{t_i} I_{[t_i, t_{i+1})}(t),$$

we denote in this little subsubsection as $S(\overline{X}, \overline{t})$ the $r.v.$ $\int_a^b X_s dB_s$, with $\overline{X} = (X_{t_1}, .., X_{t_n})$ and $\overline{t} = (t_1, .., t_n)$, and we denote as $X(\overline{X}, \overline{t})$ the $E.P.$, because we want to highlight the random variables and the partition. We can do even more, we can define $S(\overline{X}, \overline{t})$ even if $t_1 > 0$. This is useful to simplify the notation (that is already very heavy).

We have to prove that

$$X_t = \sum_{i=1}^{n-1} X_{t_i} I_{[t_i, t_{i+1})}(t) = \sum_{i=1}^{n-1} Y_{p_i} I_{[p_i, p_{i+1})}(t) \implies S(\overline{X}, \overline{t}) = S(\overline{Y}, \overline{p}).$$

Our strategy is this. We prove that the integrals does not change if we add one point, then we use a lemma that we are going to prove.

Lemma 11.4 (Equality Lemma)**.** *Let $X = X(\overline{X}, \overline{t}) = (\overline{Y}, \overline{p})$ be an E.P. Let us suppose that $\overline{t} = \overline{p}$. Then $\overline{X} = \overline{Y}$, that is for all k, $X_{t_k} = Y_{t_k}$.*

Proof. We have that for all $k = 1, .., n-1$,

$$X_{t_k} = \sum_{i=1}^{n-1} X_{t_i} I_{[t_i, t_{i+1})}(t_k) = \sum_{i=1}^{n-1} Y_{t_i} I_{[t_i, t_{i+1})}(t_k) = Y_{t_k}.$$

□

Lemma 11.5 (Addition Lemma)**.** *Let us have the following setting.*

- *Let $X = X(\overline{X}, \overline{t})$ be an E.P. with $\overline{t} = (t_1, .., t_n)$ and $\overline{X} = (X_{t_1}, .., X_{t_{n-1}})$ and $n \geq 2$.*
- *Let $c \in \mathbb{R}$, with $c \geq 0$, and $c \neq t_1, .., t_n$ and $c < t_n$.*
- *Let us consider m the highest index such that $t_m < c < t_{m+1}$.*
- *Let us consider $\hat{t} = (t_1, .., t_m, c, t_{m+1}, .., t_n)$ and $\hat{X} = (X_{t_1}, .., X_{t_m}, X_{t_m}, X_{t_{m+1}}, .., X_{t_n})$.*

Then $S(\overline{X}, \overline{t}) = S(\hat{X}, \hat{t})$.

Proof. We have that

$$\begin{aligned} S(\overline{X}, \overline{t}) = & S((X_{t_1}, .., X_{t_{m-1}}), (t_1, .., t_m)) + \\ & S(X_{t_m}, (t_m, t_{m+1})) + \\ & S((X_{m+1}, .., X_{t_n}), (t_{m+1}, .., t_n)). \end{aligned}$$

Now it is immediate (trivial substitution) to show that

$$S(X_{t_m},(t_m,t_{m+1})) = S(X_{t_m},(t_m,c)) + S(X_{t_m},(c,t_{m+1})) = S((X_{t_m},X_{t_m},(t_m,c,t_{m+1})).$$

Now, if we substitute we obtain

$$\begin{aligned} S(\overline{X},\overline{t}) =& S((X_{t_1},..,X_{t_{m-1}}),(t_1,..,t_m))+ \\ & S((X_{t_m},X_{t_m},(t_m,c,t_{m+1}))+ \\ & S((X_{t_{m+1}},..,X_{t_n}),(t_{m+1},..,t_n)) = \\ =& S(\hat{X},\hat{t}), \end{aligned}$$

and this is the thesis. □

Remark 46. It is immediate that, if we have $c > t_n$, then

$$S((X_{t_1},..,X_{t_{n-1}}),(t_1,..,t_n)) = S((X_{t_1},..,X_{t_{n-1}},0),(t_1,..,t_n,c)).$$

From the above lemmas it follows immediately that

Theorem 11.6. *Formula (17) is well defined, that is it does not depend from the representation.*

Proof. Let us have $X = X(\overline{X},\overline{t}) = X(\overline{Y},\overline{p})$ our usual *E.P.* with two representations. From now till the end of the proof, every union must be intended as an ordered union (for example, $(1,3,4)\cup(2,3) = (1,2,3,4)$). We have that

$$X = X(\overline{X},\overline{t}) = X(\overline{X}_1,\overline{t}\cup\{p_1\}) = ... = X(\hat{X},\overline{t}\cup\overline{p}).$$

but it is also true that

$$X = X(\overline{Y},\overline{p}) = X(\overline{Y}_1,\overline{p}\cup\{t_1\}) = ... = X(\hat{X},\overline{p}\cup\overline{t}),$$

so thanks to Lemma (11.4), we have that $\hat{X} = \hat{Y}$. Now, it is straightforward that

$$S(\overline{X},\overline{t}) = S(\hat{X},\overline{p}\cup\overline{t}) = S(\hat{Y},\overline{p}\cup\overline{t}) = S(\overline{Y},\overline{p}).$$

□

11.3 Property of S.I.

Remark 47. What does the above definition mean? We simply have the following cases. Let us fix an index k.

1. $t_k < t_{k+1} \le a \le b$. Then

$$\begin{aligned} (a \overset{max}{\vee} t_{k+1}) \overset{min}{\wedge} b = a \wedge b = a, \\ (a \overset{max}{\vee} t_k) \overset{\min}{\wedge} b = a \wedge b = a. \end{aligned}$$

   So we have to compute $B_a - B_a$, that is equal to 0.

2. $a \leq b \leq t_k < t_{k+1}$. Then

$$(a \overset{max}{\vee} t_{k+1}) \overset{min}{\wedge} b = t_{k+1} \wedge b = b,$$
$$(a \overset{max}{\vee} t_k) \overset{\min}{\wedge} b = t_k \wedge b = b.$$

So we have to compute $B_b - B_b$, that is equal to 0.

3. $t_k < a < t_{k+1} \leq b$. Then

$$(a \overset{max}{\vee} t_{k+1}) \overset{min}{\wedge} b = t_{k+1} \wedge b = t_{k+1},$$
$$(a \overset{max}{\vee} t_k) \overset{\min}{\wedge} b = a \wedge b = a.$$

So we have to compute $B_{t_{k+1}} - B_a$.

4. $a \leq t_k < b < t_{k+1}$. Then

$$(a \overset{max}{\vee} t_{k+1}) \overset{min}{\wedge} b = t_{k+1} \wedge b = b,$$
$$(a \overset{max}{\vee} t_k) \overset{\min}{\wedge} b = t_k \wedge b = t_k.$$

So we have to compute $B_b - B_{t_k}$.

5. $a \leq t_k < t_{k+1} \leq b$. Then

$$(a \overset{max}{\vee} t_{k+1}) \overset{min}{\wedge} b = t_{k+1} \wedge b = t_{k+1},$$
$$(a \overset{max}{\vee} t_k) \overset{\min}{\wedge} b = t_k \wedge b = t_k.$$

So we have to compute $B_{t_{k+1}} - B_{t_k}$.

We observe that the significant points are those in $[a, b]$.

Remark 48. We want to simplify Formula (17), by eliminating case 3 and 4 in the above remark. We proceed in this way.

- We have X our $E.P.$ We have associated to it a sequence $0 = t_0 < t_1 < ... < t_n$ of points and a sequence $X_{t_1}, .., X_{t_n}$ of r.v.
- We add by following Remark (45) the numbers a and b to the sequence $\{t_i\}$. Now Formula (17) that defines $S.I.$ become

$$\int_a^b X_s dB_s = \sum_{i: a \leq t_i \leq t_{i+1} \leq b} X_{t_i}(B_{t_{i+1}} - B_{t_i}). \tag{18}$$

We observe that one of the point is a, and another one is b.

Lemma 11.7 (Splitting Formula)**.** *Let X be an $E.P.$ and let us consider the $S.I.$ $\int_a^b X_s dBs$. Let us have $c \in \mathbb{R}$ such that $a \leq c \leq b$. Then*

$$\int_a^b X_s dB_s = \int_a^c X_s dB_s + \int_c^b X_s dB_s.$$

Proof. The proof is easy and it's done by adding to the partition the points a and b and c. ☐

Remark 49. We can define

$$J_{a,b} \coloneqq \{i \in \{1,..,n\} \ : \ a \le t_i \le t_{i+1} \le b\},$$

and in this way, we have

$$\int_a^b X_s dB_s = \sum_{i \in J_{a,b}} X_{t_i}(B_{t_{i+1}} - B_{t_i}).$$

Lemma 11.8 (Linearity)**.** *Let X and Y be E.P. and let $0 \le a \le b$ and let $\lambda \in \mathbb{R}$ Then*

$$\int_a^b (X_s + Y_s) dB_s = \int_a^b X_s dB_s + \int_a^b Y_s dB_s,$$
$$\int_a^b (\lambda X_s) dB_s = \lambda(\int_a^b X_s dB_s)$$

Proof. The proof is immediate if we represent X and Y with the same partition, and this is always possible, so we do not write every detail. ☐

11.3.1 Ito isometry for E.P.

Proposition 11.9 (Ito Isometry for E.P.)**.** *Let $X = (X_t)_{t \ge 0}$ be an square integrable process. Then the following facts hold true.*

1. *The r.v. $\int_a^b X_s dB_s \in L^2(\Omega)$, that is it has first and second moment finite.*
2. $$\mathbb{E}\left[\int_a^b X_s dB_s\right] = 0.$$
3. $$\mathbb{E}\left[\left(\int_a^b X_s dB_s\right)^2\right] = \int_a^b \mathbb{E}[(X_s)^2] ds.$$

Proof. We firstly prove that our $S.I.$ belong to $L^2(\Omega)$, so it has first e second moment well defined.

- We denote for every index i, the variable $\Delta_i = B_{t_{i+1}} - B_{t_i}$. We observe that $\Delta_i \perp \mathscr{F}_{t_i}$ by the properties of our Bm.
- First of all, we compute the following

  $$\left(\int_a^b X_s dB_s\right)^2 = \left(\sum_{i \in J_{a,b}} X_{t_i} \Delta_i\right)^2 = \sum_{i \in J_{a,b}} (X_{t_i})^2 (\Delta_i)^2 + 2 \sum_{i,j \in J_{a,b}, i<j} X_{t_i} \Delta_i X_{t_j} \Delta_j.$$

  We have to prove that every term is integrable that is it belong to $L^1(\Omega)$.

- We remember that, if $X \in L^1(\Omega)$ and $Y \in L^1(\Omega)$ and $X \perp Y$, then $XY \in L^1(\Omega)$.
- So, let us consider $i \in J_{a,b}$.

  Thanks to our hypothesis we have that $X_{t_i} \in L^2(\Omega)$ for every i. But then

$$
\begin{aligned}
&X_{t_i} \in L^2(\Omega) \implies (X_{t_i})^2 \in L^1(\Omega), \\
&\Delta_i \in L^2(\Omega) \implies (\Delta_i)^2 \in L^1(\Omega) \\
&X_{t_i} \perp \Delta_i \implies (X_{t_i})^2 \perp (\Delta_i)^2,
\end{aligned}
$$

  where the last implication holds true because X_{t_i} is $\mathscr{F}_{t_i} - measurable$. Now we can conclude that $(X_{t_i})^2(\Delta_i)^2 \in L^1(\Omega)$.
- On the other and, let us consider $i, j \in J_{a,b}$, with $i < j$. So we have $t_i < t_{i+1} \leq t_j$, and we have

$$
\begin{aligned}
&X_{t_i} \;\; is \;\; \mathscr{F}_{t_i} \subseteq \mathscr{F}_{t_j} - measurable, \\
&\Delta_i = B_{t_{i+1}} - B_{t_i} \;\; \mathscr{F}_{t_{i+1}} \subseteq \mathscr{F}_{t_j} - measurable.
\end{aligned}
$$

  So the product $\Delta_i X_{t_i}$ is $\mathscr{F}_{t_j} - measurable$, and we proved some lines above that $\Delta_i X_{t_i} \in L^2(\Omega)$.
- Now, we have that

$$
\begin{aligned}
&X_{t_i}\Delta_i \in L^2(\Omega), \;\; and \;\; it \;\; is \;\; \mathscr{F}_{t_j} - measurable, \\
&X_{t_j} \in L^2(\Omega), \;\; and \;\; it \;\; is \;\; \mathscr{F}_{t_j} - measurable.
\end{aligned}
$$

  Then we can conclude that $X_{t_j} X_{t_i} \Delta_i \in L^1(\Omega)$, and it is $\mathscr{F}_{t_j} - measurable$. (It is well known that $X, Y \in L^2 \implies XY \in L^1$).
- In the end, we have that

$$
\begin{aligned}
&X_{t_j} X_{t_i} \Delta_i \in L^1(\Omega), \;\; and \;\; it \;\; is \;\; \mathscr{F}_{t_j} - measurable, \\
&\Delta_j \in L^1(\Omega), \\
&\Delta_j \perp \mathscr{F}_{t_j}.
\end{aligned}
$$

  So we can conclude that $\Delta_j X_{t_j} X_{t_i} \Delta_i \in L^1(\Omega)$, as we wanted to prove.

Now we know that $\int_a^b X_s dB_s$ has first and second moments finite, so we can compute them.

- $$\mathbb{E}\left[\int_a^b X_s dB_s\right] = \sum_{i \in J_{a,b}} \mathbb{E}[X_{t_i}]\mathbb{E}[B_{t_{i+1}} - B_{t_i}] = 0$$
- For the second moment, we just need to compute the following.

$$
\begin{aligned}
&\mathbb{E}[(X_{t_i})^2(\Delta_i)^2] = \mathbb{E}[X_{t_i}^2]\mathbb{E}[(\Delta_i)^2] = \mathbb{E}[X_{t_i}^2](t_{i+1} - t_i), \\
&\mathbb{E}[X_{t_i} X_{t_j} \Delta_i \Delta_j] = \mathbb{E}[X_{t_i} X_{t_j} \Delta_i] \underbrace{\mathbb{E}[\Delta_j]}_{=0} = 0.
\end{aligned}
$$

So we have that

$$\mathbb{E}\left[\left(\int_a^b X_s dB_s\right)^2\right] = \sum_{i\in J_{a,b}} \mathbb{E}[(X_{t_i})^2](t_{i+1}-t_i) =$$
$$= \sum_{i\in J_{a,b}} \int_a^b \mathbb{E}[(X_{t_i})^2] I_{[t_i,t_{i+1})}(s)ds.$$

Now, we observe that for all $s \in [t_i, t_{i+1})$, we have

$$\mathbb{E}[(X_s)^2] = \mathbb{E}[(X_{t_i})^2],$$

so if we substitute above we obtain

$$\sum_{i\in J_{a,b}} \int_a^b \mathbb{E}[(X_s)^2] I_{[t_i,t_{i+1})}(s)ds = \int_a^b \sum_{i\in J_{a,b}} \mathbb{E}[(X_s)^2] I_{[t_i,t_{i+1})}(s)ds = \int_a^b \mathbb{E}[(X_s)^2]ds,$$

and this is the thesis.

□

We can improve the above result.

Proposition 11.10 (Ito Isometry for E.P. Improved). *Let $X = (X_t)_{t\geq 0}$ be a square integrable E.P. The following facts hold true.*

1.

$$\mathbb{E}\left[\int_s^t X_r dB_r | \mathscr{F}_s\right] = 0.$$

2.

$$\mathbb{E}\left[\left(\int_s^t X_r dB_r\right)^2 | \mathscr{F}_s\right] = \mathbb{E}\left[\int_s^t (X_r)^2 dr | \mathscr{F}_s\right].$$

Proof. The proof is a little bit tricky with respect to the above one because we do not know that some *r.v.* are measurable with respect $\mathscr{F}_s$, but we can save ourselves.

- Let us denote as $J_{s,t} := \{i : s \leq t_i \leq t_{i+1} \leq t\}$.
- We remember that

$$\int_s^t X_r dBr = \sum_{i\in J_{s,r}} X_{t_i}\Delta_i.$$

  with $\Delta_i = B_{t_{i+1}} - B_{t_i}$.
- We denote $\int_s^t X_r dB_r$ as $\int X$ (this is simply a notation).

- From Ito Isometry, we know that $\int X$ is square integrable, so the conditional expectation is well defined both $(\int X)$ and $(\int X)^2$.
- Let us compute the first *conditional expectatoin*. We have

$$\mathbb{E}\left[\left(\int X\right)|\mathscr{F}_s\right] = \sum_{i\in J_{s,r}} \mathbb{E}[X_{t_i}\Delta_i|\mathscr{F}_s].$$

if every term of the above sum is zero, we conclude. This is true because of the followin trick,

$$\mathbb{E}[X_{t_i}\Delta_i|\mathscr{F}_s] \underbrace{=}_{tower} \mathbb{E}[\ \mathbb{E}[X_{t_i}\Delta_i|\mathscr{F}_{t_i}]\ |\mathscr{F}_s] = \mathbb{E}[X_{t_i}\mathbb{E}[\Delta_i|\mathscr{F}_{t_i}]|\mathscr{F}_s] \underbrace{=}_{\Delta_i\perp\mathscr{F}_{t_i}} \mathbb{E}[X_{t_i}\underbrace{\mathbb{E}[\Delta_i]}_{=0}|\mathscr{F}_s] = 0.$$

We have use moreover thar X_{t_i} is $\mathscr{F}_{t_i} - measurable$, and that for all i, we have that $\mathscr{F}_s \subseteq \mathscr{F}_{t_i}$.

- Now let us compute the second conditional expectation. We have

$$\mathbb{E}[(\int X)^2|\mathscr{F}_s] = \sum_{i\in J_{s,t}} \mathbb{E}[(X_{t_i})^2(\Delta_{t_i})^2|\mathscr{F}_s] + 2\sum_{i\in J_{s,t},i<j} \mathbb{E}[X_iX_j\Delta_i\Delta_j|\mathscr{F}_s].$$

- If we can compute the single elements, we complete easily. We have

$$\mathbb{E}[(X_{t_i})^2(\Delta_i)^2|\mathscr{F}_s] = \mathbb{E}[(X_{t_i})^2\mathbb{E}[(\Delta_i)^2]|\mathscr{F}_s] = \mathbb{E}[(X_{t_i})^2|\mathscr{F}_s](t_{i+1}-t_i).$$

We have used the tower property with $\mathscr{F}_{t_i}$, then that $\Delta_i\perp\mathscr{F}_{t_i}$ and in the end that $\mathbb{E}[\Delta_i|\mathscr{F}_{t_i}] = \mathbb{E}[\Delta_i]$.

On the other and, we obtain for $i<j$ that

$$\mathbb{E}[X_iX_j\Delta_i\Delta_j|\mathscr{F}_s] = \mathbb{E}[X_iX_j\Delta_i\underbrace{\mathbb{E}[\Delta_j]}_{=0}|\mathscr{F}_s] = 0,$$

where we have used the tower property wrt $\mathscr{F}_{t_j}$, and that $\Delta_j\perp\mathscr{F}_{t_j}$.

If we put everything together we obtain

$$\begin{aligned}\mathbb{E}[(\int X)^2|\mathscr{F}_s] =&\mathbb{E}[\sum_{i\in J_{s,t}}(X_{t_i})^2(t_{i+1}-t_i)|\mathscr{F}_s] = \mathbb{E}[\sum_{i\in J_{s,t}}\int_a^b (X_{t_i})^2 I_{[t_i,t_{i+1})}(r)dr|\mathscr{F}_s] = \\ =&\mathbb{E}[\int_a^b\sum_{i\in J_{s,t}}(X_{t_i})^2 I_{[t_i,t_{i+1})}(r)dr|\mathscr{F}_s] = \mathbb{E}[\int_a^b (X_r)^2dr|\mathscr{F}_s],\end{aligned}$$

and this conclude. □

Corollary 11.11. *Let $X = (X_t)_{t\geq 0}$ be a square integrable process. Then*

$$M_t := \int_0^t X_s dB_s \text{ and } N_t := M_t^2 - \int_0^t X_s^2 ds$$

are martingle wrt (with respect to) the Brownian Filtration $(\mathscr{F}_t)_{t\geq}$ in Setting (2).

Proof. We firstly consider M_t, then N_t but it is all really easy. We begin remembering that

$$\forall t \geq 0, \ \ X_t = \sum_{i=1}^{n} X_{t_i} I_{[t_i,t_{i+1})}(t)$$

for some $n \geq 2$, and X_{t_i} a $\mathscr{F}_{t_i} - measurable$ r.v.

- We have for every $t \geq 0$ that

$$M_t = \sum_{i\in J_{0,t}} X_{t_i}(B_{t_{i+1}} - B_{t_i}),$$

  so adaptness is immediate.

- We have thanks to Proposition (11.9) (Ito Isometry) that M_t is integrable for every t.

- It remain to prove the martingles property, but this is just an immediate application of Proposition (11.10) and the splitting formula (11.7), because we have

$$\mathbb{E}[M_t - M_s|\mathscr{F}_s] = \mathbb{E}[\int_s^t X_r dB_r|\mathscr{F}_s] = 0,$$

  so we have that M_t is a martingle.

Now we prove the same result for N_t.

- It is immediate that N_t is adapted wrt $(\mathscr{F}_t)_t$ (we can write explicitly who is $\int_0^t (X_s)^2 ds$).

- Thanks to (11.9), we have that M_t^2 is integrable for every t, so we just need to check that $\int_0^t (X_s)^2 ds$ is integrable. This is true thanks to Fubini' theorem. In fact

$$\mathbb{E}[\int_0^t X_s^2 ds] = \int_0^t \mathbb{E}[(X_s)^2]ds = \mathbb{E}[(M_t)^2] < +\infty.$$

- Now we have to prove the martingles equality. We have

$$\mathbb{E}[N_t|\mathscr{F}_s] = \mathbb{E}[(M_t)^2 - \int_0^t (X_r)^2 dr|\mathscr{F}_s] =$$
$$= \mathbb{E}[(M_t)^2 - \int_0^s (X_r)^2 dr|\mathscr{F}_r] - \mathbb{E}[\int_s^t (X_r)^2 ds|\mathscr{F}_s] \underbrace{=}_{(A)} .$$

By Proposition (11.10) we have $\mathbb{E}[\int_s^t (X_r)^2 ds|\mathscr{F}_s] = \mathbb{E}[(M_t - M_s)^2|\mathscr{F}_s]$, so if we substitute we obtain

$$\underbrace{=}_{(A)} \mathbb{E}[(M_t)^2 - \int_0^s (X_r)^2 dr|\mathscr{F}_s] - \mathbb{E}[(M_t - M_s)^2|\mathscr{F}_s]$$
$$= -(M_s)^2 - \int_0^s X_r dr + \mathbb{E}[2M_t M_s|\mathscr{F}_s] \underbrace{=}_{(B)}$$

Now we have proved some lines above that M_t is a martingle, so $\mathbb{E}[M_t|\mathscr{F}_s] = M_s$, and if we substitue we obtain

$$\underbrace{=}_{(B)} -(M_s)^2 - \int_0^s X_r dr + 2(M_s)^2 = (M_s)^2 - \int_0^s X_r dr = N_s$$

So we have the thesis.

□

11.4 Ito Integrals

- Let $B = (B_t)_{t\geq 0} = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ be a Bm.
- Let us consider $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$.
- Let $X : [0, +\infty) \times \Omega \to \mathbb{R}$ be a $S.P.$

Definition 67. We denote as $M_B^2(a, b)$ the set of $S.P.$ $X = (X_t)_{t\geq 0}$ such that

1. X is progressively measurable in $[a, b]$. (Definition 36).
2. $\mathbb{E}[\int_a^b (X_t)^2 dt] < +\infty$.

Remark 50. We observe that the hypotheses of Fubini's theorem hold true, in fact

- X *progressively measurable* $\implies X|_{[a,b]\times\Omega}$ is $\mathfrak{B}([a, b]) \otimes \mathscr{F}_b - measurable.$
- $(X)^2$ is always positive.

So we can write that

$$\mathbb{E}[\int_a^b (X_t)^2 dt] = \int_{[a,b]\times\Omega} (X)^2 d(\mathcal{L} \otimes \mathbb{P}) = \int_a^b \underbrace{\mathbb{E}[(X_t)^2]}_{\|(X_t)^2\|_{L^2}^2} dt,$$

where $\mathcal{L}$ denote the *Lebesgue* measure of $[a, b]$.

More generally, the following proposition holds.

- Let us have $(E, \mathfrak{B}(E))$ a topological space (but in my opinion metric is more understandable).

- Let us have $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ an $(E, \mathfrak{B}(E)) - S.P.$

Proposition 11.12. *If*

- *X is adapted (wrt $(\mathscr{F}_t)_{t\geq 0}$),*
- *X is right (or left) continuous (everywhere, not only a.c.).*

Then X is progressively measurable (Definition (36)).

Remark 51. We have that

$$\{square\ integrable\ E.P.\} \subseteq M_B^2(a,b).$$

In fact $E.P.$ are right continuous everywhere (wrt the $Brownian - Filtration$).

Remark 52. We can define on $M_B^2(a,b)$ the following scalar product,

$$\langle X, Y \rangle_{M^2} := \mathbb{E}[\int_a^b X_s Y_s ds].$$

Remark 53. We denote as $\|X\|_{M_2} := (\langle X, X \rangle)^{\frac{1}{2}}$, that is

$$\|X\|_{M_2}^2 = \mathbb{E}[\int_a^b (X_s)^2 ds].$$

In this exact moment, THIS IS NOT A NORM because we have not identified two r.v. that are equal $a.c.$

Remark 54. We could make the identification with the following equivalence relation. For all $X \in M_B^2(a,b)$ and $Y \in M_B^2(a,b)$,

$$X \equiv Y \iff \|X - Y\|_{M^2} = \mathbb{E}[\int_a^b (X_s - Y_s)^2 ds].$$

Though it would be nice, we don't follow this convention, every element is in relation just with itself.

Theorem 11.13 (Approximation Theorem in M_B^2)**.** *Let $(X_t)_{t\geq 0} \in M_B^2(a,b)$ be a S.P. Then*

- *There is $(X^{(n)})_{n\geq 1}$ a sequence of E.P. such that*
  - *for all n, we have that $X^{(n)} \in M_B^2(a,b)$,*
  - $\left\| X - X^{(n)} \right\|_{M_B^2} \to 0$ *if* $n \uparrow +\infty$.
- *We can find even a sequence of continuous processes that converges to X in $M_B^2(a,b)$.*

Proof. We do not prove this theorem (so this is not a proof). □

Remark 55. More general, if we have a sequence $(X^{(n)})_{n\in\mathbb{N}}$ of elements of $M_B^2(a,b)$ that converges to an element of $X \in M_B^2(a,b)$, we write that

$$X^{(n)} \underset{M_B^2(a,b)}{\overset{n\to+\infty}{\to}} X.$$

We do not write $n \to +\infty$ or the space $M_B^2(a,b)$ if it is clear what we are saying.

Now, we introduce this notation to simplify the following theorems. We remember that we are always in Setting (2), so if we speak about Ω and $\mathscr{F}$ we know what we are referring to.

- We denote as EP the set of *all* the elementary processes.
- Now, let us fix $0 \le a \le b$.
- We define the following function,

$$\varphi : EP \to L^2(\Omega, \mathscr{F}, \mathbb{P}), \;\; \varphi(X) := \int_a^b X_s dB_s,$$

  that is we associate to every EP its SI, that we remember briefly it is the following

$$X_t = \sum_{i=1}^{n} X_{t_i} I_{[t_i,t_{i+1})}(t) \implies \varphi(X) = \sum_{i\in J_{a,b}} X_{t_i}(B_{t_{i+1}} - B_{t_i})$$

  For more detail, see *Definition* (66). We note that, since Linearity of SI (Lemma (11.8), we have that $\varphi(X+Y) = \varphi(X) + \varphi(Y)$.
- We observe that Ito Isometry (Proposition (11.9)) become, thanks also to Remark (50),

$$\mathbb{E}[(\varphi(X))^2] = \left\|\varphi(X)\right\|_{L^2}^2 = \|X\|_{M^2}^2 = \int_a^b \|X_s\|_{L^2}^2 \, ds.$$

Now we are ready to enunciate and prove the theorem that also defines the *Ito Integral*.

Proposition 11.14. *Let us have*

- $(X^{(n)})_n$ *a sequence of E.P. in* $M_B^2(a,b)$*,*
- $X \in M_B^2(a,b)$*,*

and let us suppose that $X^{(n)} \to X$ *in* $M_B^2(a,b)$*. Then*

1. *the sequence* $(\varphi(X^{(n)}))_n$ *is Cauchy in* $L^2(\Omega, \mathscr{F}, \mathbb{P})$*.*
2. *If* $(Y^{(n)})_n$ *is another sequence of E.P. such that* $Y^{(n)} \to X$*, then*

$$\lim_n \varphi(X^{(n)}) = \lim_n \varphi(Y^{(n)}).$$

Proof. We prove first point 1.and then point 2.

1. We have for all $m \in \mathbb{N}$ and $n \in \mathbb{N}$ that

$$\left\|\varphi\left(X^{(n)}\right) - \varphi\left(X^{(m)}\right)\right\|_{L^2} = \left\|\varphi\left(X^{(n)} - X^{(m)}\right)\right\|_{L^2} = \left\|X^{(n)} - X^{(m)}\right\|_{M_B^2}$$

and this implies immediately that $(\varphi(X^{(n)}))$ is Cauchy in $L^2(\Omega, \mathscr{F}, \mathbb{P})$.

2. We have, for all $n \in \mathbb{N}$ that

$$\left\|\varphi\left(X^{(n)}\right) - \varphi\left(Y^{(n)}\right)\right\|_{L^2} = \left\|X^{(n)} - Y^{(n)}\right\|_{M_B^2} \leq \left\|X^{(n)} - X\right\|_{M_B^2} + \left\|Y^{(n)} - X\right\|_{M_B^2},$$

and this permit us to conclude.

□

Remark 56. So, given the proposition above, we can extend φ to a new function $\tilde{\varphi}$ defined in this way,

$$\tilde{\varphi} : M_B^2(a, b) \to L^2(\Omega, \mathscr{F}, \mathbb{P}), \ \ \tilde{\varphi}(X) := \lim_n \varphi(X^{(n)}),$$

with $(X^{(n)})$ a sequence that converges to X in M_B^2. We observe that $\tilde{\varphi}|_{EP} = \varphi$. From now on, we make no difference between $\tilde{\varphi}$ and φ.

Definition 68 (Stochastic Integral in M_B^2)**.** Given $X \in M_2^B$, we define $\varphi(X)$ the *Stochastic Integral* (*S.I.*) of X (wrt the fixed *Bm* in Setting (2)).

Remark 57. Given $X \in M_2$, we have that X is a measurable function, but $\varphi(X)$ is a *class of equivalence* of measurable function, so expression like $\varphi(X)(\omega)$ make no sense.

11.5 Property of SI

Theorem 11.15 (Ito Isometry in M_B^2(a,b))**.** *Let us have $X \in M_B^2$. Then*

- *If $T > 0$, we have*

$$\mathbb{E}[\int_0^T X_s dB_s] = 0,$$
$$\mathbb{E}[(\int_0^T X_s dB_s)^2] = \mathbb{E}[\int_0^T (X_s)^2 ds].$$

- *More generally, if $0 \leq s \leq t \leq T$, we have*

$$\mathbb{E}[\int_s^t X_r dB_r] = 0,$$
$$\mathbb{E}[(\int_s^t X_s dB_s)^2 | \mathscr{F}_s] = \mathbb{E}[\int_s^t (X_s)^2 ds | \mathscr{F}_s].$$

- *Moreover,*

$$M_t := \int_0^t X_s dB_s, \ \ \textit{and} \ \ N_t := (M_t)^2 - \int_0^t (X_s)^2 ds,$$

are martingles wrt $(\mathscr{F}_t)_{t\geq 0}$.

Proof. Da fare. □

Remark 58. If it is not obvious the interval of integration in $\int_a^b X_s dB_s$, we denote it as $\varphi_{a,b}(X)$.

Remark 59. Let us have $X, Y \in M_B^2(0,T)$. Then by Ito Isometry and the polarization formula (cool name)

$$(X+Y)^2 - (X-Y)^2 = 4XY$$

we obtain

$$\begin{aligned} 4\langle X, Y\rangle_{M_B^2} &= 4\mathbb{E}[\int_0^T X_s Y_s] = \\ &= \|X+Y\|^2_{M_B^2} - \|X-Y\|^2_{M_B^2} = \\ &= \left\|\varphi(X+Y)\right\|_{L^2} - \left\|\varphi(X-Y)\right\|_{L^2} = \\ &= 4\langle\varphi(X), \varphi(Y)\rangle_{L^2}. \end{aligned}$$

So the Ito Isometry conserves the scalar products.

11.5.1 Continuous Version of S.I.

- We have the following family of class of equivalence $\{\int_0^t X_s dB_s, \ t \in [0,T]\}$.
- We would like to find a *S.P.* $M = (M_t)_{t\geq 0}$ that has the following properties,
  1. For every $t \geq 0$, we have that $M_t \in \int_0^t X_s dB_s$.
  2. M is *a.c.* continuous.

One day we find it. Now We give this result without proof.

11.6 More general class of processes which we want to integrate

Let us have $B = (\Omega, \mathscr{F}, \mathbb{P}, (\mathscr{F}_t)_{t\geq 0})$. Let us have $0 \leq a \leq b$, and let us have $(X_t)_{t\geq 0}$ a *S.P.*

Definition 69. We say that a $X \in \Lambda_B^2(a,b)$ if

- X is progressively measurable,
- $\mathbb{P}(\{\omega : \int_a^b X_s(\omega) ds < +\infty\}) = 1$.

Definition 70 (Convergence in $\Lambda_B^2(a,b)$)**.** Let us suppose to have the following setting.

- Let us have $(X^{(n)})_n$ a sequence of $S.P.$ in $\Lambda^2_B(a,b)$.
- Let us have $X \in \Lambda^2_B(a,b)$.
  We have that $X^{(n)} \to X$ in $\Lambda^2_B(a,b)$ if
  $$\mathbb{P}\left(\left\{\omega : \lim_n \left\|X^{(n)}(\omega) - X(\omega)\right\|_{L^2(a,b)}\right\}\right) = 1.$$
  with
  $$\left\|X^{(n)}(\omega) - X(\omega)\right\|_{L^2(a,b)} = \int_a^b (X_s^{(n)}(\omega) - X_s(\omega))^2 ds.$$

Remark 60. **WE DO NOT CONSIDER $\Lambda^2_B(a,b)$ AS A CLASS OF EQUIVALENCE CLASSES, EVEN THOUGH WE COULD DO SO.** In this case the relation would be
$$X \equiv Y \ \ in\ \ \Lambda^2_B(a,b) \iff \int_a^b (X_s(\omega) - Y_s(\omega))^2 < +\infty \ \ a.s.\ \ \omega \in \Omega.$$

Theorem 11.16 (Approximation Theorem). *Let us have $X = (X_t)_{t\geq 0} \in \Lambda^2_B(a,b)$. Then*

- *We can find $(X^{(n)})_n$ a sequence of E.P. such that $X^{(n)} \to X$ in $\Lambda^2_B(a,b)$.*
- *We can find a sequence of continuous processes tha converges to X in $\Lambda^2_B(a,b)$.*

Proof. It is just a technical lemma, so we do not prove this. □

Now we want to prove a crucial lemma that we use to define the S.I. of a function in $\Lambda^2_B(a,b)$.

Lemma 11.17. *Let X be an E.P. Then for every $\epsilon > 0$ and $\rho > 0$, we have*
$$\mathbb{P}\left(|\int_a^b X_s dB_s| \geq \epsilon\right) \leq \mathbb{P}\left(\int_a^b X_s^2 ds \geq \rho\right) + \frac{\rho}{\epsilon^2}.$$

Proof. The proof follow the following steps.

- Let us define
  $$A := \{\int_a^b X_s^2 ds \geq \rho\}, \ \ so\ that\ \ A^C = \{\int_a^b X_s^2 ds < \rho\}.$$
  Moreover, let us set
  $$B := \{|\int_a^b X_s dB_s| \geq \epsilon\}.$$
- We have the simple estimate
  $$\mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^C) \leq \mathbb{P}(A) + \mathbb{P}(B \cap A^C).$$
  If we prove that $\mathbb{P}(B \cap A^C) \leq \frac{\rho}{\epsilon^2}$, we finish.
-

□