Stochastic Calculus

an introduction
to option pricing with martingales

Christophe Giraud[1]

Lecture notes IMAMIS
Manila, august 2005

[1]University of Nice (France). E-mail: cgiraud@math.unice.fr

Contents

Abbreviations and notations

$\mathbb{E}(X)$	expectation of X
$X \perp Y$	X is independent of Y
i.e.	namely
i.i.d.	independent and identically distributed
a.s.	almost surely (with probability 1)
$\mathbb{P}^*$	risk neutral probability
iif	if and only if
s.t.	stopping time
r.v.	random variable

Chapter 0

Goal of the lecture

The goal: to give a comprehensive introduction to option pricing with martingale. The first two chapters give some mathematical stuff (conditional expectation, martingale). Chapter 3,4,5 focus on the completely discrete setting. We use the risk neutral probability approach to compute the price of european and american options.

0.1 Risk hedging - options

0.1.1 Problematic

Some financial or industrial agents may wish to eliminate some risks, either as a commercial choice or because it does not enter into their fields of competence.

Examples:

1. A european company manages its activity in euros, but signs its contracts in dollar, payable on receipt. Between today and the receipt, the euro/dollar exchange rate may fluctuate. The company thus have to face an exchange risk. If it does not want to take charge of it, the company will sign a contract that protects itself against this risk.

2. The market price of copper fluctuate dramatically. A copper mine may wish to protect itself against this fluctuation. The manager of the mine will then sign a contract, that warrants him a minimum price for its copper.

0.1.2 Some examples of options

European call : contract that gives the right (but not the obligation) to its owner to buy an asset at the fixed price K (strike) at time N (maturity). This contract has a price C (prime).

We write S_n for the price of the asset at time n. Two cases may occurs at maturity:

- when $S_N < K$: the owner of the option has the right to buy at price K an asset that he could buy for less in the market. This right is not interesting. He does not exercise it : nothing happens.
- when $S_N \geq K$: the owner of the call can buy an asset for less than the market price, which is interesting. The seller of the option, must then buy an action at price S_N and sell it at price K to the owner of the option. Things goes as if the seller was giving $S_N - K$ to the owner.

In short: at time $n = 0$ the owner gives C to the seller of the call. At time N, he receives the maximum between $S_N - K$ and 0, we denote by $(S_N - K)_+$. We call *payoff*, the function $f = (S_N - K)_+$.

Figure 1: Profit/loss of the owner of the call

European put: contract that gives the right (but not the obligation) to its owner to sell an asset at the fixed price K (strike) at time N (maturity). This contract has a price C (prime). In this case the payoff is $f = (K - S_N)_+$.

Figure 2: Profit/loss of the owner of the put

American call (resp. put): contract that gives the right (without obligation) to buy (resp. sell) an asset at any time before time N (maturity) at the fixed price K. This contract has a price.

exotic options: there exists many other options, called exotic options. Example given: the collar option, with payoff $f = \min(\max(K_1, S_N), K_2)$, the Boston option, with payoff $f = (S_N - K_1)_+ - (K_2 - K_1)$, where $K_1 < K_2$, etc.

0.1.3 Option pricing - hedging

The determination of the fair price C of an option and the way to hedge it is the main goal of the mathematical finance.

The seller of the option will invest on the market in such a way that, whatever the evolution of the market, he will be able to face its engagement without loosing money. The fair price of the option will then correspond to the minimal initial investment needed to carry out this condition. The investment strategy of the seller is called hedging strategy.

Fundamental remark: the hedging strategy of an option differs fundamentally from the hedging strategy of a classical insurance (against robbery, fire, etc). Indeed, in the present case the seller of the option must hedge the risk of a *single* contract. The risk must disappear. A contrario, the seller of a classical insurance sell many similar contracts. His strategy is then to face his engagement on *average*. He hopes that there will not be two many claims in the same time. Such an hedging strategy is so-called a hedging strategy via diversification.

0.2 Stochastic models

The future is uncertain, but there exit some models to describe its evolution. These model take into account the uncertainty of the future. They are so-called stochastic models (in opposition to deterministic models). A financial market may follow different evolutions, each evolution having a given probability to happen. For example, two basic parameters to describe the evolution of a stock is its trend and its volatility.

A stochastic process is a "value" that evolves randomly when time passes (NASDAC, quote of copper, exchange rates, etc). We will investigate in the next chapters some stochastic processes that appear when we model the financial market. A special attention will be given on martingales and brownian motions.

0.3 "One step, two states" model

The following model is the simplest one: there exists two dates (today and tomorrow) and two possible states for tomorrow. Its analysis uses (in this very simple setting) the methods that we will further develop for more involved cases.

We focus on an option with maturity $N = 1$ and payoff of the form $f = g(S_N)$ (ex: for a european call $g(x) = (x - K)_+$). Today ($n = 0$), the quote of the underlying asset is S_0. Tomorrow ($n = 1$), the quote can take two possible values S_- and S_+. The seller of the option creates a *portfolio* $\pi = (\beta, \gamma)$ made of β units of bond and γ units of asset. Today, the value of the portfolio is $X_0 = \beta + \gamma S_0$. Tomorrow, it will be $X_N = \beta e^{rN} + \gamma S_N$ (where r represents the interest rate of the bond).
To be able to face to its obligation, the value of the seller's portfolio at maturity (time N) must not be smaller than the payoff, namely

$$X_N \geq g(S_N).$$

Assume that either when $S_N = S_+$, or when $S_N = S_-$, the value of the portfolio is larger than the payoff $g(S_N)$. Then the seller could earn money with positive probability and no risk of loosing money. We assume that this is impossible (market at equilibrium). We say in this case that the market is *arbitrage-free.* As a consequence, we must have $X_N = g(S_N)$, which in turn enforces (β, γ) to satisfy the equations

$$\begin{cases} \beta e^{rN} + \gamma S_+ &= g(S_+) \\ \beta e^{rN} + \gamma S_- &= g(S_-). \end{cases}$$

The unique solution to this equation is given by

$$\gamma = \frac{g(S_+) - g(S_-)}{S_+ - S_-} \tag{1}$$

and

$$\beta = \frac{1}{2} e^{-rN} \left(g(S_+) + g(S_-) - \frac{S_+ + S_-}{S_+ - S_-} (g(S_+) - g(S_-)) \right).$$

Formula (1) is commonly called "delta hedging formula". The initial value of the portfolio is $X_0 = \beta + \gamma S_0$ which is positive. This value corresponds to the cost of the contract, namely its fair *price* C.

In short: at time $t = 0$, the owner of the option gives $C = X_0 = \beta + \gamma S_0$ to the seller. At time N, either the asset quotes S_+, and then the seller gives $g(S_+)$ to the owner, either the asset quotes S_-, in which case the seller gives $g(S_-)$ to the owner.

Fundamental Remark: *neither the price C nor the portfolio (β, γ) depend on the probability that the asset S takes value S_+ or S_- !*

why?

Because our hedging strategy works whatever the evolution of the market, not in average. If we want to face our obligation in every case, the probability of rise or fall does not matter. Our hedging must work both when the price rises and when it falls. There is no probabilistic arguments here ($\neq$ average hedging via diversification).

Nevertheless, the price C of the option can be view as the expected value of the payoff under some artificial probability *the risk neutral probability*. This remark is the corner stone of the pricing methods developed below.

Let us focus on a european call. We define the risk neutral probability $\mathbb{P}^*$ by

$$\mathbb{P}^*(S_N = S_+) = \frac{e^{rN}S_0 - S_-}{S_+ - S_-} =: p^* \quad \text{and} \quad \mathbb{P}^*(S_N = S_-) = 1 - p^*.$$

We then have

$$S_0 = \mathbb{E}^*\left(e^{-rN}S_N\right) \quad \text{and} \quad C = \mathbb{E}^*\left(e^{-rN}(S_N - K)_+\right),$$

where $\mathbb{E}^*$ represents the expectation under probability $\mathbb{P}^*$.
The first equality says that the discounted value[1] S_n^* of S, defined by $S_n^* = e^{-rn}S_n$, keeps a constant expected value under $\mathbb{P}^*$. This is the fundamental property of a risk neutral probability: it annihilates the trend of the discounted quote of the stock S. The second equality says that the fair price of the option suits with the expected value of the discounted payoff under $\mathbb{P}^*$. We shall see in the following, that this formula is very general.

Another remark: we have made two major hypotheses in the previous analysis. First, we assume that the market is arbitrage-free, which fits more or less with the reality. Second, we assume to be in an ideal market: no transaction costs, possibility to buy/sell any quantity of asset, possibility to borrow as much money as needed, etc. Such a market is so-called complete. In this case the fair price is unique, which falls without theses hypotheses.

Exercise: compute the price of a european call, when $S_0 = 120$, $K = 80$, $r = 0$ and $S_+ = 180$, $S_- = 60$.

[1] also called "present value"

Chapter 1

Conditional expectation

The goal: to formalize the notion of conditional expectation given some information $\mathcal{I}$.

Informally, the conditional expectation of a random variable (r.v.) X given $\mathcal{I}$ represents the 'average value expected for X when one knows the information $\mathcal{I}$'.

Example: throw two dices. Write X_1 and X_2 for the value of first and second dice, and set $S = X_1 + X_2$. If you have no information, the average value expected for S is

$$\mathbb{E}(S) = \mathbb{E}(X_1) + \mathbb{E}(X_2) = 7.$$

Assume now that you know the value of X_1. Then, the average value expected for S knowing the value of X_1 (our information) is

$$X_1 + \mathbb{E}(X_2) = X_1 + 3.5\,.$$

The latter quantity is what we call the 'conditional expectation of S knowing X_1'.

Notations: Ω will represent the universe of the possibilities, we will write $\mathcal{F}$ for the set of all the possible events ($\mathcal{F}$ is a so-called σ-algebra) and $\mathbb{P}(A)$ for the probability that the event $A \in \mathcal{F}$ occurs.

1.1 Discrete setting

We associate to an event $A \in \mathcal{F}$ and an event $B \in \mathcal{F}$ of positive probability

$$\mathbb{P}(A \mid B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

which represents "the probability that the event A occurs, knowing that the event B has occurred". In the same way, we define the conditional expectation of a real random variable $X : \Omega \to \{x_1, \ldots, x_m\}$ knowing that the event B has occurred by

$$\mathbb{E}\,(X \mid B) := \sum_{i=1}^{m} x_i\, \mathbb{P}(X = x_i \mid B). \tag{1.1}$$

This quantity is a real number that represents "the average value expected for X knowing that the event B has occurred".

- **Example:** throw a dice with 6 faces, and set X for the obtained value. Write B for the event "X is not smaller than 3". Then, the expected value of X knowing B is $\mathbb{E}(X \mid B) = \sum_{i=1}^{6} i\,\mathbb{P}(X = i \mid B)$. Since $\mathbb{P}(B) = 2/3$, we have

$$\begin{aligned} \mathbb{P}(X = i \mid B) &= \frac{3}{2}\,\mathbb{P}(\{X = i\} \cap B) \\ &= \begin{cases} 0 & \text{if } i = 1 \text{ or } 2 \\ 1/4 & \text{else.} \end{cases} \end{aligned}$$

  Finally, $\mathbb{E}(X \mid B) = \frac{1}{4}(3 + 4 + 5 + 6) = 9/2$.

1.1.1 Conditioning against a random variable

We will introduce in this section the notion of conditioning against a random variable. To start with, let's have look to an example.

- **Example:** we play at a game in two stages: first stage we throw a dice and call Z the obtained value. Second stage, we throw Z times the dice and multiply the Z obtained values. We write X for this product.
  Note that we are able to compute for example $\mathbb{E}(X \mid Z = 5)$. If $Z(\omega) = 5$, we throw 5 times the dice, independently, so $\mathbb{E}(X \mid Z = 5) = m^5$, where $m = \frac{1}{6}(1+2+3+4+5+6) = 7/2$ is the expected value for a throw. More generally, for $i \in \{1, \ldots, 6\}$, we have $\mathbb{E}(X \mid Z = i) = m^i$. In this case, we will define the conditional expectation of X knowing Z as

$$\mathbb{E}(X \mid Z) := m^Z.$$

  Then, $\mathbb{E}(X \mid Z)$ is a random variable such that: if i is the value of Z, i.e. if $Z(\omega) = i$, then $\mathbb{E}(X \mid Z)(\omega) = \mathbb{E}(X \mid Z = i)$.

We now generalize this notion to two general random variables $X : \Omega \to \{x_1, \ldots, x_m\}$ and $Z : \Omega \to \{z_1, \ldots, z_n\}$, with $x_1, \ldots, x_m, z_1, \ldots, z_n \in \mathbb{R}$.

Definition 1.1 Conditional expectation of X given Z.
The conditional expectation of X given Z, we note $\mathbb{E}(X \mid Z)$, is a random variable, defined by

$$\mathbb{E}(X \mid Z)(\omega) := h(Z(\omega)),$$

where h is the function defined by $h(z_j) = \mathbb{E}(X \mid Z = z_j)$ for any $j \in \{1, \ldots, n\}$. The quantity $\mathbb{E}(X \mid Z = z_j)$ is the one defined by (1.1) *(see also the following exercise).*

Warning: $\mathbb{E}(X \mid Z = z_j)$ is a real number, but $\mathbb{E}(X \mid Z)$ is a random variable. Indeed, $\mathbb{E}(X \mid Z)(\omega)$ depends on ω, since the value of $Z(\omega)$ depends on ω.

- **Example:** Let Z be a r.v. with uniform law on $\{1, \dots, n\}$ (i.e. $\mathbb{P}(Z = i) = 1/n$ for $i = 1, \dots, n$) and ε be a r.v. independent of Z, such that $\mathbb{P}(\varepsilon = 1) = p$ and $\mathbb{P}(\varepsilon = -1) = 1 - p$.
  We set $X = \varepsilon Z$. Then, X is a r.v. with value in $\{-n, \dots, n\}$. Let us compute $\mathbb{E}\left(X \mid Z\right)$. For $j \in \{1, \dots, n\}$,

$$\mathbb{E}\left(X \mid Z = j\right) = \sum_{i=-n}^{n} i\, \mathbb{P}(X = i \mid Z = j)$$

  and $\mathbb{P}(X = i \mid Z = j) = 0$ when $i \neq j$ or $i \neq -j$ (since $X = \pm Z$), so it remains

$$\begin{aligned}\mathbb{E}\left(X \mid Z = j\right) &= -j\mathbb{P}(X = -j \mid Z = j) + j\mathbb{P}(X = j \mid Z = j) \\ &= -j\mathbb{P}(\varepsilon = -1 \mid Z = j) + j\mathbb{P}(\varepsilon = 1 \mid Z = j).\end{aligned}$$

  Now, since the r.v. ε and Z are independent

$$\mathbb{P}(\varepsilon = -1 \mid Z = j) = \frac{\mathbb{P}(\varepsilon = -1, Z = j)}{\mathbb{P}(Z = j)} = \frac{\mathbb{P}(\varepsilon = -1)\mathbb{P}(Z = j)}{\mathbb{P}(Z = j)} = \mathbb{P}(\varepsilon = -1) = 1 - p$$

  and in the same manner $\mathbb{P}(\varepsilon = 1 \mid Z = j) = \mathbb{P}(\varepsilon = 1) = p$.
  Finally, $\mathbb{E}\left(X \mid Z = j\right) = -j(1 - p) + jp = j(2p - 1)$, which exactly means that $\mathbb{E}\left(X \mid Z\right) = (2p - 1)Z$.

Exercise: Check that $\mathbb{E}\left(X \mid Z\right)$ can also be written in the following form:

$$\begin{aligned}\mathbb{E}\left(X \mid Z\right)(\omega) &= \sum_{j=1}^{n} \mathbf{1}_{\{Z(\omega) = z_j\}} \mathbb{E}\left(X \mid Z = z_j\right) \\ &= \sum_{j=1}^{n} \mathbf{1}_{\{Z(\omega) = z_j\}} \left[\sum_{i=1}^{m} x_i \mathbb{P}(X = x_i \mid Z = z_j)\right].\end{aligned}$$

Remark: we usually write $\{Z = z_j\} := \{\omega \in \Omega \,:\, Z(\omega) = z_j\}$, for the event Z takes the value z_j. We also usually write $\mathbb{P}(X = x_i, Z = z_j)$ for $\mathbb{P}(\{X = x_i\} \cap \{Z = z_j\})$.

1.1.2 Conditioning against several random variables

Let us consider $n + 1$ random variables

$$\begin{aligned} X &: \Omega \to \{x_1, \dots, x_m\}, \\ Z_1 &: \Omega \to \{z_1, \dots, z_k\}, \\ &\vdots \\ Z_n &: \Omega \to \{z_1, \dots, z_k\}. \end{aligned}$$

We will extend the previous definition to the case where we deal with several random variables. In this case, formula (1.1) reads $\mathbb{E}\left(X \mid Z_1 = z_{j_1}, \ldots, Z_n = z_{j_n}\right) = \sum_{i=1}^{m} x_i \mathbb{P}(X = x_i \mid Z_1 = z_{j_1}, \ldots, Z_n = z_{j_n})$. Remind that this quantity is a real number.

Definition 1.2 Conditional expectation of X given $Z_1, \ldots, Z_n$.
We call conditional expectation of X given $Z_1, \ldots, Z_n$ the random variable *defined by:*

$$\mathbb{E}\left(X \mid Z_1, \ldots, Z_n\right)(\omega) := h(Z_1(\omega), \ldots, Z_n(\omega)),$$

where $h : \mathbb{R}^n \to \mathbb{R}$ is defined by $h(z_{j_1}, \ldots, z_{j_n}) = \mathbb{E}\left(X \mid Z_1 = z_{j_1}, \ldots, Z_n = z_{j_n}\right)$.

1.2 General case

1.2.1 Extension to the case where there exists a joint density

Reminder: Two random variables $X : \Omega \to \mathbb{R}$ and $Z : \Omega \to \mathbb{R}$, are said to have a density with respect to the Lebesgue measure, if there exists a function $f_{(X,Z)} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^+$ such that for any domain $\mathcal{D} \subset$ plane,

$$\mathbb{P}\big((X, Z) \in \mathcal{D}\big) = \int_{\mathcal{D}} f_{(X,Z)}(x, z)\, dx\, dz\,.$$

In particular,

$$\mathbb{P}(X \in dx, Z \in dz) := \mathbb{P}(X \in [x, x + dx], Z \in [z, z + dz]) = f_{(X,Z)}(x, z)\, dx\, dz\,.$$

Note that the margins can be computed in the following way

$$\mathbb{P}(X \in dx) = \left(\int_{z\in\mathbb{R}} f_{(X,Z)}(x, z)\, dz\right) dx \quad \text{and} \quad \mathbb{P}(Z \in dz) = \left(\int_{x\in\mathbb{R}} f_{(X,Z)}(x, z)\, dx\right) dz\,.$$

When we deal with random variables that have a density, then $\mathbb{P}(X = x)$ and $\mathbb{P}(Z = z)$ are equal to zero, so we cannot define the quantity $\mathbb{P}(X = x \mid Z = z)$. To bypass this difficulty we will define another quantity $\mathbb{P}(X \in dx \mid Z = z)$ that will represent "the probability that $X \in [x, x + dx]$ given $Z = z$."
In view of $\mathbb{P}(Z \in dz) = \left(\int_{x\in\mathbb{R}} f_{(X,Z)}(x, z)dx\right) dz$ and $\mathbb{P}(X \in dx, Z \in dz) = f_{(X,Z)}(x, z)\, dz\, dx$, it is natural to define $\mathbb{P}(X \in dx \mid Z = z)$ by

$$\mathbb{P}(X \in dx \mid Z = z) := \frac{\mathbb{P}(X \in dx, Z \in dz)}{\mathbb{P}(Z \in dz)} = \frac{f_{(X,Z)}(x, z)}{\int_{y\in\mathbb{R}} f_{(X,Z)}(y, z)dy}\, dx$$

Definition 1.3 *We call conditional expectation of X given Z the* random variable *defined by*

$$\mathbb{E}\left(X \mid Z\right)(\omega) = h(Z(\omega))$$

where $h : \mathbb{R} \to \mathbb{R}$ is defined by $h(z) = \int_{x\in\mathbb{R}} x\, \mathbb{P}(X \in dx \mid Z = z)$.

It is a "continuous space" version of the formula of 1.1.1.

1.2.2 General case

In the general case, we will give a more formal definition of the conditional expectation: we will define it by its properties, instead by a formula.

Reminder: let us consider $n+1$ r.v. $X : \Omega \to \mathbb{R}$, $Z_1 : \Omega \to \mathbb{R}, \dots, Z_n : \Omega \to \mathbb{R}$.
The r.v. X is so-called "$\sigma(Z_1, \dots, Z_n)$-measurable" iif[1] there exists a (measurable) function $h : \mathbb{R}^n \to \mathbb{R}$ such that $X(\omega) = h(Z_1(\omega), \dots, Z_n(\omega))$.
Let us draw a little scheme to illustrate the case $n = 1$: a random variable X is $\sigma(Z)$-measurable iif there exists a (measurable) function $h : \mathbb{R} \to \mathbb{R}$ such that $X = h(Z)$. In short:

$$
\begin{array}{ccc}
\Omega & \xrightarrow{X} & \mathbb{R} \\
 & \searrow \qquad \nearrow & \\
Z & \mathbb{R} & h
\end{array}
$$

We are now ready for giving a general definition of the conditional expectation.

Definition (and theorem) 1.1 *Let us consider $n+1$ random variables $X : \Omega \to \mathbb{R}$, $Z_1 : \Omega \to \mathbb{R}, \dots, Z_n : \Omega \to \mathbb{R}$, and set $\mathcal{F}_n := \sigma(Z_1, \dots, Z_n)$.*
There exists a unique random variable, called conditional expectation of X given $Z_1, \dots, Z_n$ (or given $\mathcal{F}_n$), we note $\mathbb{E}(X \mid Z_1, \dots, Z_n)$ (or $\mathbb{E}(X \mid \mathcal{F}_n)$), fulfilling the two properties:
a) $\mathbb{E}(X \mid \mathcal{F}_n)$ is $\mathcal{F}_n$-measurable
b) for any Y which is $\mathcal{F}_n$-measurable, we have $\mathbb{E}(\mathbb{E}(X \mid \mathcal{F}_n) Y) = \mathbb{E}(XY)$.

Comments: we admit the existence of the conditional expectation. Nevertheless, when $X \in L^2(\mathcal{F})$, we can check that $\mathbb{E}(X \mid \mathcal{F}_n)$ corresponds to the orthogonal projection of X onto the subspace $L^2(\mathcal{F}_n)$. See exercise 5 below for more details.
Besides, modifying a random variable on a set of null probability, does not change its law. Therefore, the uniqueness claimed in the above definition is to be understood as "uniqueness up to a modification on a set of null probability".
Remark 1: condition a) is equivalent to
a') there exists a (measurable) function h such that $\mathbb{E}(X \mid \mathcal{F}_n) = h(Z_1, \dots, Z_n)$.
In the same way, condition b) is equivalent to
b') $\mathbb{E}(\mathbb{E}(X \mid \mathcal{F}_n) f(Z_1, \dots, Z_n)) = \mathbb{E}(X f(Z_1, \dots, Z_n))$ for any (measurable) function $f : \mathbb{R}^n \to \mathbb{R}$.

Remark 2: In the discrete setting $\mathbb{E}(X \mid Z)$ and $\mathbb{E}(X \mid Z_1, \dots, Z_n)$ are given by the formula of the section 1.1.1 and 1.1.2.
In the same manner, if (X, Z) as a density with respect to the Lebesgue measure, $\mathbb{E}(X \mid Z)$ is given by the formula of section 1.2.1.

---

[1] if and only if

Let us check for example that we find back the formula of Section 1.1.1 For $X : \Omega \to \{x_1, \ldots, x_m\}$ and $Z : \Omega \to \{z_1, \ldots, z_k\}$, we define h by

$$\begin{cases} h(z_1) &= \mathbb{E}\left(X \mid Z = z_1\right) \\ &\vdots \\ h(z_k) &= \mathbb{E}\left(X \mid Z = z_k\right) \\ h(z) &= 0 \quad \text{if } z \notin \{z_1, \ldots, z_k\} \end{cases}$$

The formula of the section 1.1.1 defines the conditional expectation of X given Z by $\mathbb{E}\left(X \mid Z\right)(\omega) := h(Z(\omega))$. Does this definition fits with the one given above (1.2.2)? Let us check that $h(Z)$ fulfills conditions a') and b'), (which implies that the two definitions fit, since there exists a unique random variable fulfilling this two conditions).
Condition a') is clearly satisfied by $h(Z)$!
Let us check condition b'): do we have $\mathbb{E}\left(h(Z)f(Z)\right) \stackrel{?}{=} \mathbb{E}\left(Xf(Z)\right)$ for any function f? Check that

$$h(Z(\omega)) = \sum_{i=1}^{k} \underbrace{\mathbf{1}_{\{Z(\omega)=z_j\}}}_{\text{random variable}} \underbrace{\mathbb{E}\left(X \mid Z = z_j\right)}_{\text{non}-\text{random}},$$

so that

$$\begin{aligned} \mathbb{E}\left(h(Z)f(Z)\right) &= \mathbb{E}\Big[\textstyle\sum_{j=1}^{k} \mathbf{1}_{\{Z=z_j\}} \mathbb{E}\left(X \mid Z = z_j\right) \underbrace{f(Z)}_{=f(z_j)}\Big] \\ &= \textstyle\sum_{j=1}^{k} f(z_j) \underbrace{\mathbb{E}\left(X \mid Z = z_j\right) \underbrace{\mathbb{E}\left(\mathbf{1}_{\{Z=z_j\}}\right)}_{=\mathbb{P}(Z=z_j)}}_{=\mathbb{E}\left(X\mathbf{1}_{\{Z=z_j\}}\right)} \\ &= \mathbb{E}\left[\textstyle\sum_{j=1}^{k} f(z_j) X \mathbf{1}_{\{Z=z_j\}}\right] \end{aligned}$$

Note that $\sum_{j=1}^{k} f(z_j)\mathbf{1}_{\{Z(\omega)=z_j\}} = f(Z(\omega))$, which leads to $\mathbb{E}\left(h(Z)f(Z)\right) = \mathbb{E}\left(Xf(Z)\right)$. Condition b') is then satisfied. To conclude, we have check that the two definitions fit.

Informally, $\mathbb{E}\left(X \mid \mathcal{F}_n\right)$ represents the average value expected for X when one knows the values of $Z_1, \ldots, Z_n$. Let us come back to the example of the introduction.

Example: We throw two dice :
X_1 = value of the first dice
X_2 = value of the second dice
S = total value = $X_1 + X_2$.
Since X_1 et X_2 are independent we expect that $\mathbb{E}\left(S \mid X_1\right) = X_1 + \mathbb{E}\left(X_2\right)$, since "$X_1$ gives no information on X_2".
Write $h(x) = x + \mathbb{E}\left(X_2\right)$, and check that $h(X_1)$ fulfills conditions a) and b).

- Condition a): no problem.

- Condition b): let us check that $\mathbb{E}\left(h(X_1)f(X_1)\right) = \mathbb{E}\left(Sf(X_1)\right)$:

$$\mathbb{E}\left(h(X_1)f(X_1)\right) = \mathbb{E}\Big[(X_1 + \mathbb{E}(X_2))f(X_1)\Big] = \mathbb{E}\left(X_1 f(X_1)\right) + \underbrace{\mathbb{E}(X_2)\,\mathbb{E}(f(X_1))}_{\substack{=\mathbb{E}(X_2 f(X_1)) \\ \text{since } X_1 \perp X_2}}$$

  (Remind that if X_1 et X_2 are independent, $\mathbb{E}\left(f(X_1)g(X_2)\right) = \mathbb{E}\left(f(X_1)\right)\mathbb{E}\left(g(X_2)\right)$).
  Conclusion : $\mathbb{E}\left(h(X_1)f(X_1)\right) = \mathbb{E}\left((X_1+X_2)f(X_1)\right) = \mathbb{E}\left(Sf(X_1)\right)$, so condition b) is fulfilled and thus

$$\mathbb{E}\left(S \mid X_1\right) = h(X_1) = X_1 + \mathbb{E}(X_2) = X_1 + \frac{7}{2}.$$

In Practice, to compute a conditional expectation, we use the following properties.

Properties 1.1 (fundamental properties)

1. $\mathbb{E}\left(\mathbb{E}\left(X \mid \mathcal{F}_n\right)\right) = \mathbb{E}(X)$,
2. *When X is $\mathcal{F}_n$-measurable,* $\mathbb{E}\left(X \mid \mathcal{F}_n\right) = X$,
3. *When X is independent of $Z_1, \dots, Z_n$ then* $\mathbb{E}\left(X \mid Z_1, \dots, Z_n\right)(\omega) = \mathbb{E}(X)$ *a.s.*
4. *For any $a, b \in \mathbb{R}$,* $\mathbb{E}\left(aX + bY \mid \mathcal{F}_n\right) = a\,\mathbb{E}\left(X \mid \mathcal{F}_n\right) + b\,\mathbb{E}\left(Y \mid \mathcal{F}_n\right)$,
5. *When Y is $\mathcal{F}_n$-measurable (i.e. $Y = f(Z_1, \dots, Z_n)$) then* $\mathbb{E}\left(YX \mid \mathcal{F}_n\right) = Y\mathbb{E}\left(X \mid \mathcal{F}_n\right)$
6. $\mathbb{E}\big[\mathbb{E}\left(X \mid \mathcal{F}_{n+p}\right) \mid \mathcal{F}_n\big] = \mathbb{E}\left(X \mid \mathcal{F}_n\right)$.

Proof of some properties :
1. let Y be the random variable such that $Y(\omega) = 1, \quad \forall \omega \in \Omega$. Y is $\mathcal{F}_n$-measurable since $Y = h(Z_1, \dots, Z_n)$, where $h(z_1, \dots, z_n) = 1$. Condition b) gives $\mathbb{E}\left(\mathbb{E}\left(X \mid \mathcal{F}_n\right)Y\right) = \mathbb{E}(XY)$, and since $Y = 1$, we get $\mathbb{E}\left(\mathbb{E}\left(X \mid \mathcal{F}_n\right)\right) = \mathbb{E}(X)$.
2. X satisfies a) and if Y is $\mathcal{F}_n$ measurable, we have $\mathbb{E}(XY) = \mathbb{E}(XY)$(!), so X fulfills b), and $X = \mathbb{E}\left(X \mid \mathcal{F}_n\right)$.
6. We set $W = \mathbb{E}\left(X \mid \mathcal{F}_{n+p}\right)$. The random variable $\mathbb{E}\left(W \mid \mathcal{F}_n\right)$ is $\mathcal{F}_n$-measurable, so it fulfills condition a).
For any $\mathcal{F}_n$-measurable random variable Y, we have

$$\mathbb{E}\left(Y\mathbb{E}\left(W \mid \mathcal{F}_n\right)\right) \overset{\text{(condition b)}}{=} \mathbb{E}(YW) = \mathbb{E}\left(Y\mathbb{E}\left(X \mid \mathcal{F}_{n+p}\right)\right).$$

Now, since Y is $\mathcal{F}_n$ measurable, Y is $\mathcal{F}_{n+p}$ measurable (check it!), and condition b) gives $\mathbb{E}\left(Y\mathbb{E}\left(X \mid \mathcal{F}_{n+p}\right)\right) = \mathbb{E}(YX)$. As a consequence, the random variable $\mathbb{E}\left(W \mid \mathcal{F}_n\right)$ fulfills condition b), so that

$$\mathbb{E}\left(X \mid \mathcal{F}_n\right) = \mathbb{E}\left(W \mid \mathcal{F}_n\right) = \mathbb{E}\left(\mathbb{E}\left(X \mid \mathcal{F}_{n+p}\right) \mid \mathcal{F}_n\right).$$

Properties 1.2 (other properties)

1. *When* $Y_1, \ldots, Y_p$ *are independent of* X *and* $Z_1, \ldots, Z_n$*, then*
$$\mathbb{E}\left(X \mid Z_1, \ldots, Z_n, Y_1, \ldots, Y_p\right) = \mathbb{E}\left(X \mid Z_1, \ldots, Z_n\right),$$
2. *If* $X \leq Y$, $\mathbb{E}\left(X \mid \mathcal{F}_n\right) \leq \mathbb{E}\left(Y \mid \mathcal{F}_n\right)$.

What can we say about $\mathbb{E}\left(f(X) \mid \mathcal{F}_n\right)$?

- Finite setting : $\mathbb{E}\left(f(X) \mid Z_1, \ldots, Z_n\right)(\omega) = h_f(Z_1(\omega), \ldots, Z_n(\omega))$, where
$$h_f(z_1, \ldots, z_n) = \sum_{i=1}^{n} f(x_i)\mathbb{P}(X = x_i \mid Z_1 = z_1, \ldots, Z_n = z_n).$$
- Continuous setting : $\mathbb{E}\left(g(X) \mid Z\right)(\omega) = h_g(Z(\omega))$ where
$$h_g(z) = \frac{\int g(x) f_{(X,Z)}(x, z)dx}{\int f_{(X,Z)}(x, z)dx}$$

At last, when $\Phi : \mathbb{R} \to \mathbb{R}$ is a convex function, we have the **Jensen inequality**:
$$\Phi(\mathbb{E}\left(X \mid \mathcal{F}_n\right)) \leq \mathbb{E}\left(\Phi(X) \mid \mathcal{F}_n\right),$$
that ensures for example that $\mathbb{E}\left(X \mid \mathcal{F}_n\right)^2 \leq \mathbb{E}\left(X^2 \mid \mathcal{F}_n\right)$.

1.3 Exercises

1. We throw a dice and write N for the result (between 1 and 6). Then, we throw N^2 times the dice, and write S for the sum of the results (including the first throw). Compute $\mathbb{E}(S|N)$ and then $\mathbb{E}(S)$.
2. Assume that the random variables (X, Y) have a density
$$f(x, y) := n(n-1)(y-x)^{n-2}\mathbf{1}_{(x,y)\in A}$$
where $A := \{(x, y) \in \mathbb{R}^2 / 0 \leq x \leq y \leq 1\}$. Check that
$$\mathbb{E}(Y|X) = \frac{n-1+X}{n}.$$
3. Let $(X_n; n \in \mathbb{N})$ be a sequence of independent random variables. We focus on the random walk $S_n := X_1 + \ldots + X_n$ and set $\mathcal{F}_n = \sigma(S_1, \ldots, S_n)$.
a) Compute $\mathbb{E}(S_{n+1}|\mathcal{F}_n)$.
b) For any $z \in \mathbb{C}$, check that
$$\mathbb{E}\left(z^{S_{n+1}}|\mathcal{F}_n\right) = z^{S_n}\mathbb{E}\left(z^{X_{n+1}}\right).$$

4. With the Jensen inequality:
   a) For any $p > 1$ prove that

$$\mathbb{E}(|X||\mathcal{F}_n)^p \leq \mathbb{E}(|X|^p|\mathcal{F}_n).$$

   b) Prove the Jensen inequality in the discrete setting.

5. We assume that $X \in L^2(\mathcal{F})$. Check that

$$\min_h \mathbb{E}\left((X - h(Z))^2\right) = \mathbb{E}\left((X - \mathbb{E}(X|Z))^2\right).$$

   Conclude that $\mathbb{E}(X|Z)$ is the orthogonal projection of X onto $L^2(\sigma(Z))$.

Chapter 2

Martingales

Example: a gambler in a casino.
Let us investigate the fortune of a gambler in a casino:

- J = the set of all the possible games

  A game : $j \in J$

  ex: – to stake on 2 at the roulette
  – not to stake

- we write $R_j^{(n)}$ for the return of game j at time n. This is a random variable.

  ex: – if j = to stake on 2, $R_j^{(n)} = \begin{cases} 36 \text{ if 2 occurs} \\ 0 \text{ else} \end{cases}$
  – if j = not to stake, $R_j^{(n)} = 1$

  We also write $R^{(n)} := (R_j^{(n)}; j \in J)$ for the random vector made of the return at time n of each game.

- The gambler will stake at time n on each game: we write $M_j^{(n)}$ for its stake (possibly zero) on game j at time n.

- We write X_n for the fortune of the gambler after time n. X_0 is then its initial fortune. Since, "not to stake" is considered as a game, we have

$$X_{n-1} = \sum_{j \in J} M_j^{(n)}.$$

  After time n its fortune is $X_n = \sum_{j \in J} M_j^{(n)} R_j^{(n)}$.

- The information that have the gambler after time n is the value of the returns $R^{(1)}, \ldots, R^{(n)}$. We write $\mathcal{F}_n := \sigma(R^{(1)}, \ldots, R^{(n)})$ for this information.

Hypotheses of the game:

- The returns $R^{(1)}, \ldots, R^{(n)}, \ldots$ are independent.
- The stakes $M_j^{(n)}$ are $\mathcal{F}_{n-1}$-measurable: this means that the gambler choose its stakes at time n only on the basis of the information he has collected after time $n-1$. He is enable to predict the future!
- On average, the casino does not loose money, i.e. : $\mathbb{E}\left(R_j^{(n)}\right) \leq 1, \quad \forall n \in \mathbb{N},\ \forall j \in J.$

Problematic:

1. Is there a way for the gambler to choose its stakes, so that he wins money on average, viz so that $\mathbb{E}(X_n) > X_0$?
2. Is there a way to "stop gambling" so that, if T is the (random) time at which ones stop, $\mathbb{E}(X_T) > X_0$?

Answer to the first question:
Let us compute the conditional expectation of the gambler's fortune after time n, given $\mathcal{F}_{n-1}$:

$$\begin{aligned}
\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) &= \mathbb{E}\left(\sum_{j\in J} M_j^{(n)} R_j^{(n)} \mid \mathcal{F}_{n-1}\right) \\
&= \sum_{j\in J} \mathbb{E}\left(M_j^{(n)} R_j^{(n)} \mid \mathcal{F}_{n-1}\right) \\
&= \sum_{j\in J} M_j^{(n)} \mathbb{E}\left(R_j^{(n)} \mid \mathcal{F}_{n-1}\right) \quad \text{since } M_j^{(n)} \text{ is } \mathcal{F}_{n-1}\text{-measurable} \\
&= \sum_{j\in J} M_j^{(n)} \mathbb{E}\left(R_j^{(n)}\right) \quad \text{since } R_j^{(n)} \text{ is independent of } \mathcal{F}_{n-1} \\
&\leq \sum_{j\in J} M_j^{(n)} = X_{n-1} \quad \text{since } \mathbb{E}\left(R_j^{(n)}\right) \leq 1\,.
\end{aligned}$$

We thus have proved that

$$\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) \leq X_{n-1} \tag{2.1}$$

and we say that $(X_n)_n$ is a surpermartingale
Taking the expectation of inequality (2.1), leads to $\mathbb{E}(\mathbb{E}(X_n \mid \mathcal{F}_{n-1})) \leq \mathbb{E}(X_{n-1})$. According to the property 1.1.1 of the conditional expectation, it follows that $\mathbb{E}(X_n) \leq \mathbb{E}(X_{n-1})$. Iterating this inequality gives $\mathbb{E}(X_n) \leq \mathbb{E}(X_{n-1}) \leq \ldots \leq X_0$. The answer to the first question is thus "no" (which does not surprise anybody).

With regard to the second question, we need to define the notion of "stop gambling", which is the purpose of the next section.

2.1 Information, filtration, stopping time

2.1.1 Information, filtration

In the previous example, the gambler's information after time n is the return of each game at times $1, \ldots, n$, viz $R^{(1)}, \ldots, R^{(n)}$. We write $\sigma(R^{(1)}, \ldots, R^{(n)})$ for that information.
In the general case, when we observe a temporal random phenomenon (ex: the quotation of a share), the information we have at time n is the value of $X_1, \ldots, X_n$ (ex: the daily quotation of a share). We write $\sigma(X_1, \ldots, X_n)$ for this information. To get lighter notations, we usually set $\mathcal{F}_n := \sigma(X_1, \ldots, X_n)$. As time passes, we collect more and more information so that

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \ldots$$

Comment: In mathematics, $\mathcal{F}_n$ corresponds to a *σ-algebra*, and a filtration corresponds to an increasing sequence $(\mathcal{F}_n)_{n\in\mathbb{N}}$ of sigma-algebra.

2.1.2 Stopping time

In this section, we will formalize the notion of "stop gambling". The gambler is not expected to be able to predict the future. As a consequence, when he decides to stop gambling at time n, he takes its decision only from the information $\mathcal{F}_n$ he has at this time. For example, he can decide to stop when his fortune has reached twice its initial fortune, or when 5 consecutive blacks occur in the roulette, etc. A contrario, he can't decide to stop when his fortune is maximal (the ideal!), since he should be able to predict the future. Informally, a stopping time is a time where we decide to stop only from the information we have (at *this* time).

Let us give another example of stopping time: a selling order to your banker. You own a share and you ask your banker to sell it at a given time T depending on the evolution of the market. Unfortunately, your banker can't predict the future, so you can't ask him to sell it when the quotation is at its maximum. The time T at which he will sell your share shall only depends on the information he has at this time. For example, T can be the first time where the quotation pass over 100 \$, or the first time where the quotation has increased of 15% in the last 100 days, etc. In both cases, your banker does not need more information than the one he has at time T.

What are the characteristic of a stopping time T? Let us focus on the selling order and write $\mathcal{F}_n$ for the information we have at time n. Note that the time T is a random variable (it depends on the evolution of the market). The key idea is that we are able to say at time n (namely from the information $\mathcal{F}_n$) whether $T = n$ or not. Before going further, let us give some notation.

Notation 1: to a set A, we associate the indicator function $\mathbf{1}_A$, which is a random variable defined by:

$$\begin{array}{rccl} \mathbf{1}_A: & \Omega & \to & \{0,1\} \\ & \omega & \mapsto & \mathbf{1}_A(\omega) = \begin{cases} 1 \text{ if } \omega \in \Omega \\ 0 \text{ if } \omega \notin \Omega \end{cases} \end{array}$$

In the same way, if $X : \Omega \to \mathbb{R}$ is a random variable and B is a subset of $\mathbb{R}$, we set $\mathbf{1}_{\{X\in B\}}$ for the random variable:

$$\begin{array}{rccl} \mathbf{1}_{\{X\in B\}}: & \Omega & \to & \{0,1\} \\ & \omega & \mapsto & \begin{cases} 1 \text{ if } X(\omega) \in B \\ 0 \text{ if } X(\omega) \notin B \end{cases} \end{array}$$

Notation 2: we write "$A \in \mathcal{F}_n$" for "$\mathbf{1}_A$ is $\mathcal{F}_n$-measurable". For example "$\{T = n\} \in \mathcal{F}_n$" means that "$\mathbf{1}_{\{T=n\}}$ is $\mathcal{F}_n$-measurable".

We have the following properties (properties of a σ-algebra).

Properties 2.1 (de $\mathcal{F}_n$)

1. *If $A \in \mathcal{F}_n$, then $\overline{A} \in \mathcal{F}_n$ where $\overline{A}$ = complementary of A.*
2. *If $A_1, A_2, \ldots \in \mathcal{F}_n$, then $\bigcup_i A_i \in \mathcal{F}_n$ and $\bigcap_i A_i \in \mathcal{F}_n$.*

We are now ready for defining precisely a stopping time.

Definition 2.1 *A stopping time is a random variable $T : \Omega \to \mathbb{N} \cup \{+\infty\}$ such that $\{T = n\} \in \mathcal{F}_n$ for any n.*

Informally, this says that at time n, we are able to say whether $T = n$ or not.

Proposition 2.1 *Let us consider a stochastic process (X_n). Write $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ for our information at time n and $T = \inf\{n \geq 1$ such that $X_n \geq a\}$. The time T represents the first time where X_n passes over a.*
Then, the time T is a stopping time. This result still holds true when replacing $\geq$ by $=$, $\leq$, $<$ or $>$.

A contrario, The first time M where X_n reaches it maximum, namely $M = \inf\{n \geq 1$ such that$X_n = \max_{\{p\in\mathbb{N}\}} X_p\}$ is not a stopping time. Indeed, at time n, you do not know whether $M = n$ or not (you need to known the future)
Proof : Note that:

$$\{T = n\} = \{X_1 < a\} \cap \cdots \cap \{X_{n-1} < a\} \cap \{X_n \geq a\}.$$

Moreover for $k = 1 \ldots n-1$, we have $\{X_k < a\} \in \mathcal{F}_k \subset \mathcal{F}_n$ and also $\{X_n \geq a\} \in \mathcal{F}_n$. According to the property 2.1.2, we conclude to $\{T = n\} \in \mathcal{F}_n$. □

We now define the information we have at time T.

Definition 2.2 *The information $\mathcal{F}_T$ we have at time T is defined as follows:*

$$A \in \mathcal{F}_T \Leftrightarrow A \cap \{T = n\} \in \mathcal{F}_n, \quad \forall n \in \mathbb{N}.$$

Next proposition enumerates the properties of $\mathcal{F}_T$.

Proposition 2.2

1. *If $\sigma(X_1, \ldots, X_n) \subset \mathcal{F}_n$ and T is a stopping time,*

   then the random variable X_T defined by $X_T(\omega) = \begin{cases} X_1(\omega) & \text{if } T(\omega) = 1 \\ \vdots & \\ X_n(\omega) & \text{if } T(\omega) = n \\ \vdots & \\ 0 & \text{if } T(\omega) = +\infty \end{cases}$

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

2. *If $T(\omega) = n$, then $\mathbb{E}\left(Y \mid \mathcal{F}_T\right)(\omega) = \mathbb{E}\left(Y \mid \mathcal{F}_n\right)(\omega)$.*

3. *If $S \leq T$ are two stopping times, then $\mathcal{F}_S \subset \mathcal{F}_T$.*

Proof :
1. we shall check that $X_T : \Omega \to \mathbb{R}$ is $\mathcal{F}_T$-measurable, i.e. that $\{X_T \in A\} \in \mathcal{F}_T$ for any $A \subset \mathbb{R}$:

$$\begin{aligned} \{X_T \in A\} \in \mathcal{F}_T \quad &\Leftrightarrow \quad \{X_T \in A\} \bigcap \{T = n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N} \\ &\Leftrightarrow \quad \underbrace{\{X_n \in A\}}_{\in \mathcal{F}_n} \bigcap \underbrace{\{T = n\}}_{\in \mathcal{F}_n} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}, \quad \text{which is true.} \end{aligned}$$

2. et 3. : exercise. □

2.2 Martingales

2.2.1 Definition

We observe a random process $(X_n)_{n \in \mathbb{N}}$. We write $\mathcal{F}_n$ for our information at time n. This information contains $\sigma(X_0, X_1, \ldots, X_n)$, given by the values $X_0, X_1, \ldots, X_n$, that is to say $\sigma(X_0, X_1, \ldots, X_n) \subset \mathcal{F}_n$. Usually, our only information is $X_0, X_1, \ldots, X_n$ so that $\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$.

Definition 2.3 (Martingale)

- *When $\mathbb{E}\left(X_{n+1} \mid \mathcal{F}_n\right) = X_n, \ \forall n \geq 0$, we say that $(X_n)_{n \in \mathbb{N}}$ is a martingale (more precisely a $\mathcal{F}_n$-martingale).*
- *When $\mathbb{E}\left(X_{n+1} \mid \mathcal{F}_n\right) \geq X_n, \ \forall n \geq 0$, we say that $(X_n)_{n \in \mathbb{N}}$ is a submartingale.*
- *When $\mathbb{E}\left(X_{n+1} \mid \mathcal{F}_n\right) \leq X_n, \ \forall n \geq 0$, we say that $(X_n)_{n \in \mathbb{N}}$ is a supermartingale.*

Informally, a martingale is a stochastic process, such that:
"the average value expected for tomorrow given the information we have today, is equal to the today's value" (no rise nor fall in average)

Exercise: Let $(M_n)_{n\in\mathbb{N}}$ be a martingale. Check that for any positive n and p

$$\mathbb{E}(M_{n+p}|\mathcal{F}_n) = M_n \ .$$

Example:

- The gambler's fortune in a casino is a supermartingale.

  In this case, $\sigma(X_1, \ldots, X_n) \subsetneq \mathcal{F}_n$

- The random walk $S_n = Y_1 + \ldots + Y_n - nm$, where the (Y_i)'s are iid and $m = \mathbb{E}\left(Y_1\right)$. The process $(S_n)_{n\in\mathbb{N}}$ is a martingale, with $\mathcal{F}_n = \sigma(S_0, S_1, \ldots, S_n)$. Indeed,

  $$\begin{aligned}\mathbb{E}\left(S_{n+1} \mid \mathcal{F}_n\right) &= \mathbb{E}\left(Y_1 + \ldots + Y_{n+1} - (n+1)m \mid \mathcal{F}_n\right) \\ &= \mathbb{E}\left(S_n + Y_{n+1} - m \mid \mathcal{F}_n\right),\end{aligned}$$

  with S_n $\mathcal{F}_n$-measurable, and Y_{n+1} independent of $\mathcal{F}_n$. Therefore

  $\mathbb{E}\left(S_{n+1} \mid \mathcal{F}_n\right) = S_n + \underbrace{\mathbb{E}\left(Y_{n+1}\right)}_{=m} - m = S_n.$

- Assume that X is a random variable fulfilling $\mathbb{E}\left(|X|\right) < \infty$ and consider a given filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ (ex: $\mathcal{F}_n = \sigma(Z_0, \ldots, Z_n)$). Then the process defined by $M_n = \mathbb{E}\left(X \mid \mathcal{F}_n\right)$ is a martingale, since $\mathbb{E}\left(M_{n+1} \mid \mathcal{F}_n\right) = \mathbb{E}\left(\mathbb{E}\left(X \mid \mathcal{F}_{n+1}\right) \mid \mathcal{F}_n\right) = \mathbb{E}\left(X \mid \mathcal{F}_n\right) = M_n$.

We now state the fundamental property of the martingale:

Property 2.1 (fundamental) *The expected value of a martingale is constant as time passes, viz for any* $n \in \mathbb{N}$

$$\mathbb{E}\left(X_n\right) = \mathbb{E}\left(X_0\right) \ . \tag{2.2}$$

Proof : We check it by iterating the equality

$$\mathbb{E}\left(X_{n+1}\right) \overset{P1}{=} \mathbb{E}\left(\mathbb{E}(X_{n+1}|\mathcal{F}_n)\right) \overset{\text{martingale}}{=} \mathbb{E}\left(X_n\right).$$

□

One may wonder if equality (2.2), which is true at a fixed time n, still holds true at a random time T, viz do we have

$$\mathbb{E}\left(X_T\right) = \mathbb{E}\left(X_0\right) \ ?$$

For sure, this is not true at any random time T. For exemple, when T is the time where X_n reaches its maximum, we can't have $\mathbb{E}\left(X_T\right) = \mathbb{E}\left(X_0\right)$, except if $X_0 = X_T$ a.s. Indeed, $X_T \geq X_0$. Nevertheless, we shall see below, that this equality turns to be true when T is a stopping time (modulo some restrictions).

2.2.2 Optional stopping theorem

Notations: we set $n \wedge T := \min(n,T)$ and write $X_{n\wedge T}$ for the random variable

$$X_{n\wedge T}: \quad \Omega \to \mathbb{R}$$
$$\omega \mapsto X_{n\wedge T(\omega)}(\omega).$$

The next result due to the mathematician J.L. Doob, is one of the two fundamental results in martingale theory.

Theorem 2.1 (Optional stopping theorem) (Doob)
Assume that (X_n) is a martingale (respectively a supermartingale) and T is a stopping time. Then :

1. *The process $(X_{n\wedge T})_{n\in\mathbb{N}}$ is a martingale (resp. supermartingale),*
2. *When T is bounded a.s., i.e. when there exists $N \in \mathbb{N}$ such that $\mathbb{P}(T \le N) = 1$,*

$$\mathbb{E}\left(X_T\right) = \mathbb{E}\left(X_0\right) \quad (resp.\ \le),$$

3. *If $\mathbb{P}(T < \infty) = 1$ and if there exists Y such that $|X_{n\wedge T}| \le Y$ for any $n \in \mathbb{N}$, with $\mathbb{E}(Y) < \infty$*

$$\textit{then} \qquad \mathbb{E}\left(X_T\right) = \mathbb{E}\left(X_0\right) \quad (resp.\ \le).$$

Remarks:

1. Concerning the gambler, whatever his choice of T,
   $\mathbb{E}\left(X_T\right) \le \mathbb{E}\left(X_0\right) = X_0 =$ its initial fortune.
2. The hypotheses of 3. are necessary. Indeed, assume that we stake on a coin:
   - When tail occurs, we double our fortune.
   - When head occurs, we loose everything.

   Write $p_n \in \{tail, head\}$ for the result at time n and X_n for our fortune. We have

$$X_{n+1} = \begin{cases} 2X_n & \text{if } p_{n+1} = tail \\ 0 & \text{if } p_{n+1} = head \end{cases}$$

   viz $X_{n+1} = 2X_n\mathbf{1}_{\{p_{n+1}=tail\}}$. Since $X_0 = 1$, iterating previous equality gives $X_n = 2^n\mathbf{1}_{\{p_1=tail,\dots,p_n=tail\}}$. Let us check that (X_n) is a martingale:

$$\begin{aligned} \mathbb{E}\left(X_{n+1} \mid \mathcal{F}_n\right) &= \mathbb{E}\left(2X_n\mathbf{1}_{\{p_{n+1}=tail\}} \mid \mathcal{F}_n\right) \\ &= 2X_n\mathbb{E}\left(\mathbf{1}_{\{p_{n+1}=tail\}}\right) \quad \text{since} \begin{cases} 2X_n \text{ is } \mathcal{F}_n\text{-measurable} \\ \mathbf{1}_{\{p_{n+1}=tail\}} \text{ is independent of } \mathcal{F}_n \end{cases} \\ &= X_n \qquad \text{since } \mathbb{P}(p_{n+1} = tail) = 1/2 \end{aligned}$$

The time $T = \inf\{n \geq 0 \text{ such that } X_n = 0\}$ is a stopping time fulfilling $\mathbb{P}(T < \infty) = 1$. Since $X_{T(\omega)}(\omega) = 0$, we have $\mathbb{E}(X_T) = 0$, so that

$$\mathbb{E}(X_T) = 0 \neq \mathbb{E}(X_0) = 1.$$

Why? the preceding result cannot be applied since $|X_{n\wedge T}|$ cannot be dominated by any Y with finite mean.

Proof : of the optional stopping theorem
1. We focus on the case where (X_n) is a martingale. We have

$$X_{n\wedge T} = \begin{cases} X_n & \text{if } T > n-1 \\ X_T & \text{if } T \leq n-1 \end{cases}$$

or equivalently $X_{n\wedge T} = X_n \mathbf{1}_{\{T>n-1\}} + X_T \mathbf{1}_{\{T\leq n-1\}}$, so that

$$\begin{aligned}
\mathbb{E}(X_{n\wedge T} \mid \mathcal{F}_{n-1}) &= \mathbb{E}\big(X_n \underbrace{\mathbf{1}_{\{T>n-1\}}}_{\in \mathcal{F}_{n-1}} \mid \mathcal{F}_{n-1}\big) + \mathbb{E}\big(\underbrace{X_T \mathbf{1}_{\{T\leq n-1\}}}_{\substack{\mathcal{F}_{n-1}\text{-measurable} \\ \text{(prop 2.1)}}} \mid \mathcal{F}_{n-1}\big) \\
&= \mathbf{1}_{\{T>n-1\}} \underbrace{\mathbb{E}(X_n \mid \mathcal{F}_{n-1})}_{\substack{= X_{n-1} \text{ since} \\ (X_n) \text{ martingale}}} + X_T \mathbf{1}_{\{T\leq n-1\}} \\
&= X_{n-1}\mathbf{1}_{\{T>n-1\}} + X_T \mathbf{1}_{\{T\leq n-1\}} \\
&= X_{(n-1)\wedge T}.
\end{aligned}$$

We have proved that $X_{n\wedge T}$ is a martingale.
2. When T is bounded a.s.: let $N \in \mathbb{N}$ be such that $T \leq N$ a.s. The process $(X_{n\wedge T})$ is a martingale, so at time N:

$$\begin{array}{ccc}
\mathbb{E}(X_{N\wedge T}) & = & \mathbb{E}(X_{0\wedge T}) \\
\| & & \| \\
\mathbb{E}(X_T) & & \mathbb{E}(X_0)
\end{array}$$

3. If $T < \infty$ a.s. and $X_{n\wedge T}$ is dominated. The process $(X_{n\wedge T})$ is a martingale, so $\mathbb{E}(X_{n\wedge T}) = \mathbb{E}(X_0)$. Let n goes to infinity. First, note that $X_{n\wedge T} \to X_T$. Second, we can apply the theorem of dominated convergence, so that $\mathbb{E}(X_{n\wedge T}) \to \mathbb{E}(X_T)$. As a consequence $\mathbb{E}(X_T) = \mathbb{E}(X_0)$. □

Exercise: With the help of the first part of the optional stopping theorem and Fatou's lemma, prove that when (X_n) is a positive (super)martingale and T a stopping time, then we always have $\mathbb{E}(X_T) \leq \mathbb{E}(X_0)$ (without conditions on T)

In practice: how to use it ?

Usually, we choose some suitable martingale (X_n) and stopping time T and then apply the equality $\mathbb{E}(X_T) = \mathbb{E}(X_0)$ (warning: check the hypotheses of the optional stopping theorem!).

Example: The bankruptcy problem. Here is a simple example of use of the optional stopping theorem. Two persons A and B stake 1\$ on a coin. A the initial time A has a\$ and B has b\$. They play until one of them has been bankrupted.
A's fortune : $X_n = a + \varepsilon_1 + \ldots + \varepsilon_n$, with (ε_i) iid, $\varepsilon_i = \begin{cases} 1 & \text{if A wins at time } i \\ -1 & \text{if A looses at time } i \end{cases}$
(X_n) is a martingale and the time T of bankruptcy, namely $T = \inf\{n \geq 0 \text{ such that } X_n = 0 \text{ or } a+b\}$ is a stopping time. Furthermore, $|X_{n\wedge T}| \leq a + b$, so we can apply the optional stopping theorem, which ensures that

$$a = \mathbb{E}(X_0) = \mathbb{E}(X_T) = 0\,\mathbb{P}(\text{A bankrupted}) + (a+b)\,\mathbb{P}(\text{B bankrupted}).$$

As a consequence

$$\mathbb{P}(\text{A wins B's fortune}) = \frac{a}{a+b}\,.$$

2.3 Exercises

2.3.1 Galton-Watson genealogy

We focus henceforth on a simple model of population. At the initial time $n = 0$, there is one individual (the ancestor). We write Z_n for the number of individual at the n-th generation. We assume that each individual gives birth to children independently of the others. We also assume that the number of children of an individual follows some "universal" law μ. Universal means that it is the same for everybody.
Let us formalize this situation. We write $X_k^{(n)}$ for the number of children of the individual k present at generation n. Previous hypothesis says that the random variables

$$\left(X_k^{(n)}; k \in \mathbb{N}, n \in \mathbb{N}\right)$$

are independent and that their law is μ, viz $\mathbb{P}\left(X_k^{(n)} = i\right) = \mu(i)$. Furthermore, note that we have the formula

$$Z_{n+1} = X_1^{(n)} + \cdots + X_{Z_n}^{(n)}.$$

Let X be a random variable distributed according to μ (i.e. $\mathbb{P}(X = i) = \mu(i)$). We set $m := \mathbb{E}(X)$, and write G for its generative function, defined by

$$G(s) := \mathbb{E}\left(s^X\right), \text{ for } s \in [0,1].$$

We assume henceforth that $\mu(0) > 0$.

1. Express $G(s)$ in terms of the $\mu(k)$, $k \in \mathbb{N}$. What is the value of $G(1)$? $G(0)$?

2. Check that G is non-decreasing and convex. What is the value of $G'(1)$? Draw the functions $x \mapsto G(x)$ and $x \mapsto x$ in the case where $m \leq 1$ and $m > 1$. Take care of the behavior of G around 1.

3. We set $\mathcal{F}_n := \sigma\left(X_k^{(1)}, \dots, X_k^{(n-1)}; k \in \mathbb{N}\right)$, which represents the information contained in the genealogical tree up to generation $n-1$. Compute $\mathbb{E}(s^{Z_{n+1}}|\mathcal{F}_n)$. Deduce that
$$G_n(s) := \mathbb{E}\left(s^{Z_n}\right) = \underbrace{G \circ \cdots \circ G}_{n \text{ times}}(s).$$

4. Express with words what $G_n(0)$ corresponds to. We focus now on the probability of extinction, viz on $p := \mathbb{P}(\exists\, n \in \mathbb{N} : Z_n = 0)$. Express p in terms of the $G_n(0)$'s.

5. With the help of the figure of 2., check that $p = 1$ when $m \leq 1$ and $p = G(p) < 1$, when $m > 1$.

6. We set $M_n := Z_n/m^n$. Check that M_n is a martingale. What is the value of $\mathbb{E}(M_n)$?

7. We admits that the limit $M_\infty(\omega) := \lim_{n\to\infty} M_n(\omega)$ exists a.s. When $m \leq 1$, what is the value of M_∞? $\mathbb{E}(M_\infty)$? Compare $\mathbb{E}(M_\infty)$ and $\lim_{n\to\infty} \mathbb{E}(M_n)$. Give a heuristic explanation to this result.

8. Prove that: $\mathbb{E}(\exp(-\lambda M_n)) = G_n(\exp(-\lambda/m^n))$.

9. For any $\lambda > 0$, prove with the theorem of dominated convergence that
$$\lim_{n\to\infty} \mathbb{E}(\exp(-\lambda M_n)) = \mathbb{E}(\exp(-\lambda M_\infty)) =: L(\lambda).$$
Derive from 8. (and $G_n = G \circ G_{n-1}$) the functional equation satisfied by L:
$$L(\lambda) = G(L(\lambda/m)).$$

Remark: This equation completely determine L and thus the distribution of M_∞.

2.3.2 With the optional stopping theorem

- Maximal inequality (Doob)

Assume that $(X_n)_{n\in\mathbb{N}}$ is a non-negative martingale. To any real $a > 0$, we associate $\tau_a := \inf\{k \in \mathbb{N} : X_k \geq a\}$, the first passage time of X_n above a.

a) Is τ_a a stopping time?

b) Fix an integer n. Prove that
$$\mathbb{E}(X_{\min(\tau_a,n)}) = \mathbb{E}(X_0) = \mathbb{E}(X_n)$$

and derive the equality

$$\mathbb{E}(X_{\tau_a}\mathbf{1}_{\tau_a\le n}) = \mathbb{E}(X_n\mathbf{1}_{\tau_a\le n}).$$

c) Derive from this last formula, the upper bound (so-called "Doob's maximal inequality")

$$\mathbb{P}\left(\max_{k=1\dots n} X_k \ge a\right) \le \frac{1}{a}\,\mathbb{E}\left(X_n\mathbf{1}_{\max_{k=1\dots n} X_k\ge a}\right).$$

- Wald identity

We focus on the random walk $S_n = X_1+\cdots+X_n$, with $(X_n)_{n\in\mathbb{N}}$ i.i.d. We set $m := \mathbb{E}(X_n)$ and $T := \inf\{n\in\mathbb{N} : S_n \ge a\}$ for a given $a > 0$.
a) What can you say about T?
b) Assume that $m > 0$. What can we say about $(S_n - nm)_{n\in\mathbb{N}}$? Derive the equality

$$m\,\mathbb{E}(T\wedge n) = \mathbb{E}(S_{T\wedge n}) = \mathbb{E}(S_T\mathbf{1}_{T\le n}) + \mathbb{E}(S_n\mathbf{1}_{T>n}).$$

c) Prove that $\mathbb{E}(S_n\mathbf{1}_{T>n}) \le a\,\mathbb{P}(T>n)$, and then justify with the law of large numbers the equality (so-called Wald identity)

$$\mathbb{E}(T) = \frac{1}{m}\,\mathbb{E}(S_T) \qquad \text{(focus only on the case where } X_n\ge 0 \text{ for all } n).$$

d) We assume now that $m = 0$. Prove that (S_n) is a martingale. Do we have $\mathbb{E}(S_T) = \mathbb{E}(S_0)$? Why? Fix $\epsilon > 0$, and set $T_\epsilon = \inf\{k\in\mathbb{N} : S_k + \epsilon k \ge a\}$. Prove that $T \ge T_\epsilon$ and (use the Wald identity)

$$\mathbb{E}(T) \ge \mathbb{E}(T_\epsilon) = \frac{1}{\epsilon}\,\mathbb{E}(S_{T_\epsilon} + \epsilon T_\epsilon) \ge \frac{a}{\epsilon}.$$

What is the value of $\mathbb{E}(T)$?

Chapter 3

The B-S market

Goals: to define the B-S market and to rely the mathematical concept of "neutral risk probability" to the economical concepts of "arbitrage" and "completeness".

3.1 The B-S market

3.1.1 Evolution

We focus henceforth on a market made of two assets:

- a non risky asset B, or bond (with predictable evolution),
- a risky asset S, or Stock (with unpredictable evolution).

We write $\Delta B_n := B_n - B_{n-1}$ and $\Delta S_n := S_n - S_{n-1}$ for the variation of B ans S between the time $n-1$ and n. The assets evolve according to the dynamic:

$$\begin{cases} \Delta B_n &= r_n \, B_{n-1} \\ \Delta S_n &= \rho_n \, S_{n-1} \end{cases} \tag{3.1}$$

with r_n (interest rate) and ρ_n (return) random in general. We write $\mathcal{F}_n$ for the information we have at time n. Since $S_0, \ldots, S_n$ are known at time n, we have $\sigma(S_0, \ldots, S_n) \subset \mathcal{F}_n$. The asset B is said non risky, because its evolution is predictable: at time $n-1$ we know the value of the interest rate r_n for the time n. The variable r_n is thus $\mathcal{F}_{n-1}$-measurable. A contrario, S is a risky asset: at time $n-1$ we do not know the value of ρ_n. The random variable ρ_n is thus $\mathcal{F}_n$-measurable, but not $\mathcal{F}_{n-1}$-measurable.

Remark: Here the time n is discrete. Note that we can see $\Delta X_n = X_n - X_{n-1}$ as the derivative of X in discrete time. We also assume henceforth that B and S can only take a finite number of value. In other words, we are in a completely discrete setting.
Note that this corresponds to the reality. First the quotation are given with only a finite precision. Second, the quotations occur in discrete time.

We can rewrite the dynamic as follows: equation (3.1) gives $B_n - B_{n-1} = r_n B_{n-1}$ so that $B_n = (1+r_n)B_{n-1}$.
Iterating this formula gives : $B_n = (1+r_n)\dots(1+r_1)B_0$. Setting

$$U_n = \sum_{k=1}^{n} r_k \quad \text{and} \quad V_n = \sum_{k=1}^{n} \rho_k$$

we obtain

$$\begin{cases} B_n &= B_0\, \varepsilon_n(U) \\ S_n &= S_0\, \varepsilon_n(V) \end{cases}$$

with $\varepsilon_n(U) = (1+\Delta U_n)\dots(1+\Delta U_1)$. In analogy with the continuous time setting, we will call the random variable $\varepsilon_n(U)$ "stochastic exponential".

3.1.2 Portfolio

Let us consider a portfolio $\Pi = (\beta_n, \gamma_n)_{n\le N}$ made of β_n units of B and γ_n units of S at time n. Its value at time n is

$$X_n^{\Pi} = \beta_n B_n + \gamma_n S_n.$$

We manage the portfolio in the following way. At time n, we have β_n units of asset B and γ_n units of asset S. We then decide to reinvest for the next quotation, namely we choose β_{n+1} and γ_{n+1}. This choice occurs at time n. This means that β_{n+1} and γ_{n+1} are $\mathcal{F}_n$-measurable (or equivalently β_n and γ_n are $\mathcal{F}_{n-1}$-measurable).

At time n, the value of the portfolio is $X_n^{\Pi} = \beta_n B_n + \gamma_n S_n$. After the reinvestment, its value is $\beta_{n+1}B_n + \gamma_{n+1}S_n$. It is natural to assume that when we reinvest money, no value is added or lost, viz

$$\beta_n B_n + \gamma_n S_n = \beta_{n+1}B_n + \gamma_{n+1}S_n.$$

Such a trading strategy is said self-financed. In the following definition, the self-financing condition is expressed at time $n-1$ instead of n.

Definition 3.1 *A portfolio (or trading strategy) is self-financed when*

$$\beta_{n-1}B_{n-1} + \gamma_{n-1}S_{n-1} = \beta_n B_{n-1} + \gamma_n S_{n-1}$$

which can also be written as

$$\Delta\beta_n\ B_{n-1} + \Delta\gamma_n\ S_{n-1} = 0.$$

Warning: the value of the portfolio does not change during the *reinvestment*, but it changes between two consecutive time, due to the fluctuation of B and S. As an exercise, check that Π is self-financed if and only if the fluctuation of X^{Π} between time $n-1$ and time n is

$$\Delta X_n^{\Pi} = \beta_n \Delta B_n + \gamma_n \Delta S_n.$$

3.2 Arbitrage and martingale

3.2.1 Risk neutral probability

We have seen on the model "one step, two states" that it could be judicious not to investigate the evolution of the market under the *real world probability* but under another artificial probability the so-called *risk neutral probability*. This probability tends to annihilate the average rise or fall of the stock. We generalize here this point of view.

Definition 3.2 *A probability $\mathbb{P}^*$ is a risk neutral probability if*

- $\mathbb{P}^*$ *is equivalent to the real world probability* $\mathbb{P}$*, namely* $\mathbb{P}^*(A) = 0$ *iif* $\mathbb{P}(A) = 0$*,*
- $(S_n/\varepsilon_n(U))_{n\leq N}$ *is a martingale under* $\mathbb{P}^*$*.*

The second condition is equivalent to "$(S_n/B_n)_{n\leq N}$ *is a martingale under* $\mathbb{P}^*$". Next proposition gives a condition on U_n and V_n that ensures that a probability is "risk neutral".

Proposition 3.1 *Assume that* $\mathbb{P}^*$ *is equivalent to* $\mathbb{P}$*. Then*
$\mathbb{P}^*$ *is a risk neutral probability* *iif* $(V_n - U_n)$ *is a martingale under* $\mathbb{P}^*$*.*

To prove this result we need the following properties of the stochastic exponential.

Lemma 3.1 Properties of $\varepsilon_n(\cdot)$:

1. $\varepsilon_n(X)\varepsilon_n(Y) = \varepsilon_n(X + Y + [X, Y])$*, where* $[X, Y]_n = \sum_{k=1}^n \Delta X_k\, \Delta Y_k$ *is the quadratic variation of* X *and* Y*.*
2. $\frac{1}{\varepsilon_n(X)} = \varepsilon_n(-X^*)$*, where* $\Delta X_n^* = \Delta X_n - \frac{(\Delta X_n)^2}{1+\Delta X_n}$.
3. $(\varepsilon_n(X))_{n\leq N}$ *is a martingale* *iif* $(X_n)_{n\leq N}$ *is a martingale.*

Proof : of the Lemma.
1. We have

$$\begin{aligned}\varepsilon_n(X)\varepsilon_n(Y) &= \varepsilon_{n-1}(X)\,(1+\Delta X_n)\,\varepsilon_{n-1}(Y)\,(1+\Delta Y_n)\\ &= \varepsilon_{n-1}(X)\varepsilon_{n-1}(Y)\,(1+\Delta X_n+\Delta Y_n+\Delta X_n\Delta Y_n)\\ &\overset{\text{iterate}}{=} \varepsilon_n(X+Y+[X,Y])\end{aligned}$$

2. Define X^* by $\Delta X_n^* = \Delta X_n - \frac{(\Delta X_n)^2}{1+\Delta X_n}$. In the view of part 1.:

$$\varepsilon_n(X)\varepsilon_n(-X^*) = \varepsilon_n(X - X^* - [X, X^*]).$$

Besides

$$\begin{aligned}\Delta(X - X^* - [X, X^*])_n &= \Delta X_n - \Delta X_n + \frac{(\Delta X_n)^2}{1+\Delta X_n} - \Delta X_n\left(\Delta X_n - \frac{(\Delta X_n)^2}{1+\Delta X_n}\right)\\ &= 0.\end{aligned}$$

Since for any process Z, we have $Z_n = Z_0 + \sum_{k=1}^{n} \Delta Z_k$, it follows that $(X - X^* - [X, X^*])_n = 0$ and finally $\varepsilon_n(X)\varepsilon_n(-X^*) = \varepsilon_n(0) = 1$.
3. Using that $\varepsilon_{n-1}(X)$ is $\mathcal{F}_{n-1}$-measurable:

$$\begin{aligned}\mathbb{E}(\varepsilon_n(X) \mid \mathcal{F}_{n-1}) &= \mathbb{E}((1 + \Delta X_n)\varepsilon_{n-1}(X) \mid \mathcal{F}_{n-1}) \\ &= \varepsilon_{n-1}(X)(1 + \mathbb{E}(\Delta X_n \mid \mathcal{F}_{n-1}))\end{aligned}$$

so

$$\begin{aligned}\mathbb{E}(\varepsilon_n(X) \mid \mathcal{F}_{n-1}) = \varepsilon_{n-1}(X) \quad &\underline{\text{iif}} \quad \mathbb{E}(\Delta X_n \mid \mathcal{F}_{n-1}) = 0 \\ &\underline{\text{iif}} \quad \mathbb{E}(X_n \mid \mathcal{F}_{n-1}) = X_{n-1}\end{aligned}$$

□

Proof : of the proposition.
We only have to check that $(S_n/\varepsilon_n(U))$ is a martingale under $\mathbb{P}^*$ if and only if $(V_n - U_n)$ is a martingale under $\mathbb{P}^*$. We have

$$\begin{aligned}\frac{S_n}{\varepsilon_n(U)} &= \frac{S_0\,\varepsilon_n(V)}{\varepsilon_n(U)} \\ &\overset{\text{Lemma 2.}}{=} S_0\,\varepsilon_n(V)\,\varepsilon_n(-U^*) \\ &\overset{\text{Lemma 1.}}{=} S_0\,\varepsilon_n(V - U^* - [V, U^*])\end{aligned}$$

so according to Lemma 3.1.3, the process $(S_n/\varepsilon_n(U))_{n\leq N}$ is a martingale under $\mathbb{P}^*$ iif $(\varepsilon_n(V - U^* - [V, U^*])_{n\leq N}$ is a martingale under $\mathbb{P}^*$. A computation gives

$$V_n - U_n^* - [V, U^*]_n = \sum_{k=1}^{n} \frac{\Delta V_k - \Delta U_k}{1 + \Delta U_k}$$

so $(\varepsilon_n(V - U^* - [V, U^*])_{n\leq N}$ is a martingale under $\mathbb{P}^*$ iif

$$\mathbb{E}^*\left(\sum_{k=1}^{n} \frac{\Delta V_k - \Delta U_k}{1 + \Delta U_k} \mid \mathcal{F}_{n-1}\right) = \sum_{k=1}^{n-1} \frac{\Delta V_k - \Delta U_k}{1 + \Delta U_k}.$$

Moreover $\Delta U_k = r_k$ is $\mathcal{F}_{k-1}$-measurable so

$$\begin{aligned}\mathbb{E}^*\left(\sum_{k=1}^{n} \frac{\Delta V_k - \Delta U_k}{1 + \Delta U_k} \mid \mathcal{F}_{n-1}\right) &= \sum_{k=1}^{n} \frac{1}{1 + \Delta U_k}\mathbb{E}^*\left(\Delta V_k - \Delta U_k \mid \mathcal{F}_{n-1}\right) \\ &= \sum_{k=1}^{n-1} \frac{\Delta V_k - \Delta U_k}{1 + \Delta U_k} + \frac{1}{1 + \Delta U_n}\mathbb{E}^*(\Delta V_n - \Delta U_n \mid \mathcal{F}_{n-1})\end{aligned}$$

and $(\varepsilon_n(V - U^* - [V, U^*])_{n\leq N}$ is a martingale iif $\mathbb{E}^*(\Delta V_n - \Delta U_n \mid \mathcal{F}_{n-1}) = 0$, namely iif $V - U$ is a martingale. Putting pieces together leads to the claimed result. □

Next lemma investigates the statistical evolution of the value of a portfolio based on a self-financed strategy under a risk neutral probability.

Proposition 3.2 *Assume that there exists a risk neutral probability* $\mathbb{P}^*$. *Then, if* Π *is a self-financed portfolio, its so-called* discounted value *(or* present value*)* $\varepsilon_n(U)^{-1}X_n^\Pi$ *is a martingale under* $\mathbb{P}^*$.

Proof : Since $B_n = B_0\,\varepsilon_n(U)$, the discounted value of the portfolio is given by

$$X_n^\Pi/\varepsilon_n(U) = \beta_n B_0 + \gamma_n S_n/\varepsilon_n(U).$$

Taking conditional expectation on both side leads to

$$\begin{aligned}
\mathbb{E}^*\left(\frac{X_n^\Pi}{\varepsilon_n(U)} \mid \mathcal{F}_{n-1}\right) &= \mathbb{E}^*\left(\beta_n B_0 \mid \mathcal{F}_{n-1}\right) + \mathbb{E}^*\left(\gamma_n \frac{S_n}{\varepsilon_n(U)} \mid \mathcal{F}_{n-1}\right) \\
&= \beta_n B_0 + \gamma_n \mathbb{E}^*\left(\frac{S_n}{\varepsilon_n(U)} \mid \mathcal{F}_{n-1}\right) \quad \text{since } \beta_n \text{ and } \gamma_n \text{ are } \mathcal{F}_{n-1}\text{-measurable} \\
&= \beta_n B_0 + \gamma_n \frac{S_{n-1}}{\varepsilon_{n-1}(U)} \qquad \text{since } (S_n/\varepsilon_n(U)) \text{ is a martingale under } \mathbb{P}^* \\
&= \left(\beta_n B_{n-1} + \gamma_n S_{n-1}\right)/\varepsilon_{n-1}(U) \qquad \text{since } B_{n-1} = B_0\,\varepsilon_{n-1}(U) \\
&= \left(\beta_{n-1} B_{n-1} + \gamma_{n-1} S_{n-1}\right)/\varepsilon_{n-1}(U) \qquad \text{self-financing strategy} \\
&= X_{n-1}^\Pi/\varepsilon_{n-1}(U).
\end{aligned}$$

Finally $(X_n^\Pi/\varepsilon_n(U))$ is thus a martingale under $\mathbb{P}^*$. □

3.2.2 Arbitrage

Henceforth, we focus on the evolution of the market until a fixed time N.
In economy, we say that there is an arbitrage opportunity if there is an opportunity to make a profit without any risk of loosing money. In mathematical words this become:

Definition 3.3 *There exists an arbitrage opportunity, when there exists a self-financed strategy* Π *such that*

$$X_0^\Pi = 0, \quad X_n^\Pi \geq 0 \quad \forall n \leq N, \quad \textit{and} \;\; \mathbb{P}(X_N^\Pi > 0) > 0.$$

Next fundamental theorem relies the economic notion of "arbitrage opportunity" to the mathematical notion of "neutral risk probability".

Theorem 3.1

There exists no arbitrage opportunity $\Longleftrightarrow$ *There exists* at least *one risk neutral probability.*

In this case one says that the market is arbitrage-free.
Proof :
($\Leftarrow$) Assume that $\mathbb{P}^*$ is a risk neutral probability and Π is a self-financed trading strategy such that $X_0^\Pi = 0$. According to Proposition 3.2 $(X_n^\Pi/\varepsilon_n(U))_{n\leq N}$ is a martingale under $\mathbb{P}^*$. It follows that

$$\mathbb{E}^*\left(\frac{X_N^\Pi}{\varepsilon_N(U)}\right) = \mathbb{E}^*\left(\frac{X_0^\Pi}{\varepsilon_0(U)}\right) = 0.$$

This equality together with $X_N^\Pi/\varepsilon_N(U) \geq 0$ imply that $\mathbb{P}^*(X_N^\Pi > 0) = 0$. Moreover $\mathbb{P}$ and $\mathbb{P}^*$ are equivalent, which implies in turns $\mathbb{P}(X_N^\Pi > 0) = 0$. In conclusion, there exists no arbitrage opportunities.
($\Rightarrow$) Admitted. See, example given [5] Chap. V Section 2.d. □

3.3 Complete market

In economy, a so-called complete market corresponds to an ideal market, without any constraints, nor transaction costs, where every assets can be found at any time and any quantities. In mathematics, this translates as follows.

Definition 3.4 *The B-S market is so-called complete, when for any random variable* $f : \Omega \to \mathbb{R}^+$, *there exists a self-financed trading strategy* Π *such that*

$$X_N^\Pi(\omega) = f(\omega), \quad \forall \omega \in \Omega.$$

Next theorem links this concept to the uniqueness of risk neutral probability.

Theorem 3.2 *Let us consider an arbitrage-free B-S market. Then*

The market is complete $\Longleftrightarrow$ *There exists a* unique *risk neutral probability* $\mathbb{P}^*$.

Proof : ($\Rightarrow$) Assume that the market is complete and that there exists two different risk neutral probabilities $\mathbb{P}^*$ and $\mathbb{P}'$. Since $\mathbb{P}^* \neq \mathbb{P}'$, there exists an event A such that

$$\mathbb{P}^*(A) \neq \mathbb{P}'(A) \ . \tag{3.2}$$

Set $f(\omega) = \varepsilon_N(U)(\omega)\mathbf{1}_{\{A\}}(\omega)$. Due to the completeness of the market there exists a trading strategy Π such that $X_N^\Pi(\omega) = f(\omega)$ whatever $\omega \in \Omega$. According to Proposition 3.2, the discounted value of the portfolio Π is a martingale under $\mathbb{P}^*$ and $\mathbb{P}'$. As a consequence,

$$\mathbb{E}^*\left(\varepsilon_N(U)^{-1}X_N^\Pi\right) = \mathbb{E}^*\left(\varepsilon_0(U)^{-1}X_0^\Pi\right) \quad \text{and} \quad \mathbb{E}'\left(\varepsilon_N(U)^{-1}X_N^\Pi\right) = \mathbb{E}'\left(\varepsilon_0(U)^{-1}X_0^\Pi\right).$$

Now, on the one hand $\varepsilon_0(U)^{-1} = 1$ and on the other hand $\varepsilon_N(U)^{-1}X_N^\Pi = \varepsilon_N(U)^{-1}f = \mathbf{1}_{\{A\}}$. Therefore

$$\mathbb{E}^*(\mathbf{1}_{\{A\}}) = X_0^\Pi = \mathbb{E}'(\mathbf{1}_{\{A\}}),$$

i.e. $\mathbb{P}^*(A) = \mathbb{P}'(A)$ which contradicts (3.2). There cannot exist two different risk neutral probabilities. Besides, since the market is arbitrage-free, there exists at least one risk neutral probability.
($\Leftarrow$) Admitted. See, example given [5] Chap. V Section 4. $\square$

3.4 Girsanov Lemma

We have seen that arbitrage and completeness are closely linked to risk neutral probabilities. Usually, it is hard to compute $\mathbb{P}^*$. Nevertheless, next lemma gives a way to compute $\mathbb{P}^*$ from $\mathbb{P}$.

There exists a very simple way to construct a probability $\mathbb{P}'$ from $\mathbb{P}$. Let us consider a random variable $Z > 0$ such that $\mathbb{E}(Z) = 1$. We define $\mathbb{P}'$ in setting

$$\mathbb{P}'(A) := \mathbb{E}(Z1_A) \quad \text{for any } A \in \mathcal{F}. \tag{3.3}$$

It is easily checked that $\mathbb{P}'$ is a probability measure. Furthermore, we can prove that any probability $\mathbb{P}'$ equivalent to $\mathbb{P}$ is of the previous form (3.3). Besides, note that the expectation under $\mathbb{P}'$ of any random variable X is given by

$$\mathbb{E}'(X) = \mathbb{E}(ZX).$$

Assume that you have a process which is not a martingale under $\mathbb{P}$. Next lemma gives (in some situations) a way to construct a probability $\mathbb{P}'$ such that the process becomes a martingale under $\mathbb{P}'$.

Lemma 3.2 Girsanov formula
Let $(M_n)_{n\leq N}$ be a $\mathcal{F}_n$-martingale under $\mathbb{P}$ and Z_N be a ($\mathcal{F}_N$-measurable) random variable, fulfilling $\mathbb{E}(Z_N) = 1$ et $Z_N(\omega) > 0$, $\forall \omega \in \Omega$.
Let us define $\mathbb{P}'$ by

$$\mathbb{P}'(A) := \mathbb{E}(Z_N 1_A),$$

where $\mathbb{E}$ represents the expectation under the probability $\mathbb{P}$. Then the process $(M_n^*)_{n\leq N}$ defined by

$$M_n^* = M_n - \sum_{k=1}^{n} \mathbb{E}\left(\frac{Z_k}{Z_{k-1}}\, \Delta M_k \middle| \mathcal{F}_{k-1}\right), \quad \textit{where} \quad Z_n = \mathbb{E}(Z_N|\mathcal{F}_n),$$

is a martingale under $\mathbb{P}'$.

See exercises below for a proof. Next example shows how to use this formula.

Example: Assume that the interest rates r_n are constant, viz $r_n = r > 0$ for any n. Assume also that the ρ_n's are i.i.d. and set

$$m = \mathbb{E}(\rho_n) \geq 0, \quad \text{and} \quad \sigma^2 = \mathbb{E}((\rho_n - m)^2) > 0.$$

Then, Girsanov formula ensures that the probability $\mathbb{P}^*$ defined by

$$\mathbb{P}^*(A) := \mathbb{E}(Z_N 1_A) \tag{3.4}$$

with $Z_N = \mathcal{E}_N(G)$ and $G_n := \frac{r-m}{\sigma^2}(\rho_1 + \cdots + \rho_n - mn)$, is a risk neutral probability (see exercises below).

3.5 Exercises

3.5.1 Change of probability: Girsanov lemma

Let us prove Girsanov formula. First note that for any random variable X, we have

$$\mathbb{E}'(X) := \mathbb{E}(Z_N X).$$

1. Set $Z_n = \mathbb{E}(Z_N|\mathcal{F}_n)$. Prove that $(Z_n)_{n\leq N}$ is a martingale under $\mathbb{P}$.
2. Check that for any $\mathcal{F}_p$-measurable random variable W, we have:
$$\mathbb{E}'(W) = \mathbb{E}(Z_N W) = \mathbb{E}(Z_p W).$$
3. Check that for any $\mathcal{F}_n$-measurable random variable Y, we have
$$\mathbb{E}'(Y|\mathcal{F}_{n-1}) = \frac{1}{Z_{n-1}}\mathbb{E}(Y Z_n|\mathcal{F}_{n-1}).$$
4. Set $\alpha_n = Z_n/Z_{n-1}$. Check that the process $(M_n^*)_{n\leq N}$ defined by
$$M_n^* = M_n - \sum_{k=1}^{n}\mathbb{E}(\alpha_k \Delta M_k|\mathcal{F}_{k-1}),$$
is a martingale under $\mathbb{P}'$ (Girsanov lemma).

3.5.2 Application: computation of a risk neutral probability

Example 3.4 (continued): assume that the interest rates r_n are constant, viz $r_n = r > 0$ for any n (and $U_n = rn$) and that the ρ_n are i.i.d. We set

$$m = \mathbb{E}(\rho_n) \geq 0, \quad \text{et} \quad \sigma^2 = \mathbb{E}((\rho_n - m)^2) > 0.$$

1. We set $V_n = \rho_1 + \ldots + \rho_n$ and $M_n = V_n - mn$. Check that $(M_n)_{n\leq N}$ is a martingale under $\mathbb{P}$.
2. We set $G_n := \frac{r-m}{\sigma^2}M_n$ and $Z_N = \mathcal{E}_N(G)$. Check that $(G_n)_{n\leq N}$ and $(\mathcal{E}_n(G))_{n\leq N}$ are martingales under $\mathbb{P}$.

3. Derive that $Z_n = \mathcal{E}_n(G)$ (with the notations of 3.5.1). Compute α_n.
4. We define $\mathbb{P}^*$ according to (3.4). Check (with Girsanov formula) that $(V_n - U_n)_{n \leq N}$ is a martingale under $\mathbb{P}^*$.
5. Conclude that $\mathbb{P}^*$ is a risk neutral probability (think to Proposition 3.1).

Chapter 4

European option pricing

The goal: to compute the price of a european option in a B-S market.

4.1 Problematic

We want to compute the fair price of a european option with payoff f at maturity N.

Examples of european options:

- european call: $f(\omega) = (S_N(\omega) - K)_+$
- european put: $f(\omega) = (K - S_N(\omega))_+$
- Collar option: $f(\omega) = \min(S_N(\omega) - K_1, K_2)$
- "look-back" option: $f(\omega) = S_N(\omega) - \min(S_0(\omega), \ldots, S_N(\omega))$

Definition 4.1 (Hedging portfolio) *A hedging portfolio (or strategy)* Π, *is a portfolio such that its value at maturity is larger than the payoff* f, *namely*

$$X_N^\Pi(\omega) \geq f(\omega) \quad \forall \omega \in \Omega.$$

The fair price of an option will then corresponds to the minimal initial value X_0^Π that can have a self-financed hedging portfolio. In short:

Definition 4.2 (Price of an option) *The fair price of a european option with payoff* f *and maturity* N *is*

$$C := \inf \left\{ X_0^\Pi \text{ such that } \begin{array}{l} -\ \Pi \text{ is self-financed} \\ -\ X_N^\Pi(\omega) \geq f(\omega), \quad \forall \omega \in \Omega \end{array} \right\}.$$

In next section, we solve this minimization problem in an arbitrage-free and complete market.

4.2 Pricing a european option in an arbitrage-free and complete market

We assume in this section that the market is arbitrage-free and complete. According to the previous chapter, there exists in this case a unique risk neutral probability $\mathbb{P}^*$. Next result establishes the price of an option in terms of $\mathbb{P}^*$.

Theorem 4.1 Price of a european option.

1. *The price of a european option with payoff f at maturity N is*

$$C = \mathbb{E}^* \left(\varepsilon_N(U)^{-1} f \right).$$

2. *There exists a self-financed hedging strategy Π^* with initial value C. The value at time n of the portfolio Π^* is*

$$X_n^{\Pi^*} = \mathbb{E}^* \left(\varepsilon_N(U)^{-1} \varepsilon_n(U) f \mid \mathcal{F}_n \right).$$

Remarks:

- There exist theoretical formulas giving the composition $(\beta_n^*, \gamma_n^*)_{n \le N}$ of the portfolio Π^*. But they are seldom useful.
- Usually, the hard task is to compute $\mathbb{P}^*$. This may be done with the Girsanov formula. When $\mathbb{P}^*$ is known, then we can compute C at least numerically (using Monte-Carlo methods).

Proof : Let us consider a self-financed portfolio Π. Under $\mathbb{P}^*$, the discounted value of the portfolio $(X_n^\Pi / \varepsilon_n(U))$ is a martingale (Proposition 3.2). As a consequence

$$\mathbb{E}^*(\varepsilon_N(U)^{-1} X_N^\Pi) = X_0^\Pi.$$

Moreover $X_N^\Pi(\omega) \ge f(\omega)$ $\forall \omega \in \Omega$ and thus

$$X_0^\Pi \ge \mathbb{E}^*(\varepsilon_N(U)^{-1} f),$$

and therefore

$$C := \inf \left\{ X_0^\Pi \text{ such that } \begin{array}{l} -\ \Pi \text{ is self-financed} \\ -\ X_N^\Pi(\omega) \ge f(\omega), \quad \forall \omega \in \Omega \end{array} \right\} \ge \mathbb{E}^* \left(\varepsilon_N(U)^{-1} f \right). \tag{4.1}$$

Besides, the market is complete, so there exists a self-financed strategy Π^* such that $X_N^{\Pi^*}(\omega) = f(\omega)$ for any $\omega \in \Omega$. Since Π^* is self-financed, $(X_n^{\Pi^*} / \varepsilon_n(U))$ is a martingale. As a consequence

$$X_0^{\Pi^*} = \mathbb{E}^* \left(\varepsilon_N(U)^{-1} X_N^{\Pi^*} \right) = \mathbb{E}^* \left(\varepsilon_N(U)^{-1} f \right).$$

Now, Π^* is a self-financed portfolio, so due to the very definition of C we have $X_0^{\Pi^*} \geq C$. It follows that

$$C \leq \mathbb{E}^* \left(\varepsilon_N(U)^{-1} f \right). \tag{4.2}$$

Combining (4.1) and (4.2) we get

$$C = \mathbb{E}^* \left(\varepsilon_N(U)^{-1} f \right).$$

We have proved the first part of the theorem. Concerning the second part, note that Π^* suits. Moreover, since $(X_n^{\Pi^*}/\varepsilon_n(U))$ is a martingale under $\mathbb{P}^*$ we have

$$\frac{X_n^{\Pi^*}}{\varepsilon_n(U)} = \mathbb{E}^* \left(\varepsilon_N(U)^{-1} X_N^{\Pi^*} \mid \mathcal{F}_n \right) = \mathbb{E}^* \left(\varepsilon_N(U)^{-1} f \mid \mathcal{F}_n \right).$$

We get the value of Π^* at time n in multiplying the equality by $\varepsilon_n(U)$, and then make $\varepsilon_n(U)$ enter into the conditional expectation ($\varepsilon_n(U)$ is $\mathcal{F}_n$-measurable). □

4.3 And what for an incomplete market?

What happens when the market is incomplete? In this case, there is no unique fair price, but a *spread* of fair price. Let us define the following two quantities

$$\begin{aligned}
\text{ask} : \quad C^+ &:= \inf \left\{ X_0^\Pi \text{ such that } \begin{array}{l} - \Pi \text{ is self-financed} \\ - X_N^\Pi(\omega) \geq f(\omega), \quad \forall \omega \in \Omega \end{array} \right\} \\
\text{bid} : \quad C^- &:= \sup \left\{ X_0^\Pi \text{ such that } \begin{array}{l} - \Pi \text{ is self-financed} \\ - X_N^\Pi(\omega) \leq f(\omega), \quad \forall \omega \in \Omega \end{array} \right\}
\end{aligned}$$

First note that when $\mathbb{P}^*$ is risk neutral:

$$C^- \leq \mathbb{E}^* \left(\varepsilon_N(U)^{-1} f \right) \leq C^+. \tag{4.3}$$

Indeed, if Π is a self-financed portfolio, $(X_n^\Pi/\varepsilon_n(U))$ is a martingale under $\mathbb{P}^*$ (Proposition 3.2), so $X_0^\Pi = \mathbb{E}^* \left(\varepsilon_N(U)^{-1} X_N^\Pi \right)$. In particular:

- if $X_N^\Pi \geq f$, then $X_0^\Pi \geq \mathbb{E}^* (\varepsilon_N(U)^{-1} f)$ and thus $C^+ \geq \mathbb{E}^* (\varepsilon_N(U)^{-1} f)$,
- if $X_N^\Pi \leq f$, then $X_0^\Pi \leq \mathbb{E}^* (\varepsilon_N(U)^{-1} f)$ and thus $C^- \leq \mathbb{E}^* (\varepsilon_N(U)^{-1} f)$.

Note that when the market is complete, then $C^+ = C^-(= C)$. Indeed, we have seen the existence of a self-financed portfolio Π^* such that $X_0^{\Pi^*} = C^+ = \mathbb{E}^* (\varepsilon_N(U)^{-1} f)$ et $X_N^{\Pi^*} = f$. Besides Π^* also fulfills the condition of C^-, so $C^+ = X_0^{\Pi^*} \leq C^-$. Finally $C^+ = C^-$, since we always have $C^- \leq C^+$ (see (4.3)).

In view of (4.3), we have in an incomplete market

$$C^- \leq \inf_{\mathbb{P}^* \text{ risk neutral}} \mathbb{E}^* \left(\varepsilon_N(U)^{-1} f \right) \leq \sup_{\mathbb{P}^* \text{ risk neutral}} \mathbb{E}^* \left(\varepsilon_N(U)^{-1} f \right) \leq C^+.$$

This means that usually $C^- < C^+$: the spread is positive.

What are the meaning of C^- and C^+?

- imagine an option with price $x > C^+$: the seller can then earn money without risk. Indeed, at the end, he will have at least $(x - C^+) \times \mathbb{E}^*(\varepsilon_N(U)) > 0$.
- imagine an option with price $x < C^-$: the seller is sure to loose money. Indeed, he will loose at least $(C^- - x) \times \mathbb{E}^*(\varepsilon_N(U)) > 0$.

We thus conclude that $[C^-, C^+]$ corresponds to the fair prices.

It is actually interesting to develop other strategies in incomplete market, see for example [3, 4].

4.4 Exercises

4.4.1 Cox-Ross-Rubinstein model

We assume that the interest rates are constant, i.e. $r_n = r \geq 0$ and that the return $(\rho_n)_{n\leq N}$ are i.i.d. We also assume that they can only take two values a and b. Our aim is to compute the price and hedging of an option with payoff $f := g(S_N)$ and maturity N.

1. Prove that $\mathbb{P}^*$ is risk neutral $\iff \mathbb{E}^*(\rho_n) = r$
   $\iff p^* := \mathbb{P}^*(\rho_n = b) = \frac{r-a}{b-a}$

2. Compute the probability $\mathbb{P}^*\left(S_N = S_0(1+b)^k(1+a)^{N-k}\right)$ in terms of p^*. Conclude that $C = (1+r)^{-N} F_N^*(S_0)$ where
$$F_N^*(x) := \sum_{k=0}^{N} g\left(x(1+b)^k(1+a)^{N-k}\right) C_N^k \, p^{*k}(1-p^*)^{N-k}.$$

3. Assume that π^* is the optimal hedging strategy. Check that its value at time n is
$$X_n^{\pi^*} = (1+r)^{-(N-n)} F_{N-n}^*(S_n).$$

4. Assume that we are at time $n-1$. We know $S_0, S_1, \ldots, S_{n-1}$ and we shall choose β_n^* et γ_n^*. Check that they must satisfy:
$$\begin{cases} \beta_n^* B_0(1+r)^n + \gamma_n^* S_{n-1}(1+a) = (1+r)^{-(N-n)} F_{N-n}^*(S_{n-1}(1+a)) \\ \beta_n^* B_0(1+r)^n + \gamma_n^* S_{n-1}(1+b) = (1+r)^{-(N-n)} F_{N-n}^*(S_{n-1}(1+b)). \end{cases}$$

5. Derive that
$$\gamma_n^* = (1+r)^{-(N-n)} \frac{F_{N-n}^*(S_{n-1}(1+b)) - F_{N-n}^*(S_{n-1}(1+a))}{(b-a)S_{n-1}}$$
$$\text{and } \beta_n^* = \frac{X_{n-1}^{\pi^*} - \gamma_n^* S_{n-1}}{B_0(1+r)^{n-1}} = \frac{(1+r)^{-(N-n+1)} F_{N-n+1}^*(S_{n-1}) - \gamma_n^* S_{n-1}}{B_0(1+r)^{n-1}}.$$

6. In the case of a european call, viz when $g(S_N) = (S_N - K)_+$, check that

$$C = S_0 \mathbb{B}(k_0, N, p') - (1+r)^{-N} K \, \mathbb{B}(k_0, N, p^*)$$

$$\text{where} \qquad p' = \frac{1+b}{1+r} p^*, \qquad \mathbb{B}(k_0, N, p) = \sum_{k=k_0}^{N} C_N^k p^k (1-p)^{N-k}$$

and $k_0 := \min\{k \in \mathbb{N} : S_0(1+a)^{N-k}(1+b)^k > K\} = 1 + \left[\log \frac{K}{S_0(1+a)^N} / \log \frac{1+b}{1+a}\right]$.

4.4.2 Black-Scholes and Merton models

We will now rely the Cox-Ross-Rubinstein model to the Black-Scholes and Merton models, in letting the time-step tends to zero (continuous limit). Assume that $N = T/\Delta$, $r = \rho\Delta$, $a = -\sigma\sqrt{\Delta}$ and $b = \sigma\sqrt{\Delta}$. We investigate the limit $\Delta \to 0$ with $T = N\Delta$ constant.

1. Use the central limit theorem to prove

$$\mathbb{B}(k_0, N, p^*) \overset{\Delta \to 0}{\sim} \Phi\left(\frac{Np^* - k_0}{\sqrt{Np^*(1-p^*)}}\right)$$

$$\mathbb{B}(k_0, N, p') \overset{\Delta \to 0}{\sim} \Phi\left(\frac{Np' - k_0}{\sqrt{Np'(1-p')}}\right)$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} \, dt$.

2. Check then that

$$\begin{aligned}
k_0 &\sim \frac{\log(K/S_0) + T\sigma/\sqrt{\Delta}}{2\sigma\sqrt{\Delta}} \\
Np^* &\sim \frac{T(\rho - \sigma^2/2) + T\sigma/\sqrt{\Delta}}{2\sigma\sqrt{\Delta}} \\
Np' &\sim \frac{T(\rho + \sigma^2/2) + T\sigma/\sqrt{\Delta}}{2\sigma\sqrt{\Delta}} \\
\sqrt{Np^*(1-p^*)} &\sim \frac{1}{2}\sqrt{\frac{T}{\Delta}} \\
\sqrt{Np'(1-p')} &\sim \frac{1}{2}\sqrt{\frac{T}{\Delta}} \\
(1+r)^{-N} &\sim e^{-\rho T}.
\end{aligned}$$

3. Conclude that

$$C \overset{\Delta \to 0}{\sim} S_0 \Phi\left(\frac{T(\rho + \sigma^2/2) + \log(S_0/K)}{\sigma\sqrt{T}}\right) - Ke^{-\rho T} \Phi\left(\frac{T(\rho - \sigma^2/2) + \log(S_0/K)}{\sigma\sqrt{T}}\right).$$

We find here the Black-Scholes formula (see Chapter 8).

When one takes $a = -\sigma\Delta$ and b constant, then the limit model is the Merton model based on a Poisson process. In this case $C \to S_0 P_1 - K e^{-\rho T} P_2$ where

$$
\begin{aligned}
P_1 &= \sum_{i=k_0}^{\infty} \frac{[(1+b)(\rho+\sigma)T/b]^i}{i!} \exp[-(1+b)(\rho+\sigma)T/b)] \\
P_2 &= \sum_{i=k_0}^{\infty} \frac{[(\rho+\sigma)T/b]^i}{i!} \exp[-(\rho+\sigma)T/b)]
\end{aligned}
$$

Chapter 5

American option pricing

The goal: to compute the price of an american option and the optimal time of exercise.

5.1 Problematic

In this chapter, we will consider an arbitrage-free and complete B-S market. According to the results of Chapter 3, this implies the existence of a unique risk neutral probability $\mathbb{P}^*$. We focus henceforth on an american option with maturity N and payoff f_n at time $n \leq N$.
Example:

- american call: $f_n(\omega) = (S_n(\omega) - K)_+$
- american put: $f_n(\omega) = (K - S_n(\omega))_+$
- russian option: $f_n(\omega) = \sup_{k \leq n} S_k(\omega)$

The owner of the option can exercise his option at any time before time N. He then receives f_n. For hedging, the seller thus must have a portfolio Π such that

$$X_n^{\Pi}(\omega) \geq f_n(\omega) \quad \forall n \geq N, \ \forall \omega \in \Omega. \tag{5.1}$$

Definition 5.1 *An hedging portfolio (or strategy) is a portfolio Π fulfilling (5.1).*

As for european options, the fair price of an american option corresponds to the minimal initial value that can have a self-financed hedging portfolio. In short:

Definition 5.2 *The price of an american option is*

$$C := \inf \left\{ X_0^{\Pi} \text{ such that } \begin{array}{l} - \ \Pi \text{ is self-financed} \\ - \ X_n^{\Pi}(\omega) \geq f_n(\omega), \quad \forall n \leq N, \ \forall \omega \in \Omega \end{array} \right\} \tag{5.2}$$

The goal of this chapter is to

- compute C,
- hedge our portfolio in an optimal way,
- determine the optimal (rational) time for exercising the option.

Write τ^{exc} for the exercise time. The time τ^{exc} will depend on the evolution of the market, so it is a random variable. Besides, the owner of the option cannot predict the future. When he decides to exercise his option, he takes his decision only on the information available at that time. This exactly means that τ^{exc} is a stopping time.

5.2 Pricing on american option

Assume that Π is a self-financed hedging portfolio. According to Proposition 3.2, $(X_n^\Pi/\varepsilon_n(U))$ is a martingale under $\mathbb{P}^*$. We also have seen before that τ^{exc} is a stopping time bounded by N. Optional stopping theorem thus ensures that

$$\mathbb{E}^*\left(\varepsilon_{\tau^{\text{exc}}}(U)^{-1}X^\Pi_{\tau^{\text{exc}}}\right) = \mathbb{E}^*\left(\varepsilon_0(U)^{-1}X_0^\Pi\right) = X_0^\Pi.$$

Moreover $X_n^\Pi \geq f_n$ for any $n \leq N$, so

$$X_0^\Pi \geq \mathbb{E}^*\left(\varepsilon_{\tau^{\text{exc}}}(U)^{-1}f_{\tau^{\text{exc}}}\right).$$

Since this must be true for any exercise time τ^{exc}, we must have

$$X_0^\Pi \geq \sup_{\substack{\tau \text{ s.t.} \\ \tau \leq N}} \mathbb{E}^*\left(\varepsilon_\tau(U)^{-1}f_\tau\right)$$

where "s.t." means "stopping time". As a consequence the price C defined by (5.2) admits for lower bound

$$C \geq \sup_{\substack{\tau \text{ s.t.} \\ \tau \leq N}} \mathbb{E}^*\left(\varepsilon_\tau(U)^{-1}f_\tau\right). \tag{5.3}$$

We will actually see that previous inequality is an equality.

Theorem 5.1 Price of an american option.

The price of an american option with maturity N and payoff f_n is

$$C = \sup_{\substack{\tau \text{ s.t.} \\ \tau \leq N}} \mathbb{E}^*\left(\varepsilon_\tau(U)^{-1}f_\tau\right)$$

where "s.t." means "stopping time" and where $\mathbb{E}^*$ represents the expectation under $\mathbb{P}^*$.

Formula (5.3) already gives us the inequality

$$C \geq \sup_{\substack{\tau \text{ s.t.} \\ \tau \leq N}} \mathbb{E}^* \left(\varepsilon_\tau(U)^{-1} f_\tau \right).$$

Thus to prove the theorem, all we need is to find a self-financed hedging portfolio Π^* with initial value

$$X_0^{\Pi^*} = \sup_{\substack{\tau \text{ s.t.} \\ \tau \leq N}} \mathbb{E}^* \left(\varepsilon_\tau(U)^{-1} f_\tau \right).$$

This work will be done at section 5.5. Before, we will try to better understand the optimisation problem appearing in Theorem 5.1.

5.3 Dynamic Programing Principle

The dynamic programing principle will enable us to compute the price given in Theorem 5.1. We write henceforth $X_n := \varepsilon_n(U)^{-1} f_n$.

Theorem 5.2 Dynamic Programing Principle.
We define the sequence $(Y_n)_{n \leq N}$ by a backward iteration

$$Y_N := X_N \quad \text{and} \quad Y_n := \max\left(X_n, \mathbb{E}^*\left(Y_{n+1} \mid \mathcal{F}_n\right)\right) \text{ for } n \quad \textit{going from } N-1 \quad \textit{to } 0.$$

We also set

$$T_n := \inf\{k \in [n, N], \ \textit{such that } X_k = Y_k\}.$$

Then, for any stopping time τ with value between n and N

$$\mathbb{E}^*\left(X_{T_n} \mid \mathcal{F}_n\right) = Y_n \geq \mathbb{E}^*\left(X_\tau \mid \mathcal{F}_n\right). \tag{5.4}$$

In particular,

$$Y_0 = \mathbb{E}^*\left(X_{T_0}\right) = \sup_{\substack{\tau \text{ s.t.} \\ \tau \leq N}} \mathbb{E}^*\left(X_\tau\right). \tag{5.5}$$

Proof :
a) We first prove that (5.4) implies (5.5). Taking the expectation of (5.4) gives

$$\mathbb{E}^*\left(X_{T_n}\right) = \mathbb{E}^*\left(Y_n\right) \geq \mathbb{E}^*\left(X_\tau\right).$$

Specifying previous bound for $n = 0$, gives

$$\mathbb{E}^*\left(X_{T_0}\right) = \mathbb{E}^*\left(Y_0\right) \geq \mathbb{E}^*\left(X_\tau\right)$$

for any $\tau \leq N$. Moreover $\mathbb{E}^*\left(Y_0\right) = Y_0$, so that

$$\mathbb{E}^*\left(X_{T_0}\right) = Y_0 \geq \sup_{\substack{\tau \text{ s.t.} \\ \tau \leq N}} \mathbb{E}^*\left(X_\tau\right).$$

To conclude, since T_0 is himself a stopping time, the inequality turns to be an equality.

b) Let us prove (5.4). First, we check that $Y_n = \mathbb{E}^*\left(X_{T_n} \mid \mathcal{F}_n\right)$. In view of the very definition of T_n, for any $k \in [n, T_n(\omega) - 1]$, we have $Y_k(\omega) = \mathbb{E}^*\left(Y_{k+1} \mid \mathcal{F}_k\right)(\omega)$. An iteration thus enforces $Y_n(\omega) = \mathbb{E}^*\left(X_{T_n} \mid \mathcal{F}_n\right)(\omega)$.

We now prove $Y_n \geq \mathbb{E}^*\left(X_\tau \mid \mathcal{F}_n\right)$ by a backward iteration. First, note that it holds true for $n = N$. We now assume that (5.4) holds true for any $n \geq k+1$. For any stopping time τ taking value in $[k, N]$, we have

$$\begin{aligned} \mathbb{E}^*\left(X_\tau \mid \mathcal{F}_k\right) &= \mathbf{1}_{\tau=k}\mathbb{E}^*\left(X_\tau \mid \mathcal{F}_k\right) + \mathbf{1}_{\tau>k}\mathbb{E}^*\left(X_\tau \mid \mathcal{F}_k\right) \\ &= \mathbf{1}_{\tau=k}X_k + \mathbb{E}^*\left(X_\tau \mathbf{1}_{\tau>k} \mid \mathcal{F}_k\right). \end{aligned}$$

Note that $X_\tau \mathbf{1}_{\tau>k} = X_{\tau'}\mathbf{1}_{\tau>k}$, where

$$\tau' = \begin{cases} \tau & \textit{when} \quad \tau \geq k+1 \\ k+1 & \textit{when} \quad \tau = k \end{cases}$$

is also a stopping time. We then have

$$\begin{aligned} \mathbb{E}^*\left(X_\tau \mid \mathcal{F}_k\right) &= \mathbf{1}_{\tau=k}X_k + \mathbf{1}_{\tau>k}\mathbb{E}^*\left(X_{\tau'} \mid \mathcal{F}_k\right) \\ \mathbb{E}^*\left(X_\tau \mid \mathcal{F}_k\right) &= \mathbf{1}_{\tau=k}X_k + \mathbf{1}_{\tau>k}\mathbb{E}^*\left(\mathbb{E}^*\left(X_{\tau'} \mid \mathcal{F}_{k+1}\right) \mid \mathcal{F}_k\right). \end{aligned}$$

Furthermore, we have assumed that (5.4) holds true for $n = k+1$, so that $\mathbb{E}^*\left(X_{\tau'} \mid \mathcal{F}_{k+1}\right) \leq Y_{k+1}$. Putting pieces together gives

$$\begin{aligned} \mathbb{E}^*\left(X_\tau \mid \mathcal{F}_k\right) &\leq \mathbf{1}_{\tau=k}X_k + \mathbf{1}_{\tau>k}\mathbb{E}^*\left(Y_{k+1} \mid \mathcal{F}_k\right) \\ &\leq \max(X_k, \mathbb{E}^*\left(Y_{k+1} \mid \mathcal{F}_k\right)) \;=\; Y_k. \end{aligned}$$

We thus have checked that (5.4) holds true for $n = k$. A descending iteration ensures that (5.4) holds true for any n. □

Remark: the definition of Y_n involves the conditional expectation $\mathbb{E}^*\left(Y_{n+1} \mid \mathcal{F}_n\right)$. When the risk neutral probability is of the form $\mathbb{P}^* = Z_N\mathbb{P}$, this conditional expectation is given by

$$\mathbb{E}^*\left(Y_{n+1} \mid \mathcal{F}_n\right) = \frac{1}{Z_n}\,\mathbb{E}\left(Y_{n+1}Z_{n+1} \mid \mathcal{F}_n\right)$$

where $Z_n = \mathbb{E}(Z_N \mid \mathcal{F}_n)$, see Exercise 3.5.1 Chapter 3.

5.4 Doob's decomposition and predictable representation

We will see in this section two results of the martingale theory, that we will use for the proof of Theorem 5.1. Let us first introduce the so-called "non-decreasing predictable processes".

Definition 5.3 *A process $(A_n)_{n\leq N}$ is a non-decreasing predictable process if*

1. *A_n is $\mathcal{F}_{n-1}$-measurable, for all $n \leq N$, viz (A_n) is predictable*
2. *$\Delta A_n(\omega) = A_n(\omega) - A_{n-1}(\omega) \geq 0$, for all $n \leq N$, $\omega \in \Omega$, i.e. (A_n) is non-decreasing.*

The first result gives the decomposition of a supermartingale as a martingale and a non-decreasing predictable process.

Theorem 5.3 Doob's decomposition.
Assume that $(Y_n)_{n\leq N}$ is a supermartingale (i.e. $Y_n \geq \mathbb{E}(Y_{n+1} \mid \mathcal{F}_n)$). Then, there exists a unique martingale $(M_n)_{n\leq N}$ and a unique non-decreasing predictable process $(A_n)_{n\leq N}$ such that

$$Y_n = M_n - A_n \quad \textit{and} \quad Y_0 = M_0.$$

Proof :

- Existence of M_n and A_n: set

$$\begin{aligned} M_n &= Y_0 - \sum_{k=1}^{n} \left[\mathbb{E}\left(Y_k \mid \mathcal{F}_{k-1}\right) - Y_k\right] \\ A_n &= \sum_{k=1}^{n} \left[Y_{k-1} - \mathbb{E}\left(Y_k \mid \mathcal{F}_{k-1}\right)\right] \end{aligned}$$

  and check that they suit.

- uniqueness: if $Y_n = M_n - A_n = M'_n - A'_n$, then

$$\Delta A'_n = \Delta A_n + \Delta M'_n - \Delta M_n. \tag{5.6}$$

  Since A_n and A'_n are $\mathcal{F}_{n-1}$-measurable

$$\mathbb{E}\left(\Delta A_n \mid \mathcal{F}_{n-1}\right) = \Delta A_n \quad \text{and} \quad \mathbb{E}\left(\Delta A'_n \mid \mathcal{F}_{n-1}\right) = \Delta A'_n.$$

  In the same way, since M_n and M'_n are martingales

$$\mathbb{E}\left(\Delta M_n \mid \mathcal{F}_{n-1}\right) = 0 \quad \text{and} \quad \mathbb{E}\left(\Delta M'_n \mid \mathcal{F}_{n-1}\right) = 0.$$

  So taking the conditional expectation $\mathbb{E}\left(\cdot \mid \mathcal{F}_{n-1}\right)$ of equality (5.6), gives $\Delta A'_n = \Delta A_n$, for all $n \leq N$. Since $A_0 = A'_0 = 0$, we have $A_n = \sum_{k=1}^{n} \Delta A_k$ and $A'_n = \sum_{k=1}^{n} \Delta A'_k$. So, in view of $\Delta A_k = \Delta A'_k$, we have $A_n = A'_n$. As a consequence, we also have $M_n = M'_n$.

□

The second result gives a so-called *predictable* representation of a martingale under $\mathbb{P}^*$.

Theorem 5.4 Predictable representation.
Let $(M_n)_{n\le N}$ be a martingale under the unique risk neutral probability $\mathbb{P}^*$. Then, there exists a predictable process $(\alpha_n)_{n\le N}$ (i.e. α_n is $\mathcal{F}_{n-1}$-measurable) such that

$$M_n(\omega) = M_0 + \sum_{k=1}^{n} \alpha_k(\omega)\Delta\left(\varepsilon_k(U)^{-1}S_k\right)(\omega).$$

Proof : Set $f = \varepsilon_N(U)M_N$. Since the market is complete, there exists a self-financed portfolio $\Pi = (\beta_n, \gamma_n)_{n\le N}$ such that $X_N^\Pi = f$. Moreover (M_n) is a martingale under $\mathbb{P}^*$ as well as $(\varepsilon_n(U)^{-1}X_n^\Pi)$ (Proposition 3.2), so

$$\mathbb{E}^*\left(M_N \mid \mathcal{F}_n\right) = M_n \quad \text{et} \quad \mathbb{E}^*\left(\varepsilon_N(U)^{-1}X_N^\Pi \mid \mathcal{F}_n\right) = \varepsilon_n(U)^{-1}X_n^\Pi.$$

Since $X_N^\Pi = f = \varepsilon_N(U)M_N$, we have $M_n = \varepsilon_n(U)^{-1}X_n^\Pi$ for all $n \le N$. Therefore, according to next lemma, we only need to choose $\alpha_k = \beta_k$ to conclude (remind that β_k is $\mathcal{F}_{k-1}$-measurable). □

Lemma 5.1 *Assume that $\Pi = (\beta_n, \gamma_n)_{n\le N}$ is a (generic) self-financed portfolio. Then*

$$\frac{X_n^\Pi}{\varepsilon_n(U)} = X_0^\Pi + \sum_{k=1}^{n} \gamma_k\,\Delta\left(\frac{S_k}{\varepsilon_k(U)}\right).$$

Proof : (of the lemma). Remind that a portfolio Π is self-financed if

$$\begin{cases} X_n^\Pi &= \beta_n B_n + \gamma_n S_n \\ X_{n-1}^\Pi &= \beta_{n-1}B_{n-1} + \gamma_{n-1}S_{n-1} = \beta_n B_{n-1} + \gamma_n S_{n-1} \qquad \text{(self-financing)} \end{cases}$$

As a consequence

$$\begin{aligned} \Delta\left(\frac{X_n^\Pi}{\varepsilon_n(U)}\right) &= \frac{X_n^\Pi}{\varepsilon_n(U)} - \frac{X_{n-1}^\Pi}{\varepsilon_{n-1}(U)} \\ &= \beta_n B_0 + \gamma_n\frac{S_n}{\varepsilon_n(U)} - \beta_n B_0 - \gamma_n\frac{S_{n-1}}{\varepsilon_{n-1}(U)} \\ &= \gamma_n\left(\frac{S_n}{\varepsilon_n(U)} - \frac{S_{n-1}}{\varepsilon_{n-1}(U)}\right). \end{aligned}$$

This proves the lemma. □

5.5 Proof of theorem 5.1

We set

$$\Lambda := \sup_{\substack{\tau \text{ s.t.} \\ \tau \le N}} \mathbb{E}^*\left(\varepsilon_\tau(U)^{-1}f_\tau\right).$$

All we need to prove Theorem 5.1, is to find a self-financed hedging portfolio with initial value Λ.

Let us consider the sequence (Y_n) of the dynamic programing principle. In view of

$$Y_n := \max\left(X_n, \mathbb{E}^*\left(Y_{n+1} \mid \mathcal{F}_n\right)\right),$$

we have $Y_n \geq \mathbb{E}^*\left(Y_{n+1} \mid \mathcal{F}_n\right)$, so $(Y_n)_{n\leq N}$ is a supermartingale. We write $Y_n = M_n - A_n$ for its Doob decomposition and

$$M_n = M_0 + \sum_{k=1}^{n} \alpha_k \Delta\left(\varepsilon_k(U)^{-1} S_k\right) \tag{5.7}$$

for the predictable representation of (M_n). Note that $M_0 = Y_0 = \Lambda$.
We will construct a self-financed hedging portfolio $\Pi^* = (\beta_n^*, \gamma_n^*)_{n\leq N}$ with initial value Λ. First, we set $\gamma_n^* = \alpha_n$ for all $n \leq N$. The initial value of the portfolio is $\beta_1^* B_0 + \gamma_1^* S_0$. This value is equal to Λ, when we set

$$\beta_1^* := \left(\Lambda - \gamma_1^* S_0\right) / B_0.$$

Then, the self-financing condition enforces that $X_{n-1}^{\Pi^*} = \beta_n^* B_{n-1} + \gamma_n^* S_{n-1}$, which in turns enforces that

$$\beta_n^* := \left(X_{n-1}^{\Pi^*} - \gamma_n^* S_{n-1}\right) / B_{n-1}.$$

The sequence (β_n^*) is thus defined by iteration. By construction, the portfolio Π^* is self-financed and its initial value is Λ. It remains to check that it is also an hedging portfolio.

In view of Lemma 5.1

$$\varepsilon_n(U)^{-1} X_n^{\Pi^*} = \Lambda + \sum_{k=1}^{n} \gamma_k^* \, \Delta\left(\varepsilon_k(U)^{-1} S_k\right).$$

Keep in mind that $\gamma_k^* = \alpha_k$ and compare previous equality to (5.7). It turns out that $M_n = \varepsilon_n(U)^{-1} X_n^{\Pi^*}$. In particular

$$X_n^{\Pi^*} = \varepsilon_n(U) M_n \geq \varepsilon_n(U) Y_n \qquad \text{since } M_n = Y_n + A_n \text{ with } A_n \geq 0.$$

Now according to the dynamic programing principle (5.4)

$$\varepsilon_n(U) Y_n = \sup_{\substack{\tau \text{ s.t.} \\ n \leq \tau \leq N}} \mathbb{E}^*\left(\varepsilon_\tau(U)^{-1} \varepsilon_n(U) f_\tau \mid \mathcal{F}_n\right) \geq f_n$$

(the latter equality comes from $f_n = \mathbb{E}^*\left(\varepsilon_\tau(U)^{-1}\varepsilon_n(U) f_\tau \mid \mathcal{F}_n\right)$ for $\tau = n$). Putting pieces together, we conclude that $X_n^{\Pi^*} \geq f_n$ for all $n \leq N$. The portfolio Π^* is then an hedging portfolio.
Remark: $T_0 = \inf\{k \leq N,$ such that $\varepsilon_k(U)^{-1} f_k = Y_k\}$ is the optimal time of exercise. Indeed, according to (5.5), for all stopping time τ bounded by N

$$\mathbb{E}^*\left(\varepsilon_\tau(U)^{-1} f_\tau\right) \leq \mathbb{E}^*\left(\varepsilon_{T_0}(U)^{-1} f_{T_0}\right).$$

5.6 Exercises

Assume that $\varepsilon_n(U)^{-1}f_n = g(Y_n)M_n$ with M_n a martingale under $\mathbb{P}^*$, $M_0 = 1$ and g such that $g(y) \leq g(y^*)$ $\forall y \in \mathbb{R}$.

1. check that $\varepsilon_n(U)^{-1}f_n \leq g(y^*)M_n$. Derive that for any stopping time $\tau \leq N$

$$\mathbb{E}^*\left(\varepsilon_\tau(U)^{-1}f_\tau\right) \leq g(y^*)\mathbb{E}^*\left(M_\tau\right) = g(y^*).$$

2. Conclude that $C \leq g(y^*)$.

3. Set $\tau^* := \inf\{n \leq N : \ Y_n = y^*\}$. We assume that with probability 1, Y_n takes the value y^* before time N. Prove in this case that $\mathbb{E}^*\left(\varepsilon_{\tau^*}(U)^{-1}f_{\tau^*}\right) = g(y^*)$ and $C = g(y^*)$. What is the optimal exercise time?

Remark: we usually do not have "with probability 1, Y_n takes the value y^* before time N". Nevertheless, when N goes to infinity the probability that Y_n takes the value y^* before time N tends to 1. Therefore, for large value of N, $g(y^*)$ is a good approximation of C and τ^* is a good approximation of the optimal exercise time.

Chapter 6

Brownian motion

The goal: to give a short introduction to continuous-time processes and especially to Brownian motions.

6.1 Continuous-time processes

A continuous-time process is a collection $(X_t)_{t\in I}$ of random variables indexed by some subinterval I of $\mathbb{R}$, often $I = [0, \infty[$. We usually ask some regularity properties to the trajectories $t \mapsto X_t(\omega)$.

Definition 6.1 *A process* $(X_t)_{t\geq 0}$ *is so-called* "continuous" *(respectively* "left-continuous"*) when for any* $\omega \in \Omega$ *the trajectories* $t \mapsto X_t(\omega)$ *are continuous (resp. left-continuous).*

As in discrete time, we call filtration, a collection $(\mathcal{F}_t)_{t\geq 0}$ of nested σ-algebras

$$\mathcal{F}_0 \subset \cdots \subset \mathcal{F}_t \subset \cdots$$

Typically, $\mathcal{F}_t$ will correspond to the information given by a process X up to time t, namely $\mathcal{F}_t = \sigma(X_s,\ s \leq t)$.

Definition 6.2 *A process* Y *is so-called* $\mathcal{F}_t$*-adapted, if for any* $t \geq 0$ *the random variable* Y_t *is* $\mathcal{F}_t$*-measurable. .*

Example: consider a continuous process X. Set $Y_t = \sup_{s\leq t} X_s$ and $\mathcal{F}_t = \sigma(X_s,\ s \leq t)$. Then, the process Y is $\mathcal{F}_t$-adapted.

6.2 Brownian motion

Brownian motion is one of the basic process for modeling in continuous time. In this section we will define the Brownian motion and state its basic properties. No proof are given here, we refer the interested reader to [2] for (much) more details.

6.2.1 Gaussian law

A *Gaussian* (or *normal*) real random variable with mean m and variance t is a real random variable whose law admits a density against the Lebesgue measure given by

$$\frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-m)^2}{2t}\right).$$

We write henceforth $\mathcal{N}(m,t)$ for the law of a Gaussian random variable with mean m and variance t. We remind some properties of Gaussian random variables.

Properties 6.1 (of Gaussian laws)

1. *the characteristic function of a $\mathcal{N}(m,t)$ random variable X is*

$$\mathbb{E}(\exp(izX)) = \exp(izm - tz^2/2), \quad \forall z \in \mathbb{C}.$$

2. *If X follows a $\mathcal{N}(m,t)$ law then $a+bX$ follows a $\mathcal{N}(a+bm, b^2t)$ law.*
3. *If X and Y are two independent random variables with respectively $\mathcal{N}(a,t)$ and $\mathcal{N}(b,s)$ law then $X+Y$ follows a $\mathcal{N}(a+b, t+s)$ law.*

6.2.2 Definition of Brownian motion

Definition 6.3 *A Brownian motion (or Wiener process) $W = (W_t)_{t\geq 0}$ is a real-valued process such that*

1. *W is continuous and $W_0 = 0$,*
2. *for any positive t, the random variable W_t follows a $\mathcal{N}(0,t)$ law,*
3. *For any $s,t > 0$, the increment $W_{t+s} - W_t$ is independent of $\mathcal{F}_t$ and follows a $\mathcal{N}(0,s)$ law.*

We admit the existence of such a process.

Warning! a continuous process X such that for any time $t > 0$, the variable X_t follows a $\mathcal{N}(0,t)$ law is not necessary a Brownian motion. Think, example given, to $X_t = \sqrt{t}\,N$, with N distributed according to the $\mathcal{N}(0,1)$ law. The process $(X_t)_{t\geq 0}$ fulfills the two first conditions but not the last one.

What does a Brownian motion looks like? Next result states that a Brownian motion is in some sense "a continuous-time random walk".

Theorem 6.1 Donsker's principle
Assume that $(X_i)_{i\in\mathbb{N}}$ is a sequence of i.i.d. random variables such that $\mathbb{E}(X_i) = 0$ and $\mathbb{E}(X_i^2) = 1$. Set $S_n = X_1 + \cdots + X_n$. Then the following convergence holds

$$\left(\frac{S_{[nt]}}{\sqrt{n}}\right)_{t\geq 0} \xrightarrow{(d)} (W_t)_{t\geq 0}$$

where W is a Brownian motion, $[x]$ represents the integer part of x and "$\xrightarrow{(d)}$" means "convergence in distribution" (see the Appendix for a reminder on the various types of convergence)

6.2.3 Properties of Brownian motion

We list below some basic properties of Brownian motion. Again, we refer the interested reader to [2] for proofs and finer properties.

Properties 6.2 (of Brownian motion)
Assume that W is a Brownian motion. Then,

1. Symmetry: *the process $(-W_t)_{t\geq 0}$ is also a Brownian motion.*
2. Scaling: *for any $c > 0$, the process $\left(c^{-1/2}W_{ct}\right)_{t\geq 0}$ is again a Brownian motion.*
3. Roughness: *the trajectories $t \mapsto W_t(\omega)$ are nowhere differentiable.*
4. Markov property: *for any $u > 0$, the process $(W_{u+t} - W_u)_{t\geq 0}$ is again a Brownian motion, independent of $\mathcal{F}_u$.*

(Admitted)
Some properties of the Brownian motion are rather unusual. Let us give an example. Fix $a > 0$ and write T_a for the first time where the Brownian motion W hits the value a. Then, whatever $\varepsilon > 0$, the Brownian motion W will hit again *infinitely* many times the value a during the little time interval $[T_a, T_a + \varepsilon]$!

6.3 Continuous-time martingales

As in the discrete setting, martingales will play a major role in the continuous-time theory of option pricing.

Definition 6.4 *A process $(M_t)_{t\geq 0}$ is a martingale when for any $s, t \geq 0$*

$$\mathbb{E}(M_{t+s}|\mathcal{F}_t) = M_t.$$

When a martingale is continuous, the properties we have seen in the discrete setting (fundamental property, optional stopping theorem, maximal inequality, Doob's decomposition, etc) still hold true after replacing "$n \in \mathbb{N}$" by "$t \in [0, \infty[$".

6.4 Exercises

6.4.1 Basic properties

1. Assume that N follows a $\mathcal{N}(0,1)$ law and set $X_t = \sqrt{t}\,N$ for any $t \geq 0$. Check that the process X fulfills the two first properties of Brownian motion, but not the third one.

2. Check that a Brownian motion W is a martingale.

3. Set $\mathcal{F}_t = \sigma\left(W_s, s \leq t\right)$ and $L_t = W_t^2 - t$. Prove that L is a $\mathcal{F}_t$-martingale.

4. Set $\mathcal{E}_t = \exp(W_t - t/2)$. Is $\mathcal{E}_t$ a martingale?

5. Assume that M is a martingale such that $\mathbb{E}\left(M_t^2\right) < +\infty$. Check that for any $s \leq t$:

$$\mathbb{E}\left((M_t - M_s)^2 \mid \mathcal{F}_s\right) = \mathbb{E}\left(M_t^2 - M_s^2 \mid \mathcal{F}_s\right).$$

6.4.2 Quadratic variation

Henceforth, W represents a Brownian motion. For a partition $\tau = (t_1, \ldots, t_n)$, of $[0,t]$ (i.e. $0 = t_0 < t_1 < \cdots < t_n = t$), we set $\Delta t_k = t_k - t_{k-1}$ and $\Delta W_k = W_{t_k} - W_{t_{k-1}}$.

1. For $s \leq t$, check that $\mathbb{E}\left(W_t W_s\right) = \mathbb{E}\left((W_t - W_s)W_s\right) + \mathbb{E}\left(W_s^2\right) = s$.

2. Check that $\mathbb{E}\left((\Delta W_k)^2\right) = \Delta t_k$ and $\mathbb{E}\left((\Delta W_k)^4\right) = 3\left(\Delta t_k\right)^2$.

3. Write $|\tau| = \max_{k=1\ldots n} \Delta t_k$ and $< W >_t^\tau (\omega) = \sum_{k=1}^n \left(\Delta W_k(\omega)\right)^2$. Check that

$$\begin{aligned}
\mathbb{E}\left[(< W >_t^\tau - t)^2\right] &= \sum_{j=1}^n \sum_{k=1}^n \mathbb{E}\left[\left((\Delta W_j)^2 - \Delta t_j\right)\left((\Delta W_k)^2 - \Delta t_k\right)\right] \\
&= 2\sum_{k=1}^n (\Delta t_k)^2 \\
&\leq 2|\tau| \times t \overset{|\tau|\to 0}{\longrightarrow} 0\,.
\end{aligned}$$

4. Deduce from the inequality

$$< W >_t^\tau \leq \max_{k=1\ldots n} |\Delta W_k| \times \sum_{k=1}^n |\Delta W_k|$$

and the uniform continuity of the Brownian motion on $[0,t]$ that

$$\sum_{k=1}^n |\Delta W_k| \overset{|\tau|\to 0}{\longrightarrow} +\infty.$$

We say that the Brownian motion has no bounded variations.

Chapter 7

Itô calculus

The goal: to give a short introduction to Itô calculus and its special rules.

7.1 Problematic

Let us discuss briefly and informally the motivation for introducing the Itô calculus. We want to model a "continuous-time noisy signal" $X = (X_t)_{t\geq 0}$. In analogy with the discrete time, we want the evolution on a short time interval δt to be given by

$$"\delta X_t(\omega) = a(t,\omega)\delta t + \sigma(t,\omega)\epsilon_t(\omega)" \tag{7.1}$$

where ϵ_t represents some "noise" fulfilling the properties:

- ϵ_t is independent of ϵ_s for $s \neq t$, and has the same law,
- $\mathbb{E}(\epsilon_t) = 0$.

Therefore, setting "$W_t = \sum_{s\leq t} \epsilon_s$" it is natural, in view of the Donsker's principle, to assume that W_t is a Brownian motion. Equation (7.1) then turns to

$$\delta X_t(\omega) = a(t,\omega)\delta t + \sigma(t,\omega)\delta W_t(\omega)\,.$$

Unfortunately, the Brownian motion is nowhere differential, so we cannot give a meaning to this equality in terms of differentials. We will try instead to give a meaning to it in terms of integrals:

$$X_t(\omega) = X_0 + \int_0^t a(s,\omega)\,ds + \int_0^t \sigma(s,\omega)dW_s(\omega). \tag{7.2}$$

The first integral enters into the classical field of integration theory and is well-defined. But what is the meaning of the second integral?

Should W be with bounded variations, could we define the second integral in terms of the Stieljes' integral. Unfortunately, we have seen in Exercise 6.4.2 that a Brownian motion has no bounded variations. The goal of Itô's integration theory is precisely to give a mathematical meaning to (7.2).

7.2 Itô's integral

Henceforth, W is a Brownian motion and $\mathcal{F}_t = \sigma(W_s, s \le t)$.

Definition (and theorem) 7.1 - Itô's integral -
Assume that H is a left-continuous $\mathcal{F}_t$-adapted process fulfilling $\int_0^t H_s^2\, ds < \infty$ for any $t > 0$. Then, there exists a continuous process $\left(\int_0^t H_s\, dW_s\right)_{t\ge 0}$ such that

$$\sum_{i=1}^{n} H_{(i-1)t/n}\left(W_{it/n} - W_{(i-1)t/n}\right) \xrightarrow{\mathbb{P}} \int_0^t H_s\, dW_s \qquad \text{for any } t > 0, \tag{7.3}$$

where "$\xrightarrow{\mathbb{P}}$" means "convergence in probability" (see the Appendix for a reminder).
This process fulfills the equality

$$\mathbb{E}\left(\left(\int_0^t H_s\, dW_s\right)^2\right) = \int_0^t \mathbb{E}\left(H_s^2\right)\, ds\,,$$

even when this quantity takes infinite value.
Furthermore, in the case where $\int_0^t \mathbb{E}\left(H_s^2\right)\, ds < \infty$ for any $t > 0$, the Itô's integral $\left(\int_0^t H_s\, dW_s\right)_{t\ge 0}$ is a $\mathcal{F}_t$-martingale.

We admit the existence of the process $\int_0^\cdot H_s\, dW_s$ and refer the interested reader to the Appendix 9.2 for the main lines of its construction. Before investigating the properties of Itô's integral, let us stop for a comment.

Question: does the above result still hold true when setting $\sum_{i=1}^n H_{it/n}\left(W_{it/n} - W_{(i-1)t/n}\right)$ instead of $\sum_{i=1}^n H_{(i-1)t/n}\left(W_{it/n} - W_{(i-1)t/n}\right)$?

The answer is *no* in general. Indeed, takes $H = W$. According to the exercise 6.4.2

$$\sum_{i=1}^{n} W_{it/n}\left(W_{it/n} - W_{(i-1)t/n}\right) - \sum_{i=1}^{n} W_{(i-1)t/n}\left(W_{it/n} - W_{(i-1)t/n}\right) \xrightarrow{L^2} t$$

and therefore $\sum_{i=1}^n W_{it/n}\left(W_{it/n} - W_{(i-1)t/n}\right) \xrightarrow{\mathbb{P}} t + \int_0^t W_s\, dW_s$.

7.3 Itô's processes

When K is a process whose path $t \mapsto K_t(\omega)$ are left-continuous, we can define the process $\left(\int_0^t K_s\, ds\right)_{t\ge 0}$ by

$$\left(\int_0^t K_s\, ds\right)(\omega) := \int_0^t K_s(\omega)\, ds\,, \qquad \forall\, \omega,\ t\,,$$

where in the right-hand side, the integral is a classic Riemann (or Lebesgue) integral.

Definition 7.1 *We call "Ito process" a process X of the form*

$$X_t = X_0 + \int_0^t H_s\, dW_s + \int_0^t K_s\, ds,$$

where K and H are $\mathcal{F}_t$-adapted and left-continuous, and where H fulfills $\int_0^t H_s^2\, ds < \infty$, for any $t > 0$.

For short notations, we usually write $dX_t = H_t\, dW_t + K_t\, dt$.

Next fundamental Lemma says that a smooth function of an Itô process is still an Itô process.

Theorem 7.1 - Itô's formula - *Assume that X is an Itô process and $g : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ belongs to $\mathcal{C}^2$ (twice continuously differentiable). Then*

$$g(t, X_t) = g(0, X_0) + \int_0^t \frac{\partial g}{\partial x}(s, X_s)\, dX_s + \int_0^t \frac{\partial g}{\partial t}(s, X_s)\, ds + \frac{1}{2}\int_0^t H_s^2\, \frac{\partial^2 g}{\partial x^2}(s, X_s)\, ds\,.$$

Comments:

1. The term $\int_0^t \frac{\partial g}{\partial x}(s, X_s)\, dX_s$ is a short notation for
$$\int_0^t \frac{\partial g}{\partial x}(s, X_s)\, H_s\, dW_s + \int_0^t \frac{\partial g}{\partial x}(s, X_s)\, K_s\, ds\,.$$
2. The last term in the right-hand side is unusual. This odd rule of calculus is due to the fact that Brownian motion has a quadratic variation. Let us inspect this.

Proof : We only sketch the main lines of the proof of Itô's formula: our goal is just to understand the appearance of the last integral.
Set $t_i = it/n$ and for any process Y, we write $\Delta Y_{t_i} = Y_{t_i} - Y_{t_{i-1}}$. The Taylor expansion ensures that

$$\begin{aligned} g(t, X_t) &= g(0, X_0) + \sum_i \Delta g(t_i, X_{t_i}) \\ &= g(0, X_0) + \sum_i \frac{\partial g}{\partial t}(t_{i-1}, X_{t_{i-1}})\Delta t_i + \frac{\partial g}{\partial x}(t_{i-1}, X_{t_{i-1}})\Delta X_{t_i} \\ &\quad + \frac{1}{2}\sum_i \frac{\partial^2 g}{\partial x^2}(t_{i-1}, X_{t_{i-1}})\,(\Delta X_{t_i})^2 + \text{residue}, \end{aligned}$$

where the residue goes to 0 when n goes to infinity. Furthermore, note that $(\Delta X_{t_i})^2 = H_{t_{i-1}}^2\,(\Delta W_{t_i})^2 +$ "residue", and according to exercise 6.4.2 $(\Delta W_{t_i})^2 \approx \Delta t_i$. So, the convergence of Riemann sums to Riemann integrals gives

$$\frac{1}{2}\sum_i \frac{\partial^2 g}{\partial x^2}(t_{i-1}, X_{t_{i-1}})\,(\Delta X_{t_i})^2 \approx \frac{1}{2}\int_0^t H_s^2\, \frac{\partial^2 g}{\partial x^2}(s, X_s)\, ds\,.$$

Using again the convergence of Riemann sums to Riemann integrals for the first term of Taylor expansion and the approximation formula (7.3) for the second term leads to the claimed result:

$$g(t, X_t) = g(0, X_0) + \int_0^t \frac{\partial g}{\partial x}(s, X_s)\, dX_s + \int_0^t \frac{\partial g}{\partial t}(s, X_s)\, ds + \frac{1}{2}\int_0^t H_s^2 \frac{\partial^2 g}{\partial x^2}(s, X_s)\, ds\,.$$

□

Remark: The condition $g \in \mathcal{C}^2$ is necessary. Indeed, set $g(t, x) = |x|$, which is not $\mathcal{C}^2$ at 0. Let us check that Itô's formula cannot hold true for this function. Otherwise, we would have $|W_t| = \int_0^t \operatorname{sgn}(W_s)\, dW_s$, which is impossible since the left-hand side is a submartingale (Jensen formula), whereas the right-hand side is a martingale. Indeed, the right-hand side is a stochastic integral, with the integrand fulfilling $\int_0^t \mathbb{E}\left(\operatorname{sgn}(W_s)^2\right)\, ds = t < \infty$.

Examples:

1. Check that $W_t^2 = 2\int_0^t W_s\, dW_s + t$ and $W_t^3 = 3\int_0^t W_s^2\, dW_s + \int_0^t W_s\, ds$.

2. **Geometric Brownian motion -** our goal is to find a process X such that

$$dX_t = \mu X_t\, dt + \sigma\, X_t\, dW_t,$$

with $\mu, \sigma \in \mathbb{R}$. We will search this process in the form $X_t = g(t, W_t)$. Previous equation then reads

$$dX_t = \mu g(t, W_t)\, dt + \sigma g(t, W_t)\, dW_t.$$

Besides, Itô's formula says that

$$dX_t = \frac{\partial g}{\partial x}(t, W_t)\, dW_t + \left(\frac{\partial g}{\partial t}(t, W_t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, W_t)\right) dt\,.$$

So, all we need is to solve

$$\begin{aligned} \sigma g &= \frac{\partial g}{\partial x}\,, \\ \text{and} \quad \mu g &= \frac{\partial g}{\partial t} + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}\,. \end{aligned}$$

First equation gives $g(t, x) = h(t)e^{\sigma x}$, and second equation leads to $h'(t) = \left(\mu - \frac{1}{2}\sigma^2\right) h(t)$. In conclusion, we find

$$X_t = X_0 \exp\left(\sigma W_t + \left(\mu - \frac{\sigma^2}{2}\right) t\right)\,,$$

which is so-called "geometric Brownian motion".

7.4 Girsanov formula

7.4.1 Stochastic exponentials

Definition 7.2 *We associate to an Itô process $dX_t = H_t\, dW_t$, its stochastic exponential*

$$\varepsilon_t(X) := \exp\left(\int_0^t H_s\, dW_s - \frac{1}{2}\int_0^t H_s^2\, ds\right).$$

Exercises:

1. Use Itô 's formula to check that $\varepsilon_t(X)$ is solution of the stochastic differential equation

$$\frac{dY_t}{Y_t} = dX_t\,. \tag{7.4}$$

2. Derives of this equation that $\varepsilon(X)$ is a martingale as soon as

$$\int_0^t \mathbb{E}\left((H_s\varepsilon_s(X))^2\right) ds < \infty\,, \qquad \text{for any } t > 0\,.$$

Warning! the solution of (7.4) is $Y_t = \varepsilon_t(X)$ and <u>not</u> $Y_t = \exp(X_t)$ as you may have expected. This is due to the special rules of Itô calculus.

7.4.2 Girsanov formula

We state in this section a continuous-time Girsanov formula. We will not prove any result. Again, we refer the interested reader to [2] for further details.

We have seen that under some technical conditions $\varepsilon(X)$ is a martingale. In this case, $\mathbb{E}(\varepsilon_t(X)) = \mathbb{E}(\varepsilon_0(X)) = 1$. A sufficient and practical condition for this equality to hold true is the following criterion.

Novikov's Criterion -
Assume that X is an Itô process $dX_t = H_t\, dW_t$ with H fulfilling

$$\mathbb{E}\left(\exp\left(\frac{1}{2}\int_0^T H_s^2\, ds\right)\right) < \infty$$

for some positive T. Then $\mathbb{E}\left(\varepsilon_T(X)\right) = 1$.

Under the above hypotheses, we can define (as in the discrete setting) a probability measure $\mathbb{Q}$ by setting

$$\mathbb{Q}(A) = \mathbb{E}\left(\varepsilon_T(X)\, \mathbf{1}_A\right)\,. \tag{7.5}$$

Theorem 7.2 - Girsanov formula - *Assume that X fulfills the hypotheses of Novikov's criterion (for any positive t) and define $\mathbb{Q}$ by (7.5). Then, the process $\tilde{W}_t := W_t - \int_0^t H_s\,ds$, $t \leq T$, is a Brownian motion under $\mathbb{Q}$.*

As a conseqence, for any measurable function $g : \mathbb{R}^{[0,T]} \to \mathbb{R}$, we have the formula

$$\mathbb{E}\left(g\left(\tilde{W}_t;\, t \leq T\right)\right) = \mathbb{E}\left(g(W_t; t \leq T) \exp\left(-\int_0^T H_s\,dW_s - \frac{1}{2}\int_0^T H_s^2\,ds\right)\right),$$

where $\mathbb{E}$ represents the expectation under the original probability $\mathbb{P}$.

Exercise: under the hypotheses and notations of Girsanov formula, we set

$$dY_t = K_t H_t\,dt + K_t\,dW_t,$$

where K is a left-continuous adapted process.

1. Check that $dY_t = K_t\,d\tilde{W}_t$.
2. Assume furthermore that $\int_0^T \mathbb{E}(K_s^2)\,ds < \infty$. Derives from Girsanov formula that $(Y_t)_{t \leq T}$ is martingale under $\mathbb{Q}$.

7.5 Exercises

7.5.1 Scaling functions

We consider a diffusion $dX_t = \sigma(X_t)\,dW_t + b(X_t)\,dt$, with b and σ continuous. We assume that b is bounded and that there exists ϵ and M such that $0 < \epsilon \leq \sigma(x) \leq M < \infty$, for any $x \in \mathbb{R}$. We set also $\mathcal{F}_t = \sigma(W_s,\, s \leq t)$.

1. Find a function $s \in \mathcal{C}^2$ such that $(s(X_t))_{t \geq 0}$ is a martingale. This function is so-called "scaling function".
2. For any $a < X_0 < b$, we set $T_a = \inf\{t \geq 0 : X_t = a\}$ and $T_b = \inf\{t \geq 0 : X_t = b\}$. Admitting that the equality $\mathbb{E}(s(X_{T_a \wedge T_b})) = s(X_0)$ holds true, show that

$$\mathbb{P}\left(T_b < T_a\right) = \frac{s(X_0) - s(a)}{s(b) - s(a)}\,.$$

3. Assume that $\lim_{x \to \infty} s(x) < \infty$. Show that in this case $\lim_{b \to \infty} \mathbb{P}\left(T_b < T_a\right) > 0$. Derive that with positive probability the diffusion X will never reach a.

7.5.2 Cameron-Martin formula

We set $X_t = \mu t + W_t$, where W_t is a Brownian motion under $\mathbb{P}$. Fix $T > 0$ and define

$$\mathbb{Q}(A) = \mathbb{E}\left(\exp\left(-\mu W_T - \mu^2 \frac{T}{2}\right) \mathbf{1}_A\right).$$

1. Show that $(X_t, t \leq T)$ is a Brownian motion under $\mathbb{Q}$.
2. Check that for any measurable function $g : \mathbb{R}^{[0,T]} \to \mathbb{R}$, we have

$$\mathbb{E}\left(g(X_t,\, t \leq T)\right) = e^{-\mu^2 T/2}\mathbb{E}\left(e^{\mu W_T} g(W_t, t \leq T)\right).$$

3. Replace μt by $\mu_t = \int_0^t m_s \, ds$. Find $\mathbb{Q}$ such that $(X_t,\, t \leq T)$ is a Brownian motion under $\mathbb{Q}$.

Chapter 8

Black-Scholes model

The goal: to define the Black-Scholes model, and price a european option in this setting.

8.1 Setting

8.1.1 The Black-Scholes model

As in the discrete setting, we focus on a market made of two assets:

- a non risky asset B (bond),
- a risky asset S (stock).

The two assets B and S are assumed to evolves according to

$$\begin{cases} dB_t &= rB_t\,dt \\ dS_t &= S_t(\mu\,dt + \sigma\,dW_t). \end{cases}$$

The parameter r corresponds to the *interest rate* of the bond B, whereas μ and σ corresponds to the *trend* and *volatility* of the asset S.

We can give a closed form for the value of B_t and S_t. The evolution of B is driven by an ordinary differential equation, solved by

$$B_t = B_0 e^{rt}.$$

The evolution of S is driven by the stochastic differential equation of a geometric Brownian motion (see Chapter 7, Section 3, Example 2). Therefore,

$$S_t = S_0 \exp\left(\sigma W_t + \left(\mu - \frac{\sigma^2}{2}\right) t\right).$$

Throughout this chapter, $\mathcal{F}_t$ refers to $\sigma(W_u, u \le t)$. Note that according to the previous formula, we also have $\mathcal{F}_t = \sigma(S_u, u \le t)$.

8.1.2 Portfolios

A portfolio $\Pi = (\beta_t, \gamma_t)_{t\geq 0}$ made of β_t units of B and γ_t units of S has value

$$X_t^\Pi = \beta_t B_t + \gamma_t S_t.$$

We will assume henceforth that β_t and γ_t are left-continuous, and $\mathcal{F}_t$-adapted.

In the discrete setting, a portfolio was so-called "self-financed" when its fluctuation between two consecutive times was given by

$$\Delta X_n^\Pi = \beta_n \Delta B_n + \gamma_n \Delta S_n.$$

In continuous time, this condition turns to:

Definition 8.1 *A portfolio* Π *is* self-financed *when* X_t^Π *solve*

$$dX_t^\Pi = \beta_t dB_t + \gamma_t dS_t.$$

We call *discounted value* (or present value) of a portfolio Π, the value $\tilde{X}_t^\Pi := e^{-rt} X_t^\Pi$. Next lemma gives a characterization of self-financed portfolios (compare with Lemma 5.1).

Lemma 8.1 *A portfolio* Π *is self-financed* <u>*if and only if*</u> *its discounted value* $\tilde{X}_t^\Pi$ *fulfills* $d\tilde{X}_t^\Pi = \gamma_t d\tilde{S}_t$*, with* $\tilde{S}_t := e^{-rt} S_t$*.*

Proof : Itô's formula ensures that $d\tilde{X}_t^\Pi = -re^{-rt} X_t^\Pi \, dt + e^{-rt} dX_t^\Pi$. Therefore, a portfolio is self-financed if and only if

$$\begin{aligned} d\tilde{X}_t^\Pi &= -r\left(\beta_t e^{-rt} B_t + \gamma_t e^{-rt} S_t\right) dt + e^{-rt}\left(\beta_t dB_t + \gamma_t dS_t\right) \\ &= \beta_t e^{-rt} \underbrace{\left(-rB_t \, dt + dB_t\right)}_{=0} + \gamma_t \underbrace{\left(-re^{-rt} S_t \, dt + e^{-rt} dS_t\right)}_{=d\tilde{S}_t}. \end{aligned}$$

The lemma follows. □

8.1.3 Risk neutral probability

As in the discrete setting a *risk neutral probability* $\mathbb{P}^*$ is a probability equivalent to $\mathbb{P}$ (i.e. $\mathbb{P}^*(A) = 0$ iif $\mathbb{P}(A) = 0$) such that the discounted value of the stock $(\tilde{S}_t)_{t\leq T}$ is a martingale under $\mathbb{P}^*$.

Let us focus on the evolution of $\tilde{S}_t$:

$$\tilde{S}_t = \exp\left(\sigma W_t + \left(\mu - r - \frac{\sigma^2}{2}\right) t\right) = \exp\left(\sigma W_t^* - \frac{\sigma^2}{2} t\right), \tag{8.1}$$

where we have set

$$W_t^* = W_t + \frac{\mu - r}{\sigma} t.$$

Now, according to the Girsanov-Cameron-Martin formula (see Exercise 7.5.2), the process $(W_t^*)_{t\leq T}$ is a Brownian motion under the probability $\mathbb{P}^*$ defined by

$$\mathbb{P}^*(A) := \mathbb{E}\left(Z_T \mathbf{1}_A\right), \quad \text{with} \quad Z_T := \exp\left(-\frac{\mu - r}{\sigma} W_T - \left(\frac{\mu - r}{\sigma}\right)^2 \frac{T}{2}\right).$$

Since the discounted value of the stock $\left(\tilde{S}_t\right)_{t\leq T}$ is a stochastic exponential under $\mathbb{P}^*$, it is a martingale under $\mathbb{P}^*$ (as already checked in exercise 6.4.1.4).
Conclusion: the probability $\mathbb{P}^*$ is a risk neutral probability.

Comment: Compare the probability $\mathbb{P}^*$ with the risk neutral probability computed at the end of Section 3.4.

8.2 Price of a european option in the Black-Scholes model

We compute in this section, the price of a european option in the Black-Scholes model. To be more specific, we will focus on options with maturity T and payoff of the form $f(\omega) = g(S_T(\omega))$.

Definition 8.2 *A portfolio* Π *is so-called an* hedging portfolio *when* $X_T^\Pi \geq g(S_T)$ *and*

$$\int_0^T \mathbb{E}^*\left(\left(\tilde{X}_t^\Pi\right)^2\right) dt < +\infty\,. \tag{8.2}$$

The first condition is the same as in the discrete setting, while the second condition is technical.
The price C of an option will again correspond to the minimal initial value X_0^Π that can have an hedging self-financed portfolio, namely

$$C := \inf\left\{ X_0^\Pi \text{ such that } \begin{array}{l} -\ \Pi \text{ is self-financed} \\ -\ \Pi \text{ is hedging} \end{array} \right\}.$$

Again, we will solve this minimization problem, with martingale methods. First, we will bound C from below by using that the discounted value of a self-financed hedging portfolio is a martingale under $\mathbb{P}^*$. Second, we will construct a self-financed hedging portfolio Π^*, whose initial value fits with the lower bound found at the first step.

In the next result, we assume that g is piecewise $\mathcal{C}^1$ and fulfills $\mathbb{E}\left(g(S_T)^2\right) < \infty$.

Theorem 8.1 Price of a european option.

1. *The price of a european option with payoff* $g(S_T)$ *at maturity* T *is*

$$C = \mathbb{E}^*\left(e^{-rT}g(S_T)\right) = G(0, S_0)$$

where the function G *is defined by*

$$G(t,x) := e^{-r(T-t)}\int_{-\infty}^{+\infty} g\left(x\, e^{(r-\sigma^2/2)(T-t)+\sigma y}\right)\exp\left(\frac{-y^2}{2(T-t)}\right)\frac{dy}{\sqrt{2\pi(T-t)}}\,.$$

2. *There exists a self-financed hedging portfolio* Π^* *with initial value* C*. The value at time* t *of the portfolio* Π^* *is*

$$X_t^{\Pi^*} = e^{-r(T-t)}\,\mathbb{E}^*\left(g(S_T)\mid\mathcal{F}_t\right) = G(t, S_t).$$

Its composition is given by

$$\gamma_t^* = \frac{\partial G}{\partial x}(t, S_t) \quad \text{and} \quad \beta_t^* = \frac{G(t,S_t) - \gamma_t^* S_t}{B_t}\,.$$

Proof : Let us first consider an arbitrary self-financed hedging portfolio Π. According to (8.1), we have $d\tilde{S}_t = \sigma\tilde{S}_t\, dW_t^*$. Combining with Lemma 8.1, thus leads to

$$\tilde{X}_t^\Pi = \tilde{X}_0^\Pi + \sigma\int_0^t \gamma_u\tilde{S}_u\, dW_u^*\,,$$

with W^* Brownian motion under $\mathbb{P}^*$. The discounted price $\tilde{X}_t^\Pi$ of Π is therefore a stochastic integral, and since we have assumed that $\int_0^T \mathbb{E}^*\left((\tilde{X}_t^\Pi)^2\right)dt < \infty$, it is a martingale under $\mathbb{P}^*$. In particular, we have

$$\tilde{X}_0^\Pi = \mathbb{E}^*\left(\tilde{X}_T^\Pi\right) \geq \mathbb{E}^*\left(e^{-rT}g(S_T)\right),$$

where the last inequality comes from the condition $X_T^\Pi \geq g(S_T)$. It follows that $C \geq \mathbb{E}^*\left(e^{-rT}g(S_T)\right)$.

Conversely, we will show that there exists a self-financed hedging portfolio Π^* with initial value $\mathbb{E}^*\left(e^{-rT}g(S_T)\right)$. Set $M_t = \mathbb{E}^*\left(e^{-rT}g(S_T)\mid\mathcal{F}_t\right)$.

Lemma 8.2 *When* M *is defined by the above formula, its value is given by*

$$M_t = e^{-rt}G(t, S_t) = \tilde{G}(t, \tilde{S}_t),$$

with $\tilde{G}(t,x) := e^{-rt}G(t, xe^{rt})$.

Proof : (of the lemma). The second equality is straightforward. Let us prove the first one. Due to the very definition of conditional expectation, all we need is to check that:

1. the random variable $e^{-rt}G(t,S_t)$ is $\mathcal{F}_t$-measurable (which is obvious!)
2. the equality $\mathbb{E}^*(e^{-rT}g(S_T)h(S_t)) = \mathbb{E}^*(e^{-rt}G(t,S_t)h(S_t))$ holds for any measurable h.

For the second point, note that

$$g(S_T) = g\left(S_t \exp\left((r-\sigma^2/2)(T-t) + \sigma(W_T^* - W_t^*)\right)\right)$$

with $W_T^* - W_t^*$ independent of S_t and $\mathcal{N}(0, T-t)$ distributed under $\mathbb{P}^*$. Therefore

$$\begin{aligned}\mathbb{E}^*\left(e^{-rT}g(S_T)h(S_t)\right) &= \int \mathbb{P}^*(S_t \in dx) \int_{-\infty}^{+\infty} e^{-rT}h(x)\, g\left(x\, e^{(r-\sigma^2/2)(T-t)+\sigma y}\right) \frac{e^{-y^2/2(T-t)}\, dy}{\sqrt{2\pi(T-t)}} \\ &= \mathbb{E}^*\left(e^{-rt}G(t,S_t)h(S_t)\right)\end{aligned}$$

This conclude the proof of the lemma. □

Applying Itô's formula to $\tilde{G}(t,\tilde{S}_t)$, leads to

$$M_t = M_0 + \int_0^t \frac{\partial \tilde{G}}{\partial x}(u,\tilde{S}_u)\, d\tilde{S}_u + \underbrace{\int_0^t \left(\frac{\partial \tilde{G}}{\partial t} + \frac{\sigma^2 x^2}{2}\frac{\partial^2 \tilde{G}}{\partial x^2}\right)(u,\tilde{S}_u)\, du}_{=0},$$

where the second integral turns to be 0, since

$$\frac{\partial \tilde{G}}{\partial t} + \frac{\sigma^2 x^2}{2}\frac{\partial^2 \tilde{G}}{\partial x^2} = 0. \qquad \text{(check it!)} \qquad (8.3)$$

We now define Π^* by setting

$$\gamma_t^* = \frac{\partial \tilde{G}}{\partial x}(t,\tilde{S}_t) = \frac{\partial G}{\partial x}(t,S_t)$$

and $\beta_t^* = (G(t,S_t) - \gamma_t^* S_t)/B_t$. Then, the value of Π^* is $X_t^{\Pi^*} = G(t,S_t)$ and its discounted value is

$$\tilde{X}_t^{\Pi^*} = \tilde{G}(t,\tilde{S}_t) = \tilde{X}_0^{\Pi^*} + \int_0^t \gamma_u^*\, d\tilde{S}_u. \qquad \text{(Itô's formula)}$$

Therefore, according to Lemma 8.1, Π^* is self-financed. Besides, Jensen inequality enforces

$$\mathbb{E}^*\left(\left(\tilde{X}_t^{\Pi^*}\right)^2\right) = \mathbb{E}^*\left(\mathbb{E}^*\left(e^{-rT}g(S_T) \mid \mathcal{F}_t\right)^2\right) \le e^{-2rT}\mathbb{E}^*\left(g(S_T)^2\right)$$

so that condition (8.2) holds. To conclude: since $X_T^{\Pi^*} = g(S_T)$, we have constructed a self-financed hedging portfolio with initial value $G(0,S_0) = \mathbb{E}^*\left(e^{-rT}g(S_T)\right)$. □

Comment: we derive from (8.3), that G is solution of

$$\begin{cases} \frac{\partial G}{\partial t} + \frac{\sigma^2 x^2}{2}\frac{\partial^2 G}{\partial x^2} + rx\frac{\partial G}{\partial x} = rG & \text{in } [0,T]\times\mathbb{R}^+ \\ G(T,x) = g(x) & \text{for any } x > 0. \end{cases}$$

Exercise: compute the price of a call with maturity T and strike K.

Chapter 9

Appendix

9.1 Convergence of random variables

In the following, $(X_n)_{n\geq 0}$ stands for a sequence of real random variables.

9.1.1 Convergence a.s.

(X_n) is said to converge a.s. to X, when there exists a set Ω_0 such that $\mathbb{P}(\Omega_0) = 1$ and $X_n(\omega) \overset{n\to\infty}{\longrightarrow} X(\omega)$ for any $\omega \in \Omega_0$.

9.1.2 Convergence in L^2

X_n is said to converge in L^2 to X, when $\mathbb{E}\left(|X_n - X|^2\right) \overset{n\to\infty}{\longrightarrow} 0$.

9.1.3 Convergence in probability

X_n is said to converge in probability to X, when for any $\epsilon > 0$

$$\mathbb{P}\left(|X_n - X| > \epsilon\right) \overset{n\to\infty}{\longrightarrow} 0.$$

9.1.4 Convergence in distribution

X_n is said to converge in distribution to X, when for any bounded and continuous function $f : \mathbb{R} \to \mathbb{R}$ we have

$$\mathbb{E}(f(X_n)) \overset{n\to\infty}{\longrightarrow} \mathbb{E}(f(X)).$$

9.1.5 Relationships

$$\begin{array}{ccccc} \text{Convergence a.s.} & \Longrightarrow & \text{Convergence in probability} & \Longrightarrow & \text{Convergence in distribution} \\ & & \Uparrow & & \\ & & \text{Convergence } L^2 & & \end{array}$$

9.2 Construction of Itô's integral

We sketch in this section the main lines of the construction of the Itô's integral. For comprehensive proofs, we refer to [2].

9.2.1 Setting

Henceforth, W is a Brownian motion and $\mathcal{F}_t = \sigma(W_s, s \leq t)$. Fix $T \in [0, \infty]$ and write $\mathcal{M}_T$ for the space of $\mathcal{F}_t$-martingales which are continuous and start from 0 (i.e. $M_0 = 0$) with probability one and fulfill $\mathbb{E}\,(M_T^2) < \infty$.

Proposition 9.1 *$\mathcal{M}_T$ endowed with the scalar product $(M|N)_{\mathcal{M}} = \mathbb{E}\,(N_T M_T)$ is an Hilbert space.*

Proof : Checking that $(M|N)_{\mathcal{M}}$ is a scalar product is straightforward. Proving the completeness needs both Doob's inequality and Borel-Cantelli's lemma, see [2]. □

We will also introduce the linear space spanned by the left-continuous $\mathcal{F}_t$-adapted processes $(H_t)_{t\geq 0}$, fulfilling

$$\mathbb{E}\left(\int_0^T H_s^2\, ds\right) < \infty.$$

We write L_T^2 for the closure of this space endowed with the scalar product $(H|K)_L = \mathbb{E}\left(\int_0^T H_s K_s\, ds\right)$.

9.2.2 Integration of elementary processes

Elementary processes are left-continuous and bounded piecewise constant processes, i.e. processes H of the form

$$H_t(\omega) = \sum_{i=1}^n h_i(\omega)\, \mathbf{1}_{]t_{i-1},t_i]}(t)$$

where the h_i's are bounded $\mathcal{F}_{t_{i-1}}$-measurable random variables and $0 \leq t_0 < \ldots < t_n \leq T$. We write henceforth $\mathcal{E}_T$ for the space of elementary processes endowed with the scalar product $(\cdot|\cdot)_L$. For $t \leq T$ and a process H of the previous form, we define the stochastic integral as follows

$$\left(\int_0^t H_s\, dW_s\right)(\omega) := \sum_{i=1}^n h_i(\omega)\left(W_{t\wedge t_i}(\omega) - W_{t\wedge t_{i-1}}(\omega)\right),$$

with $s \wedge t := \min(s, t)$.

Proposition 9.2 *The map $(H_t)_{t\leq T} \mapsto \left(\int_0^t H_s\, dW_s\right)_{t\leq T}$ is an isometry from $\mathcal{E}_T$ to $\mathcal{M}_T$.*

Proof : First, note that the process $\left(\sum_{i=1}^n h_i\left(W_{t\wedge t_i}-W_{t\wedge t_{i-1}}\right)\right)_{t\le T}$ is continuous and takes value 0 at $t=0$. Second, it is a martingale. Indeed, fix $0\le t_k\wedge t\le s<t_{k+1}\wedge t$. Using the martingale property of Brownian motions (see exercise 6.4.1.2), we have

$$\begin{aligned}
&\mathbb{E}\left(\sum_{i=1}^n h_i\left(W_{t\wedge t_i}-W_{t\wedge t_{i-1}}\right)\mid\mathcal{F}_s\right)\\
=&\sum_{i=1}^k\mathbb{E}\left(h_i\left(W_{t\wedge t_i}-W_{t\wedge t_{i-1}}\right)\mid\mathcal{F}_s\right)+\mathbb{E}\left(h_{k+1}\left(W_{t\wedge t_{k+1}}-W_{t\wedge t_k}\right)\mid\mathcal{F}_s\right)\\
&+\sum_{i=k+2}^n\mathbb{E}\left(h_i\left(W_{t\wedge t_i}-W_{t\wedge t_{i-1}}\right)\mid\mathcal{F}_s\right)\\
=&\sum_{i=1}^k h_i\left(W_{t\wedge t_i}-W_{t\wedge t_{i-1}}\right)+h_{k+1}\left(\mathbb{E}\left(W_{t\wedge t_{k+1}}\mid\mathcal{F}_s\right)-W_{t\wedge t_k}\right)\\
&+\sum_{i=k+2}^n\mathbb{E}\left(\mathbb{E}\left(h_i\left(W_{t\wedge t_i}-W_{t\wedge t_{i-1}}\right)\mid\mathcal{F}_{t_{i-1}\wedge t}\right)\mid\mathcal{F}_s\right)\\
=&\sum_{i=1}^k h_i\left(W_{t\wedge t_i}-W_{t\wedge t_{i-1}}\right)+h_{k+1}\left(W_s-W_{t\wedge t_k}\right)+0\\
=&\sum_{i=1}^n h_i\left(W_{s\wedge t_i}-W_{s\wedge t_{i-1}}\right).
\end{aligned}$$

We thus have checked that $\left(\int_0^t H_s\,dW_s\right)_{t\le T}$ belongs to $\mathcal{M}_T$, it remains to prove that the map $H\mapsto\int_0^\cdot H_s\,dW_s$ is an isometry.
Let us compute the norm of $\int_0^\cdot H_s\,dW_s$:

$$\begin{aligned}
\left\|\int_0^\cdot H_s\,dW_s\right\|_{\mathcal{M}}^2 &= \mathbb{E}\left(\left(\int_0^T H_s\,dW_s\right)^2\right)\\
&= \sum_i\sum_j\mathbb{E}\left(h_ih_j(W_{t_i}-W_{t_{i-1}})(W_{t_j}-W_{t_{j-1}})\right).
\end{aligned}$$

For $i<j$, the variable $h_ih_j(W_{t_i}-W_{t_{i-1}})$ is $\mathcal{F}_{t_{j-1}}$-measurable. Therefore, according to the third property of Brownian motion $h_ih_j(W_{t_i}-W_{t_{i-1}})$ is independent of $W_{t_j}-W_{t_{j-1}}$. It follows that

$$\mathbb{E}\left(h_ih_j(W_{t_i}-W_{t_{i-1}})(W_{t_j}-W_{t_{j-1}})\right)=\mathbb{E}\left(h_ih_j(W_{t_i}-W_{t_{i-1}})\right)\underbrace{\mathbb{E}\left(W_{t_j}-W_{t_{j-1}}\right)}_{=0}=0.$$

When $i=j$: since h_i^2 is $\mathcal{F}_{t_{i-1}}$-measurable, it is independent of $(W_{t_i}-W_{t_{i-1}})^2$, so that

$$\begin{aligned}
\mathbb{E}\left(h_i^2(W_{t_i}-W_{t_{i-1}})^2\right) &= \mathbb{E}\left(h_i^2\right)\left((W_{t_i}-W_{t_{i-1}})^2\right)\\
&= \mathbb{E}\left(h_i^2\right)\times(t_i-t_{i-1}).
\end{aligned}$$

Putting pieces together thus gives

$$\left|\left|\int_0^{\cdot} H_s\, dW_s\right|\right|_{\mathcal{M}}^2 = \sum_i \mathbb{E}\left(h_i^2\right) \times (t_i - t_{i-1}) = \int_0^T \mathbb{E}(H_s^2) ds = ||H||_L^2 .$$

We have check that the map $H \mapsto \int_0^{\cdot} H_s\, dW_s$ is an isometry. □

9.2.3 Extension to L^2_T

Thanks to the isometry property we can extend the previous integral to the closure of $\mathcal{E}_T$ in L^2_T, which turns to be L^2_T itself.

Proposition 9.3 *The map*

$$\begin{array}{rcl} \mathcal{E}_T & \to & \mathcal{M}_T \\ H & \mapsto & \int_0^{\cdot} H_s\, dW_s \end{array}$$

can be extend in a unique way as an isometry from L^2_T *to* $\mathcal{M}_T$.

Extending this integral to processes H fulfilling only $\int_0^T H_s^2\, ds < \infty$, needs somewhat more technical arguments. We refer the interested reader to [2] for this extension as well as for the proof of the density of $\mathcal{E}_T$ in L^2_T.

Corollary 9.1 *When* $H \in L^2_T$ *and is left-continuous, the following convergence holds:*

$$\sum_{i=1}^n H_{(i-1)t/n}\left(W_{it/n} - W_{(i-1)t/n}\right) \xrightarrow{\mathbb{P}} \int_0^t H_s\, dW_s, \qquad \text{for any } t \leq T.$$

Proof : Set $H_s^{(n)} = \sum_{i=1}^n H_{(i-1)t/n}\mathbf{1}_{](i-1)t/n, it/n]}(s)$. On the one hand

$$\int_0^t H_s^{(n)}\, dW_s = \sum_{i=1}^n H_{(i-1)t/n}\left(W_{it/n} - W_{(i-1)t/n}\right).$$

On the other hand $H^{(n)}$ converges to H in L^2_t: according to the isometry property $\int_0^t H_s^{(n)}\, dW_s$ converges to $\int_0^t H_s\, dW_s$ in $\mathcal{M}_t$ and therefore in probability. □

Remark: the Itô integral $\int_0^{\cdot} H_s\, dW_s$ constructed above belongs to the space $\mathcal{M}_T$ and is therefore continuous with probability one. Since we can modify a process on a set of probability 0 without changing its law, we can choose a version of $\int_0^{\cdot} H_s\, dW_s$ which is continuous everywhere.

Bibliography

[1] Baxter M., Rennie A. *Financial calculus: an introduction to derivative pricing.* Cambridge University Press 1999, Cambridge.

[2] Karatzas I., Shreve S.E. *Brownian Motion and Stochastic Calculus.* Springer-Verlag, New-York, 1988.

[3] Mel'nikov A.V. *Financial Markets: stochastic analysis and the pricing of derivative securities.* Translations of Mathematical Monographs 1999, AMS edition, Providence.

[4] Pliska S. *Introduction to mathematical finance.* Blackwell Publishers, Oxford UK.

[5] Shiryaev A.N. *Essentials of stochastic finance: facts, models, theory.* Advanced Series on Statistical Science & Applied Probability 1999, World Scientific, Singapore.

[6] Williams D. *Probability with martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Index