Notes for “MAT519: Introduction to mathematical finance”

Reda Chhaibi *

December 31, 2014

Contents

*reda.chhaibi@math.uzh.ch

Preamble

These lecture notes are very introductory by nature, and quite *plain vanilla*. Due to the large amount of material covering the subject, these notes do not intend to be complete in any way. They are rather intended to serve as a roadmap for the course MAT519 and are largely based on the very good books of Williams [Wil06] and Lamberton-Lapeyre [LL08] - sometimes shamelessly. The keen students can complement their knowledge by looking at the very practical book of Fries [Fri07].

In order to learn mathematical finance, my general feeling is that the students are faced with two distinct challenges:

- On the one hand, one needs to understand the mathematics and more precisely the underlying probability theory. Therefore, the prerequisites for the class are a standard course in measure theory and a first probability class. The corresponding modules at the University of Zürich are "Analysis 3" and "Probability 1". This provides sufficient tools in order to deliver a course in mathematical finance in discrete time. The notions of conditional expectation and martingales, considered more advanced, will be introduced when needed.

- On the other hand, one needs to understand how financial markets are organised. To the mathematically-minded people, this is perhaps the most difficult task, as describing financial markets cannot be done with a sequence of definitions, lemmas, propositions and theorems. There is nothing canonical about the legal texts that define financial contracts. It would be much easier if one is allowed to step inside a bank. I tried filling the gap with the first section where non-mathematical notions of mathematical finance are presented. The complete reference would be the book of Hull [Hul06].

The time variable is generally denoted t and will be discrete for most of the lectures. We adopt the convention that prices are revealed exactly at times $t = 0, 1, 2, 3, \dots$ and, in between, one is allowed to strategise and rebalance portfolios.

I would like to thanks Markus Neumann, for type-setting the first lecture; and Martina Dal Borgo for her feedback. All mistakes are mine and I will gladly correct them, once pointed out.

1 Non-mathematical notions of mathematical finance

1.1 The "universal bank" structure

No two banks are organised exactly the same way. However, one can draw a general scheme of how a generic "universal bank" is structured.

Bank			
Retail	Investment banking		
Deposits, credit card...	M& A	Private equity	Capital Markets

The capitals markets division is in charge of the business that happens on financial markets. Itself is broken down to smaller divisions depending on the different existing asset classes. Some names are self-explanatory:

- FX (Foreign eXchange).
- Commodities: Oil, Metal, Grain...
- Fixed income: Credit products and interest rates.
- Equities: Products related to stocks.

We will mainly focus on the assets and financial contracts related to equities. This is the standard entrance point to mathematical finance. Other asset classes are usually the subject of more specialised classes.

1.2 Financial markets

An aspect of finance, like any specialised field, is the prevalence of jargon, i.e, specialised vocabulary. Jargon will be indicated in bold letters with a ⚠sign as follows.

⚠**Long and short positions:** If one buys an asset or enters in a financial contract, he is said to hold a *long* position. Reciprocally, if one sells an asset or offers the financial contract, he is said to be in a *short* position.

Financial markets are the platforms where assets are traded. These tradable assets are called *securities* and we distinguish between two kinds of markets depending on the level of sophistication of the securities they trade:

- On *Primary markets*, one trades basic securities like:
  - *Stocks*. We will generally denote the value of a single stock by S_t. If more are available, we will use a vector notation $\left(S_t^1, \dots, S_t^d\right)$. This value is commonly called a ⚠**spot**.
  - *Currencies.*
  - *Bonds* are products with given or predictable interest rate in the future. Two examples we will often use are:
    - The risk free bond with value:

      $$B_t = (1+r)^t$$

      It tantamounts to a standard bank account where the risk free interest rate is compounded.
    - The zero coupon bond is the bond which gives you 1\$ at time T. At time t, its value is:

      $$B^0_{t,T} = (1+r)^{-(T-t)}$$

- On a *secondary market*, more elaborate securities are traded. Because they are based on securities from the primary markets, these more sophisticated assets are called *derivatives*. Options are the derivatives we will be dealing with: financial contracts that give the buyer the possibility but not the obligation of performing a deal at or until a maturity date T. Of course, this optionality earned them the name of "options".

  As a down-to-earth example, a perfectly standard financial contract is the option of buying 10^6 gallons of kerosene at the price of 0.55$ per gallon, in a year from now (maturity=1 year). One sees how such a contract is useful to an airline company.

⚠**Bid vs. Ask:** The bid price is the price for which agents are willing to buy the asset. The ask price is the price for which agents are willing to sell the asset. The difference between the two is called the bid-ask spread. Daily-life examples are in exchange offices in airports that ask you 1.37$ for *their* euro and bid 1.25$ for *your* euro. Here the asset in question is the euro on a *US* dollar market, and the bid-ask spread is $1.37 - 1.25 = 0.12$$.

A market is said to be *liquid* if assets are easily bought and sold. In other words, at any time, one can find a buyer or a seller without having to change too much his price. This supposes plenty of offer and demand, but also a competitive environment. A corollary of high liquidity is that the bid-ask spread is very small. We assume the bid-ask spread is zero, therefore lifting any ambiguity about what is the price an asset: assets can be bought and sold a specific price called the *spot* price.

⚠**OTC vs. non-OTC:** *OTC* stands for Over The Counter and refers to unregulated financial contracts. They are unregulated in the sense that no financial official is organising the deal. Naturally, OTC contracts are generally between large financial institutions for whom default risk is minimal.

- The **forward** is an OTC agreement to buy or sell an asset at a certain price. The forward price, decided at time t for a deal at T, is $f_{t,T}$. E.g Facebook now is $S_0 = 100$$. Would you lock the forward price $f_{0,T} = 102$$ with T being a year? The answer depends on the interest rate...

- A **future** is similar to the forward but much more regulated. It is traded on a financial exchange. For example, these follow a settlement procedure called "marking to market", detailed in the exercise class. Basically, in order to reduce default risk, the invester makes an initial deposit ($\approx 70\%$) of $F_{t,T}$ a clearing house, which will give you a margin call in case the security's value drops too low.

  As a useful approximation, the future price $F_{t,T}$ is in general assumed to be equal to the forward price $f_{t,T}$.

- **Options:** Financial product that gives you the *option* (not obligation) of buying/selling at a certain price called the *strike* and which we will denote by K. The cash flow at the time of exercise is called the ⚠**payoff** and determines the option.

  - **European options:** Exercising happens at time T (called *Maturity*). ($\tau = T$). The example of a call on kerosene was a European option with strike 0.5$ and maturity a year. We write Φ_T for the cash flow at maturity for European options.
  - **American options:** Exercising can happen any time τ until T. ($\tau \in [t_0, T]$)

- **Call:** Right (optional) of *buying* a stock S at price K.
- **Put:** Right of *selling* a stock S at price K.

1.3 Cash versus physical settlements

Consider options such as the right of selling or buying at a certain price. Upon exercising such an option, one of the parties would hand in the strike's amount and would receive in exchange a physical asset - technically. Indeed, most assets are physical in essence. This is true for stocks as owning a stock means in practice owning a legal document declaring you are the owner. This is even more true for commodities, where owning 10^6 gallons of kerosene implies you need a tanker to store it. In such a case, we speak of physical settlements, historically the only kind of settlements.

A cash settlement happens when instead of receiving the physical asset, one receives its monetary value. In all the following, we will always assume cash settlements in order to equate assets and their monetary value. Notice that equating cash and physical settlements supposes high liquidity - again.

We leave it to the reader to convince himself/herself that the cash settlement of a call option is equivalent to a monetary payoff $\Phi_T = (S_T - K)^+$. In the same fashion, a cash settlement of a put option is equivalent to a monetary payoff of $\Phi_T = (K - S_T)^+$.

1.4 Arbitrage

An *arbitrage* or an opportunity of arbitrage (OA) is an opportunity of making profit without risk.

Example: The price of an i-Phone in EU is 600e? and in the US is 600\$. But the exchange rate is $1e = 1,3\$$. An arbitrage is easily found and is known as the ⚠**cash-and-carry** arbitrage with the :

1. Borrow 600\$
2. Buy the product on the US dollars on the market
3. Sell on European market: $+600e$
4. Exchange $600e \rightarrow 800\$$
5. Pay back your dept
6. Total: +200\$

If a market is *liquid*, prices move very fast to eliminate OA. The basic line of reasoning in mathematical finance is that absence of opportunity of arbitrage (AOA) forces relations between prices of forwards, futures, calls and puts on a stock. One of the goals of mathematical finance is to establish these relations. However, unlike physics, very few laws are available. The only rule in mathematical finance, is the dominance relation: Financial products with larger payoffs must have larger prices.

Axiom 1.1 (Dominance relation)**.** *Given two financial products A and B, with payoffs $\Phi_T(A), \Phi_T(B)$, prices P_A, P_B at $t = 0$*

$$AOA \Longleftrightarrow \Big(\Phi_T(A) > \Phi_T(B) \Rightarrow P_A > P_B\Big)$$

Notice that we take this relation as a working axiom, rather than a theorem under certain hypotheses. These hypotheses would be liquidity, equality of all agents in the market and perfect symmetry in the information available. Some would argue this is the work of economists, but it is certainly not the scope of this lecture.

1.5 Applications of AOA

Here, we propose two applications of the absence of opportunity of arbitrage. The first one deals with computing the forward price of a stock S_t.

Lemma 1.1 (Forward price)**.** *By AOA,* $f_{t,T} = (1+r)^T S_t = \frac{S_t}{B^0_{t,T}}$.

Proof. In the case of $f_{t,T} > (1+r)^T S_t$, perform the following strategy:

1. At time t, borrow price of S_t, buy S_t and offer a foward contract with forward price $f_{t,T}$.

2. Wait until time T.

3. At time T, hand the stock S_T, cash in $f_{t,T}$ and pay back the debt $S_t\,(1+r)^{T-t}$.

The final value of such a strategy is $V_T = -S_t(1+r)^{T-t} + f_{t,T} > 0$. We just found an arbitrage. In the case $f_{t,T} > (1+r)^T S_t$:

1. At time t, short-sell the stock, put that money in the bank, enter a forward constract.

2. Wait until time T.

3. Pay $f_{t,T}$, recieve the stock and pass it to your broker who short-sold you the stock.

The final value of such a strategy is $V_T = -f_{t,T} + (1+r)^{T-t} S_t > 0$. Another arbitrage. $\square$

Remark 1.1. *In the second case, we supposed we are on a trading platform that allows short-selling, which basically amounts to borrowing a stock and selling it with the promise of giving it back later to the lender. This service is provided by brokers, one you open an account. A broker has the incentive to provide such services since he charges fees for the account's maintenance. However, larger players that have direct access to the market would rather not use brokers as intermediates. A different mechanism called repurchasing or "repo" is used. More about that in the exercise class.*

Remark 1.2. *We have proved that necessarily,* $f_{t,T} = (1+r)^T S_t = \frac{S_t}{B^0_{t,T}}$. *However, we cannot assert that this price is arbitrage-free. Without the appropriate tools, it is difficult to prove that arbitrages do not exist.*

The second application is a relation between the prices of European calls and puts with same strike and maturity. This is expected as the two contracts are somehow dual to each other:

Lemma 1.2 (Call-Put parity)**.** *Let $C(T, K)$ and $P(T, K)$ be the prices of a call and put with maturity T and strike K, at time $t = 0$. Then, by AOA:*

$$C(T, K) - P(T, K) = S_0 - K(1+r)^{-T}$$

Proof. Exercise. $\square$

1.6 The different players in financial markets

There are very different players on the market. Not only they have different incentives, but they also operate differently. One can list the following categories, although they are not mutually exclusive:

- Market makers serve as intermediates and their role is to quote prices of assets publicly, and continuously in time. At any moment, they should offer the service of buying or selling. Their presence is key in order to achieve liquidity, and their activity is generally restricted to the primary market.

- Arbitragists or speculators aim at identifying opportunities of arbitrage, and taking advantage of them.

- Hedgers generally deal with more complicated expositions (seconday market) and aim at neutralising the sensitivities of portfolios to risk. Supposedly, their incentive is not monetary gain.

2 Binomial or C-R-R model

C-R-R stands for Cox-Ross-Rubinstein, who were the first to introduce it. This is the simplest model for a financial market, yet with enough features to be representative of more general classes.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be our working probability space. We consider only finitely many times $t = 0, 1, 2, \ldots, T < \infty$. In this context, a discrete stochastic process $(X_t)_{0 \leq t \leq T}$ is a sequence of random variables indexed by time. Expectation under $\mathbb{P}$ is denoted $\mathbb{E}$ or $\mathbb{E}^{\mathbb{P}}$ if the reference measure is ambiguous.

2.1 Model specification

The binomial model concerns a primary market where only two assets are quoted.

- The bond with risk free interest rate r, whose deterministic dynamic is given by:

$$B_t = (1 + r)^t$$

- The stock, whose spot value S_t is written as a product of returns:

$$S_t = S_0 \prod_{i=1}^{t} \xi_i$$

We assume that, at every step, S_t jumps independently from its past to two possible values uS_t or dS_t. Here, u stands "up" while d stands for "down". Equivalently, the ξ_t are Bernouilli random variables with:

$$\forall t,\ \mathbb{P}(\xi_t = u) = 1 - \mathbb{P}(\xi_t = d) = p$$

The probability p gives the entire dynamic of the model and determines $\mathbb{P}$. $\mathbb{P}$ is refered to as the *historical probability* or real-world probability.

For concreteness, one can reduce Ω to the finite set $\{u, d\}^T$. Then $\mathcal{F} = \mathcal{P}(\Omega)$ is all the subsets of Ω and $\mathbb{P}$ is the product measure $\mathbb{P} = (p\delta_u + (1-p)\delta_d)^{\otimes T}$. If $\omega \in \Omega = \{u, d\}^T$, then:

$$\forall t, \xi_t(\omega) = \omega_t$$

$$\forall t, \forall a \in \{u, d\},\ \mathbb{P}(\xi_t = a) = \mathbb{P}\left(\omega \in \{u, d\}^{t-1} \times \{a\} \times \{u, d\}^{T-t}\right)$$

2.2 Filtrations, measurability and strategies

The notion of filtration: In order to decide for appropriate investment strategies, at time t, we need to take into account all the available information. Informally, we need a way of seeing the past increments $(\xi_1, \ldots, \xi_t)$ as deterministic and future ones $(\xi_{t+1}, \ldots, \xi_T)$ as random. This is what filtrations naturally achieve.

Definition 2.1. *A filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ is an increasing sequence of σ-algebras.*

Here we take $\mathcal{F}_t$ to be the smallest σ-algebra making $(\xi_1, \ldots, \xi_t)$ measurable, which is denoted:

$$\begin{aligned} \mathcal{F}_t &= \sigma(\xi_1, \ldots, \xi_t) \\ &= \sigma(S_1, \ldots, S_t) \end{aligned}$$

An important theorem from measure theory tells us that the functions that are measurable with respect to $\mathcal{F}_t$ are exactly functions of $(\xi_1, \ldots, \xi_t)$, and therefore, the spot dynamics up to time t. The general formulation is:

Theorem 2.1. *Let $f : (E, \mathcal{E}) \to (H, \mathcal{H})$ be a measurable map and define $\mathcal{E}_f = \sigma(f) = f^{-1}(\mathcal{H})$ to be the smallest σ-algebra so that f is $\mathcal{E}_f$-measurable. Suppose $g : (E, \mathcal{E}) \to ([0,1], Bor([0;1]))$ is a $\mathcal{E}_f$-measurable map. Then there exists a measurable $h : (H, \mathcal{H}) \to ([0,1], Bor([0;1]))$ such that $g = h \circ f$.*

Proof. Because g is numerical, we can approximate g by a g_n defined as $g_n(x) = \sum_{0 \leq k < 2^n} \frac{k}{2^n} \mathbb{1}_{\left\{g(x) \in [\frac{k}{2^n}; \frac{k+1}{2^n})\right\}}$. Since g is $\mathcal{E}_f$-measurable, every set $g^{-1}\left([\frac{k}{2^n}; \frac{k+1}{2^n})\right)$ is of the form $f^{-1}(A_{n,k})$ for a certain $A_{n,k} \in \mathcal{H}$. Hence $g_n = h_n \circ f$ where $h_n = \sum \frac{k}{2^n} \mathbb{1}_{A_{n,k}}$.
$h = \liminf h_n$ satisfies $g = h \circ f$. □

Of course, the theorem extends from $[0,1]$ to any measurably isomorphic set. Let us equate an investment strategy at time t with a vector (α_t, β_t). This vector specifies an amount α_t of stock and an amount β_t of bond to buy. Therefore, finding an optimal investment stragegy that takes into account all the information at time t is equivalently formulated as finding a vector that is $\mathcal{F}_t$-measurable.

Measurability of processes: In a sense, most of the interesting processes unfold in time as we discover their values as time flows. Examples are numerous: S_t, the weather, the temperature or even the number of students showing up at each lecture.

Definition 2.2. *A process $(X_t)_{0 \leq t \leq T}$ is called*

- *$\mathbb{F}$-adapted or $\mathcal{F}_t$-adapted when for all t, X_t is $\mathcal{F}_t$ measurable.*
- *Predictable when for all t, X_t is $\mathcal{F}_{t-1}$ measurable.*

Example 2.1. *The stock S_t is $\mathbb{F}$-adapted as $\mathbb{F}$ is the filtration it generates. B_t is adapted to any filtration as it is deterministic.*

Trading strategies:

Definition 2.3. *A trading strategy - in the primary market - is a* predictable *process $\varphi = (\varphi_t)_{0 \leq t \leq T}$ with $\varphi_t = (\alpha_t, \beta_t)$.*

As before, α_t is the number of shares of stock to hold during $(t-1; t]$ and β_t is the number of bonds to hold during that same period. Notice that $\alpha_t < 0$ is allowed and refers to short-selling (or repo); while $\beta_t < 0$ means borrowing money from the bank. The predictability hypothesis is crucial, as one rebalances his portfolios with φ_t during $(t-1; t]$. During this time period the only information available is $\mathcal{F}_{t-1}$. At exactly time t, the investor observes the new prices and does not rebalance his portfolio, since there is no time left.

Therefore, the value of a portfolio following the strategy φ is given by the process $V_t(\varphi)$ defined by:

$$V_0(\varphi) = \alpha_1 S_0 + \beta_1 B_0$$

$$\forall t \geq 1, V_t(\varphi) = \alpha_t S_t + \beta_t B_t$$

The class of strategies we allow are self-financing in the sense that we are only allowed to rebalance our portfolio, keeping its value constant:

Definition 2.4 (Self-financing condition)**.** *A self-financing strategy is a strategy φ such that:*

$$\forall t, \ \alpha_t S_t + \beta_t B_t = \alpha_{t+1} S_t + \beta_{t+1} B_t$$

This enables a formal definition of arbitrage opportunity:

Definition 2.5 (Arbitrage opportunity). *An arbitrage opportunity (in the primary market) is a self-financing trading strategy* φ *such that:*

- *No initial cost:*

$$V_0(\varphi) = 0$$

- *Always non-negative final value:*

$$V_T(\varphi) \geq 0 \ \mathbb{P} - a.s$$

- *With the possibility of a positive final gain:*

$$\mathbb{E}^{\mathbb{P}}(V_T(\varphi)) > 0 \Longleftrightarrow \mathbb{P}(V_T(\varphi) > 0) > 0$$

In fact, for the binomial model is arbitrage-free under the condition $d < 1 + r < u$. This will be apparent later using the tool of risk neutral measures. A better way to introduce such objects is via the theory of pricing.

2.3 Pricing of European options

Consider a European option with final payoff Φ_T. The matter of pricing is the question "What is its fair price P?". A true paradigm shift lies in the following answer: P would be the initial value $P = V_0(\varphi)$ of a self-financing portfolio such that

$$\Phi_T = V_T(\varphi)$$

when such a stragegy φ exists. In such a case, if a market maker charges P to his client for the European option, and executes the strategy φ, he can honor his side of the contract without further cost. Moreover, his final gain is zero. φ is then called a replicating or a hedging strategy. If such a φ exists, the option is called replicable. We will be only concerned with this setting, for now.

For all of this subsection, such an answer is valid as:

Theorem 2.2. *In the binomial model, under the natural conditions:*

$$d < 1 + r < u$$

every European option is replicable.

The focus is on how to compute the strategy φ, and whether the resulting price P gives an arbitrage-free market. As all the ideas are essentially in the one period case, we choose to present the matter as a discussion; where we progressively discover how pricing is made.

2.3.1 One period

Initial price: The payoff $\Phi_T = X$ is an $\mathcal{F}_T$-measurable random variable. Therefore, it is a deterministic function of S_1. Since S_1 can take only two values uS_0 and dS_0, the random variable X also takes two values. We denote by X^u the value of X on the event $\{\xi_1 = u\}$, and X^d the value on $\{\xi_1 = d\}$.

Let us determine a one-period replicating strategy $\varphi_1 = (\alpha_1, \beta_1)$. Because of the predictability assumption, φ is deterministic. As:

$$X = \Phi_T = \alpha S_1 + \beta B_1$$

needs to hold in all states of the universe, we obtain the linear system:

$$\begin{cases} X^u & = \alpha_1 S_0 u + \beta_1 (1+r) \\ X^d & = \alpha_1 S_0 d + \beta_1 (1+r) \end{cases}$$

Solving the linear system, one obtains:

$$\begin{cases} \alpha_1 & = \frac{X^u - X^d}{S_0(u-d)} \\ \beta_1 & = \frac{uX^d - dX^u}{(1+r)(u-d)} \end{cases}$$

Hence the initial price:

$$\begin{aligned} P &= \alpha_1 S_0 + \beta_1 B_0 \\ &= \frac{X^u - X^d}{u-d} + \frac{uX^d - dX^u}{(1+r)(u-d)} \\ &= \frac{1}{(1+r)(u-d)} \Big((1+r-d) X^u + (u-(1+r)) X^d \Big) \\ &= \frac{1}{1+r} \Big(p^* X^u + (1-p^*) X^d \Big) \end{aligned}$$

where $p^* = \frac{1+r-d}{u-d}$. Notice that it is natural to assume $d < 1+r < u$. Otherwise, we would have a stock that always outperforms the bond ($d > 1+r$) or an interest rate so high that the bond always outperforms the stock ($1+r > u$). Moreover, $d < 1+r < u$ is equivalent to $0 < p^* < 1$: p^* is a non-degenerate probability.

At this point, it is important to step back and discuss the previous computation. In the process of finding a replicating strategy and computing its initial value, one finds that the price P is obtained by weighting the discounted payoff $\frac{X}{1+r}$ by a probability p^*. We define a new probability $\mathbb{Q}$ by $\mathbb{Q}(\xi_1 = u) = p^*$, hence obtaining:

$$P = \mathbb{E}^{\mathbb{Q}} \left(\frac{X}{1+r} \right)$$

Remark 2.1.

- *For all practical purposes, the probability $\mathbb{Q}$ is only a computation device in order to find the price P. It does not yield any predictions regarding real-world spot dynamics.*
- *It is called the risk neutral measure because of the following property:*

$$\mathbb{E}^{\mathbb{Q}} \left(\frac{S_1}{1+r} \right) = S_0$$

Under $\mathbb{Q}$, stock and bond have the same expected return, and therefore an investor would have no preference between them.

Absence of arbitrage in the primary market: Aside from pricing, the risk neutral measure $\mathbb{Q}$ has a theoretical use. If non-denegenerate, it allows to prove very easily that there are no arbitrages in the primary market. Here we give a simple version of what will become the first fundamental theorem of asset pricing in section 4.

Consider a trading strategy φ. As we are in one period, this corresponds to a two-dimensional vector (α_1, β_1) where α_1 and β_1 are respectively the number of the shares of stock and the amount of bonds to hold during the period $(0;1]$. This quantity is deterministic, as it is decided at time 0. If φ is candidate for being an arbitrage opportunity, then:

$$V_0(\varphi) = \alpha_1 S_0 + \beta_1 = 0$$

$$V_1(\varphi) = \alpha_1 S_1 + \beta_1 (1+r) \geq 0 \text{a.s}$$

Taking the expectation under $\mathbb{Q}$, we have:

$$\mathbb{E}^{\mathbb{Q}}(V_1(\varphi)) = \alpha_1 \mathbb{E}^{\mathbb{Q}}(S_1) + \beta_1(1+r) = V_0(\varphi) = 0$$

Therefore, $V_1(\varphi)$ is a non-negative random variable with zero expectation. Hence, it is nothing but zero. What could have been an arbitrage opportunity is just a portfolio with zero value!

This proof is indeed remarkably simple for a result as strong as "there are no arbitrage opportunities". Virtually the same argument holds in the proof of theorem 2.6.

Absence of arbitrage in the extended market: Two natural questions are the following. Suppose I quote the price of our European option $P = \mathbb{E}^{\mathbb{Q}}\left(\frac{X}{1+r}\right)$ making a secondary market Bond-Stock-Option.

- In this extended market, can someone provide a better price? Perhaps by buying the option at favorable moments and selling it at unfavorable moments? This matter will be referred to as the question of *uniqueness*.
- We saw that the binomial model is arbitrage-free in the previous paragraph. But that statement concerns the primary market only: No arbitrages provided one trades stocks and bonds only. Now that I have introduced an option, did I break the arbitrage free property?

Theorem 2.3 (theorem 2.2.1 in [Wil06] for $T=1$ only)**.** *$P = \mathbb{E}^{\mathbb{Q}}\left(\frac{X}{1+r}\right)$ is the unique arbitrage-free initial price for the European option with payoff $\Phi_T = X$.*

We defer the proof to the general case. We will see that the market remaining arbitrage free is basically caused by the option being replicable.

2.3.2 Multi-period

In order to propagate the one period analysis to multiple periods, we will use the tools of conditional expectation and martingales given in appendix.

For any random variable X corresponding to a cash flow at time T, the actualised value at time t corresponds to the value of that cash flow in constant dollars. It is denoted $\widetilde{X}$ and $\widetilde{X} = \frac{X}{(1+r)^{T-t}}$.

Theorem 2.4 (Theorem 2.2.1 in [Wil06])**.** *Let Φ_T be an $\mathcal{F}_T$-measurable random variable corresponding to the payoff of a European option with maturity T. Then there exists a self-financing replicating strategy φ i.e $V_T(\varphi) = \Phi_T$.*

Moreover, the price of the option is given by $P = V_0(\varphi) = \mathbb{E}^{\mathbb{Q}}\left(\widetilde{\Phi_T}\right)$.

Proof. Recall that if $\varphi_t = (\alpha_t, \beta_t)$ is a strategy, then the value process is:

$$V_t(\varphi) = \alpha_t S_t + \beta_t B_t$$

We will proceed by backward induction to construct a stragegy φ such that:

$$\forall t \in \{0, 1, \ldots, T\},\ V_t(\varphi) = \frac{1}{(1+r)^{T-t}} \mathbb{E}^{\mathbb{Q}}(\Phi_T \mid \mathcal{F}_t)$$

Supposing the allocation of wealth in the portfolio has been decided for times larger than t, let us find an allocation (α_t, β_t) during $(t-1; t]$ such that:

$$V_{t-1}(\varphi) = \alpha_t S_{t-1} + \beta_t B_{t-1}$$

with the target:

$$V_t(\varphi) = \alpha_t S_t + \beta_t B_t = \frac{1}{(1+r)^{T-t}} \mathbb{E}^{\mathbb{Q}} (\Phi_T \mid \mathcal{F}_t) \tag{1}$$

We will now refer to the previous quantity as the target random variable. Notice this is just Φ_T if $t = T$. Clearly, the target random variable is $\mathcal{F}_t$ measurable and because of theorem 2.1, it is a function f of $(S_0, S_1, \ldots, S_t)$: $\frac{1}{(1+r)^{T-t}} \mathbb{E}^{\mathbb{Q}} (\Phi_T \mid \mathcal{F}_t) = f(S_0, \ldots, S_{t-1}, S_t)$ Therefore, conditionally on $\mathcal{F}_{t-1}$, one observes only two possible values:

$$V_t^u = f(S_0, \ldots, S_{t-1}, S_{t-1}u)\,; \; V_t^d = f(S_0, \ldots, S_{t-1}, S_{t-1}d)\,;$$

Asking for equation (1) to hold in both states of the universe gives the linear system of equations:

$$\begin{aligned} V_t^u &= \alpha_t S_{t-1} u + \beta_t B_t \\ V_t^d &= \alpha_t S_{t-1} d + \beta_t B_t \end{aligned}$$

Solving the system, as in the one-period case, gives:

$$\alpha_t = \frac{V_t^u - V_t^d}{(u-d) S_{t-1}}$$

$$\beta_t = \frac{1}{(1+r)^t} \frac{u V_t^d - d V_t^u}{(u-d)}$$

Now, reassembling the pieces, one has:

$$\begin{aligned} V_{t-1}(\varphi) &= \alpha_t S_{t-1} + \beta_t B_{t-1} \\ &= \frac{V_t^u - V_t^d}{(u-d)} + \frac{1}{(1+r)} \frac{u V_t^d - d V_t^u}{(u-d)} \\ &= \frac{1}{1+r} \left(p^* V_t^u + (1-p^*) V_t^d \right) \\ &= \frac{1}{1+r} \mathbb{E}^{\mathbb{Q}} (V_t(\varphi) \mid \mathcal{F}_{t-1}) \\ &= \frac{1}{1+r} \mathbb{E}^{\mathbb{Q}} \left(\frac{1}{(1+r)^{T-t}} \mathbb{E}^{\mathbb{Q}} (\Phi_T \mid \mathcal{F}_t) \mid \mathcal{F}_{t-1} \right) \\ &= \frac{1}{(1+r)^{T-t+1}} \mathbb{E}^{\mathbb{Q}} (\Phi_T \mid \mathcal{F}_{t-1}) \end{aligned}$$

$\square$

Notice that in the course of the proof, we gave a stronger result. We derived the value of the option in all intermediate times $t \in \{0, 1, \ldots, T\}$.

Corollary 2.1 (Corollary of the proof)**.** *For all times, the value of an option with final payoff Φ_T is:*

$$\frac{1}{(1+r)^{T-t}} \mathbb{E}^{\mathbb{Q}} (\Phi_T \mid \mathcal{F}_t)$$

In theory, this gives the price at which one could sell back his option before maturity.

Application to the pricing of a European call option: If we denote by $\mathcal{B}$ the number of "up" in a trajectory of the spot, then $\mathcal{B}$ is a binomial random variable with parameters (T, p^*) under $\mathbb{Q}$. Moreover:

$$S_T = S_0 \prod_{j=1}^{T} \xi_j = S_0 u^{\mathcal{B}} d^{T-\mathcal{B}}$$

Hence:

$$\forall k \in \{0, 1, \dots, T\}, \mathbb{Q}\left(S_T = S_0 u^k d^{T-k}\right) = \mathbb{P}\left(\mathcal{B} = k\right) = \binom{T}{k} \left(p^*\right)^k \left(1 - p^*\right)^{T-k}$$

In the end, the pricing formula in theorem 2.4 yields the price $C_0 = C(T, S_0)$ where:

$$\begin{aligned} C\left(T, x\right) =& \frac{1}{\left(1+r\right)^T} \mathbb{E}\left(\left(x u^{\mathcal{B}} d^{T-\mathcal{B}} - K\right)^+\right) \\ =& \frac{1}{\left(1+r\right)^T} \sum_{k=0}^{T} \binom{T}{k} \left(p^*\right)^k \left(1 - p^*\right)^{T-k} \left(x u^k d^{T-k} - K\right)^+ \end{aligned}$$

A general philosophy: The goal of a reasonable pricing theory is to construct a risk neutral measure $\mathbb{Q}$ under which the price of any European option with payoff Φ_T is given, at any time, by the conditional expectation under $\mathbb{Q}$ conditionally to $\mathcal{F}_t$. As a consequence, the discounted value of the replication portfolio satisfies:

$$\widetilde{V_t\left(\varphi\right)} = \mathbb{E}^{\mathbb{Q}}\left(\frac{\Phi_T}{\left(1+r\right)^T} | \mathcal{F}_t\right)$$

which is a closed $\mathbb{Q}$-martingale.

The following lemma goes in the same direction:

Lemma 2.1. *The processes $\widetilde{S}_t := \frac{1}{(1+r)^t} S_t$ and $\widetilde{V_t\left(\varphi\right)}$, for any bounded self-financing strategy φ, are $\mathbb{Q}$-martingales.*

Proof. Let us fix a time value t. For the first statement, we have $\mathbb{E}^{\mathbb{Q}}\left(\widetilde{S}_{t+1} | \mathcal{F}_t\right) = \widetilde{S}_t \mathbb{E}^{\mathbb{Q}}\left(\frac{\xi_{t+1}}{1+r}\right)$ and $\mathbb{E}^{\mathbb{Q}}\left(\frac{\xi_{t+1}}{1+r}\right) = \frac{up^*}{(1+r)} + \frac{d(1-p^*)}{(1+r)} = 1$.

For the second statement, write $\widetilde{V_t\left(\varphi\right)} = \alpha_t \widetilde{S}_t + \beta_t$ which is adapted and integrable. Then $\mathbb{E}^{\mathbb{Q}}\left(\widetilde{V_{t+1}\left(\varphi\right)} | \mathcal{F}_t\right) = \alpha_{t+1} \mathbb{E}^{\mathbb{Q}}\left(\widetilde{S}_{t+1} | \mathcal{F}_t\right) + \beta_{t+1} = \alpha_{t+1} S_t + \beta_{t+1} = \alpha_t S_t + \beta_t = \widetilde{V_t\left(\varphi\right)}$. The next to last step used the self financing condition. □

The risk neutral measure allows to prove that the market is arbitrage-free.

Theorem 2.5. *If $\mathbb{Q}$ is a non-denegenerate probability measure i.e $0 < p^* < 1$, then the primary market is arbitrage-free.*

This remains true even if one is allowed to trade in the option like any other asset. The previous theorem is also implied by the following.

Theorem 2.6. *Under the same assumption, the price $P = \mathbb{E}^{\mathbb{Q}}\left(\widetilde{\Phi_T}\right)$ is the unique arbitrage-free initial price for a European option with payoff Φ_T.*

Remark 2.2. *As before, the uniqueness is regarding the possibility that in an extended (secondary) market, the price might changes if other players can replicate the option for a lesser fee. The arbitrage-free statement regards the absence of opportunity of arbitrage in the entire secondary market.*

Proof. We will not give the uniqueness argument and refer to [Wil06]. Basically, one has to come up with arbitrage strategies in the extended secondary market if the quoted price is different from P. Now, for the absence of opportunities of arbitrage in the extended market, everything boils down to a martingale argument. It is virtually the same as in one-period in the primary market. Consider self-financing strategies in the extended market to be predictable processes φ^{ext} with $\varphi_t^{ext} = (\alpha_t, \beta_t, \gamma_t)$. During the period $(t-1; t]$,

- α_t is the number of share of stock to hold.
- β_t is the amount of cash to hold.
- γ_t is the nominal of options to have.

The value of our portfolio at any time t is:

$$V_t\left(\varphi^{ext}\right) = \alpha_t S_t + \beta_t B_t + \gamma_t C_t$$

where C_t is the price of the option at time t. Because the option is replicable, there exists a trading strategy in the primary market $\varphi_t^* = (\alpha_t^*, \beta_t^*)$ such that:

$$V_t\left(\varphi^*\right) = C_t = \alpha_t^* S_t + \beta_t^* B_t$$

Hence:

$$V_t\left(\varphi^{ext}\right) = \left(\alpha_t + \gamma_t \alpha_t^*\right) S_t + \left(\beta_t + \gamma_t \beta_t^*\right) B_t$$

From the fact that φ^{ext} is self-financing, it is easy to deduce that the strategy $(\alpha_t + \gamma_t\alpha_t^*, \beta_t + \gamma_t\beta_t^*)$ is self-financing. Therefore, thanks to lemma 2.1, the process $\left(V_t\left(\varphi^{ext}\right)\right)_{t\leq T}$ is a $\mathbb{Q}$-martingale, hence constant in expectation:

$$\mathbb{E}^{\mathbb{Q}}\left(V_t\left(\varphi^{ext}\right)\right) = V_0\left(\varphi^{ext}\right) = 0$$

As a consequence $V_t\left(\varphi^{ext}\right) = 0$, $\mathbb{Q}$-almost surely. Moreover, the condition $0 < p^* < 1$ shows $\mathbb{Q}$ is non-degenerate: only the empty event has probability zero. Therefore $V_t\left(\varphi^{ext}\right) = 0$, without any statement on probability: No arbitrage! □

3 American options

In this section, I will follow both [Wil06] and [LL08]. Recall that an American option is an option that can be exercised at any time $\tau \in \{0, 1, 2, \ldots, T\}$. The possible payoffs at each date are specified by $(\varphi_t)_{0\leq t\leq T}$.

For modeling purposes, τ is chosen to be a stopping time. This choice is motivated by both financial and mathematical reasons:

- The financial rationale is that the client's choice to exercise is motivated only by the past spot dynamic and not its future. Therefore, $\{\tau \leq t\}$ should be in $\mathcal{F}_t$. If τ is not a stopping time in the natural filtration $\mathbb{F}$, one could argue the investor has access to information regarding the future of the spot, making him guilty of insider trading.

- We have examined in exercise class a case where we have expended the spot filtration $\mathbb{F}$ to a filtration $\mathbb{G}$ by adding the final value of the stock. This procedure does change the martingales and allows for arbitrage opportunities. In fact, the same is true if one expands $\mathbb{F}$ thanks to a random time τ by taking $\mathbb{G}$ to be smallest filtration making τ into a stopping time.

We are aware that this explanation is perhaps involved, but it can be ignored in a first reading.

3.1 A primer in the context of the binomial model

There is an important difference between pricing methodologies of European and American options. In the case of European options, we have seen there is always a replicating strategy so that a predetermined initial cost allows to cover the payoff.

In the case of an American option with payoff $(\varphi_t)_{0\leq t\leq T}$, one cannot in general find a self-financing stragegy φ such that:

$$\forall t, V_t(\varphi) = \Phi_t$$

Then, the next best thing to achieve would to look for a *super-replicating* stragegy i.e a strategy φ such that:

$$\forall t, V_t(\varphi) \geq \Phi_t$$

and postulate that the price P of the option is minimum amount of initial wealth among super-replicating strategies:

$$P := \inf_{\varphi \text{ such that } \forall t, V_t(\varphi)\geq \Phi_t} V_0(\varphi)$$

Notice that we are not garanteed that the minimum is reached, hence the presence of an infinimum instead of a minimum. In the context of the binomial model, the minimum is attained and the value process is described in the following proposition:

Proposition 3.1. *Let U_t be the minimum amount of wealth to cover the payoff between t and T:*

$$U_t := \inf_{\varphi \text{ such that } \forall s\geq t, V_s(\varphi)\geq \Phi_s} V_t(\varphi)$$

Then, there is a self-financing strategy φ^* such that $U_0 = V_0(\varphi^*)$. Moreover, the process U_t satisfies the following backward recurrence:

$$U_T = \Phi_T$$
$$U_t = \max\left(\Phi_t, \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}\left(U_{t+1} \mid \mathcal{F}_t\right)\right)$$

Proof 1: "Quick and dirty". Let us give a quick proof of the backward recurrence. Because U_T is the minimum amount of wealth to cover Φ_T, we trivially have $U_T = \Phi_T$. Now for $t < T$, U_t has to cover Φ_t in the case of immediate exercise and cover an exercise past $t+1$. If the wealth needed at time t to cover an exercise past $t+1$ is denoted C_t, we already see:

$$U_t = \max(\Phi_t, C_t)$$

The European pricing theory tells us that C_t is obtained by actualising the amount U_{t+1} and taking conditional expectation under $\mathbb{Q}$:

$$C_t = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}\left(U_{t+1} \mid \mathcal{F}_t\right)$$

Hence the result. □

Proof 2: "Where are my stragegies?" Seen in class. □

Candidate for being an optimal stopping time Define:

$$\tau := \inf \{t \geq 0 \mid U_t = \Phi_t\}$$

As both $(U_t - \Phi_t; t \geq 0)$ is an adapted process, τ is a stopping time as any hitting time. Moreover, since that for $t < \tau$, $U_t > \Phi_t$ or equivalently $U_t = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}\left(U_{t+1} \mid \mathcal{F}_t\right)$. This means that $\widetilde{\delta}_{t+1} = 0$: No excess of wealth!

This also tells us that for $0 \leq t \leq \tau$, $V_t(\varphi^*) = U_t$: the super-replication of the minimum amount needed to cover the option is in fact a replication.

Reformulation: After discounting, we see that the discounted process $\widetilde{U}_t$ is defined by the backward recurrence:

$$\widetilde{U}_T = \widetilde{\Phi}_T$$

$$\widetilde{U}_t = \max\left(\widetilde{\Phi}_t, \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}\left(\widetilde{U}_{t+1} \mid \mathcal{F}_t\right)\right)$$

One can study such objects in the context of optimal stopping theory.

3.2 Optimal stopping problems and Snell envelopes

Let us now consider a general setting, falling possibly outside the binomial model. As usual let $(\Omega, \mathcal{F}, \mathbb{P})$ denote a probability space, along with a filtration $\mathbb{F} = \{\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_T\}$. In the following, it is good to have in mind the case of American options as an application.

Definition 3.1 (Definition and theorem)**.** *Let $(\Phi_t; t \geq 0)$ be an $\mathbb{F}$-adapted process that is integrable for all t. If we define:*

$$U_T = \Phi_T$$

$$U_t = \max\left(\Phi_t, \mathbb{E}\left(U_{t+1} \mid \mathcal{F}_t\right)\right)$$

then $(U_t; t \geq 0)$ is the smallest supermartingale dominating Φ_t. It is called the Snell envelope of $(\Phi_t; t \geq 0)$.

Remark 3.1. *Because we adopt a general setting here, we dropped the tilde used for discounting and expectation is under a generic $\mathbb{P}$, instead of $\mathbb{Q}$.*

Proof. In order to prove that U is a supermartingale, it is enough to write:

$$U_t = \max\left(\Phi_t, \mathbb{E}\left(U_{t+1} \mid \mathcal{F}_t\right)\right) \geq \mathbb{E}\left(U_{t+1} \mid \mathcal{F}_t\right)$$

In order to prove that any other supermartingale dominating Φ has to larger than U, let W be a supermartingale dominating Φ. Then necessarily, for all times t, we have:

$$W_t \geq \Phi_t$$

$$W_t \geq \mathbb{E}\left(W_{t+1} | \mathcal{F}_t\right)$$

necessarily implying $W_t \geq \max\left(\Phi_t, \mathbb{E}\left(W_{t+1} | \mathcal{F}_t\right)\right)$. Then by backward induction, one proves that $W_t \geq U_t$. □

The random time we had defined satisfies the following:

Proposition 3.2 (Proposition 2.2.1 in [LL08])**.** *The random variable $\tau := \inf\{t \geq 0 | U_t = \Phi_t\}$ is a stopping time and the stopped process U^τ is a martingale.*

Proof. Both U and Φ are adapted. Therefore τ is the first hitting time of zero for $X = U - \Phi$. Any first hitting time is a stopping time thanks to the following:

$$\{\tau = t\} = \{X_0 \neq 0\} \cap \cdots \cap \{X_{t-1} \neq 0\} \cap \{X_t = 0\}$$

As X is adapted, every event in this intersection is in $\mathcal{F}_t$.

Now, in order to prove that U^τ is a martingale, we use the same trick as in the proof of Doob's optional stopping theorem. We start by writing:

$$U_t^\tau = U_0 + \sum_{s=1}^{t} U_s^\tau - U_{s-1}^\tau = U_0 + \sum_{s=1}^{t} \mathbb{1}_{\{s \leq \tau\}} \left(U_s - U_{s-1}\right)$$

But if $s \leq \tau$, $U_{s-1} > \Phi_{s-1}$ and hence $U_{s-1} = \mathbb{E}\left(U_s | \mathcal{F}_{s-1}\right)$. This entails:

$$U_t^\tau = U_0 + \sum_{s=1}^{t} \mathbb{1}_{\{s \leq \tau\}} \left(U_s - \mathbb{E}\left(U_s | \mathcal{F}_{s-1}\right)\right)$$

From that, one concludes easily that U^τ is a martingale.

$$\mathbb{E}\left(U_{t+1}^\tau - U_t^\tau | \mathcal{F}_t\right) = \mathbb{E}\left(\mathbb{1}_{\{t+1 \leq \tau\}} \left(U_{t+1} - \mathbb{E}\left(U_{t+1} | \mathcal{F}_t\right)\right) | \mathcal{F}_t\right) = 0$$

□

If $\mathcal{T}_{t,T}$ is the set of all stopping times valued in $\{t, t+1, \ldots, T\}$. Then define for every t:

$$\tau_t = \inf\left\{k \geq t \mid U_k = \Phi_k\right\} \in \mathcal{T}_{t,T}$$

Corollary 3.1.

$$U_t = \mathbb{E}\left(\Phi_{\tau_t} \mid \mathcal{F}_t\right) = \max_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\left(\Phi_\tau \mid \mathcal{F}_t\right)$$

Remark 3.2. *By definition τ_t is the earliest optimal stopping time. The financial interpretation of this statement is as follows. The Snell envelope is the expectation of the payoff, at its best exercise moment. Then it is only logical that under $\mathbb{Q}$, this expectation becomes the price of the American option.*

Proof. Let us deal with the first equality. By adapting the proof of the previous proposition to a different starting moment t instead of 0, one sees that $(U_s^{\tau_t}; s = t, t+1, \ldots, T)$ is a martingale. Then:

$$U_t = U_t^{\tau_t} = \mathbb{E}\left(U_T^{\tau_t} | \mathcal{F}_t\right) = \mathbb{E}\left(U_{\tau_t} | \mathcal{F}_t\right) = \mathbb{E}\left(\Phi_{\tau_t} | \mathcal{F}_t\right)$$

For the second equality, consider any stopping time $\tau \in \mathcal{T}_{t,T}$. By Doob's optional stopping theorem, $(U_s^\tau; s \geq t)$ is a supermartingale. Hence $U_t \geq \mathbb{E}\left(U_\tau | \mathcal{F}_t\right) = \mathbb{E}\left(\Phi_\tau | \mathcal{F}_t\right)$. Therefore:

$$U_t \geq \max_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\left(\Phi_\tau \mid \mathcal{F}_t\right)$$

and equality is reached for the stopping time τ_t. □

4 Finite market models

4.1 Definitions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. As usual, we assume time to be discrete $1, 2, \dots, T$ with $T < \infty$. The working filtration is again denoted:

$$\mathbb{F} := \{\mathcal{F}_0 \subset \mathcal{F}_1 \dots \mathcal{F}_T\}$$

A key assumption that allows a simple treatment of financial market is finiteness of the state-space:

Definition 4.1. *A finite market model is a model where $|\Omega| < \infty$. We label the possible scenarios by:*

$$\Omega := \{\omega_1, \dots, \omega_n\}$$

Remark 4.1. *In the context of the binomial model, $\Omega = \{u, d\}^T$ and $|\Omega| = 2^T$.*

Observables: We assume that one observes a family of discrete stochastic processes $\left(S_t^i; 0 \le t \le T\right)$ with $0 \le i \le d$ which are adapted. They model the prices of $1 + d$ assets and among them, there is a distinguished risk-free asset

$$S_t^0 = (1 + r)^t$$

which is deterministic. We write in vector notation:

$$S_t = \left(S_t^0, \dots, S_t^d\right)$$

and we typically take for filtration $\mathcal{F}_t = \sigma\left(S_0, S_1, \dots, S_t\right)$.

Trading strategies are predictable $1 + d$-dimensional processes $\varphi_t = \left(\varphi_t^0, \dots, \varphi_t^d\right)$ such that φ_t^i indicates the quantity of stock i to hold during the period $(t-1; t]$. The self-financing condition is:

$$\sum_i \varphi_t^i S_t^i = \sum_i \varphi_{t+1}^i S_t^i$$

which can also be written in vector notation as an equality between scalar products in $\mathbb{R}^{d+1}$:

$$\varphi_t \cdot S_t = \varphi_{t+1} \cdot S_t$$

Value process: The value of a portfolio with self-financing strategy φ is:

$$V_t\left(\varphi\right) = \varphi_t \cdot S_t = \varphi_{t+1} \cdot S_t$$

Lemma 4.1.

$$V_t\left(\varphi\right) = V_0\left(\varphi\right) + \sum_{s=1}^{t} \varphi_s \cdot \Delta S_{s-1}$$

where $\Delta S_{s-1} = S_s - S_{s-1}$.

Remark 4.2. *The previous expression is a discrete version of a stochastic integral, a term we will loosely use. The framework of discrete time has the advantage of avoiding all the technicalities associated to the construction of this integral.*

Proof. Using successively a telescoping sum and the self-financing condition:

$$\begin{aligned}
V_t(\varphi) &= V_0(\varphi) + \sum_{s=1}^{t} \left[V_s(\varphi) - V_{s-1}(\varphi)\right] \\
&= V_0(\varphi) + \sum_{s=1}^{t} \left[\varphi_s \cdot S_s - \varphi_s \cdot S_{s-1}\right] \\
&= V_0(\varphi) + \sum_{s=1}^{t} \varphi_s \cdot \left[S_s - S_{s-1}\right]
\end{aligned}$$

$\square$

Discounted processes: For every i, the discounted asset i is given by:

$$\widetilde{S}_t^i = \frac{S_t^i}{S_t^0}$$

Clearly $\widetilde{S}_t = \left(1, \widetilde{S}_t^1, \ldots, \widetilde{S}_t^d\right)$ can be viewed as d-dimensional vector and:

$$\widetilde{V(\varphi)} = V_0(\varphi) + \sum_{s=1}^{t} \varphi_s \cdot \widetilde{\Delta S_{s-1}}$$

where the scalar product involves only risky assets ($\widetilde{\Delta S^0_{s-1}} = 0$).

Arbitrage opportunities are strategies φ such that

- (Zero initial cost):
$$V_0(\varphi)$$
- (Non-negative return):
$$V_T(\varphi) \geq 0$$
- (Possible positive return):
$$\mathbb{E}\left(V_T(\varphi)\right) > 0 \Leftrightarrow \mathbb{P}\left(V_T(\varphi) > 0\right) > 0$$

The previous equivalence is only possible because Ω is finite.

4.2 AOA and the first fundamental theorem of asset pricing

Definition 4.2. *Consider two probability measures $\mathbb{Q}$ and $\mathbb{Q}'$ on $(\Omega, \mathcal{F})$.*

- *$\mathbb{Q}$ is absolutely continuous w.r.t $\mathbb{Q}'$ ($\mathbb{Q} \ll \mathbb{Q}'$) if*
$$\forall A \in \mathcal{F}, \mathbb{Q}'(A) = 0 \Rightarrow \mathbb{Q}'(A) = 0$$
- *$\mathbb{Q}$ and $\mathbb{Q}'$ are equivalent ($\mathbb{Q} \approx \mathbb{Q}'$) if they are absolutely continuous with respect to each other, or equivalently*
$$\forall A \in \mathcal{F}, \mathbb{Q}'(A) = 0 \Leftrightarrow \mathbb{Q}'(A) = 0$$

Remark 4.3. *Equivalence between two measures tantamounts to having the same null sets.*

Definition 4.3 (EMM or risk neutral measures)**.** *An Equivalent Martingale Measure, or risk free measure is a probability $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ such that:*

- *$\mathbb{Q} \approx \mathbb{P}$ i.e has the same null sets as the historical probability.*
- *$\widetilde{S}_t$ is a $\mathbb{Q}$-martingale (for the natural working filtration).*

In order to lift any ambiguity, recall that in this setting the stock S_t is a vector and the identity:

$$\mathbb{E}^{\mathbb{Q}}\left(\widetilde{S_{t+1}}|\mathcal{F}_t\right) = \widetilde{S}_t$$

giving the martingale property has to be understood component-wise. Now for the main theorem in this subsection. It gives a particularly computable criterion for deciding whether a market model is arbitrage-free or not.

Theorem 4.1 (First fundamental theorem of asset pricing)**.** *The finite market model has no arbitrage if and only if there exists $\mathbb{Q}$, an Equivalent Martingale Measure.*

Before diving into the proof, let us give a useful lemma around martingales.

Lemma 4.2.

- *Discrete stochastic integrals against martingales are martingales: If $\widetilde{S}_t$ is a $\mathbb{Q}$-martingale then $\widetilde{V_t(\varphi)}$ is a $\mathbb{Q}$-martingale.*
- *Martingales are the adapted processes against which any stochastic integral is centered: If M is an adapted process, then*

$$M \textit{ martingale } \Leftrightarrow \left(\forall \eta \textit{ predictable}, \ \mathbb{E}\left(\sum_{t=1}^{T} \eta_t \left(M_t - M_{t-1}\right)\right) = 0\right)$$

Proof. For the first point, recall that:

$$\widetilde{V_t(\varphi)} = \widetilde{V_0(\varphi)} + \sum_{s=1}^{t} \varphi_s \cdot \Delta \widetilde{S}_s$$

Hence:

$$\begin{aligned} \mathbb{E}\left(\widetilde{V_{t+1}(\varphi)} - \widetilde{V_t(\varphi)}|\mathcal{F}_t\right) &= \mathbb{E}\left(\varphi_{t+1} \cdot \Delta\widetilde{S}_{t+1}|\mathcal{F}_t\right) \\ &\overset{\varphi \text{ predictable}}{=} \varphi_{t+1} \cdot \mathbb{E}\left(\widetilde{S}_{t+1} - \widetilde{S}_t|\mathcal{F}_t\right) \\ &= 0 \end{aligned}$$

For the second point, we need to prove an equivalence. The implication $\Leftarrow$ has in fact already been obtained. Let us now deal with $\Rightarrow$. Let $A \in \mathcal{F}_{s-1}$ and define

$$\eta_t := \begin{cases} 0 & \text{if } s \neq t \\ \mathbb{1}_A & \text{if } s = t \end{cases}$$

Applying the hypothesis for that η gives:

$$0 = \mathbb{E}\left(\mathbb{1}_A \left(M_s - M_{s-1}\right)\right),$$

a statement which holds then for all A and s, whence the martingale property. □

Proof of theorem 4.1. " $\Leftarrow$" Let us assume the existence of a martingale measure Q. We will use the usual martingale trick which we have already seen in the case of the binomial model. Let φ be a possible arbitrage opportunity, i.e:

$$V_0(\varphi) = 0, \quad V_T(\varphi) \geq 0$$

Thanks to the fact that $\widetilde{V_{\cdot}(\varphi)}$ is a $\mathbb{Q}$-martingale, we readily obtain that:

$$\mathbb{E}^{\mathbb{Q}}\left(\widetilde{V_T(\varphi)}\right) = 0$$

Since we are taking the expectation of a positive random variable, we deduce:

$$\mathbb{Q}\left(\widetilde{V_T(\varphi)} > 0\right) = 0$$

Now recall that $\mathbb{Q}$ and $\mathbb{P}$ have the same null sets. Hence:

$$\mathbb{P}\left(V_T(\varphi) > 0\right) = \mathbb{Q}\left(V_T(\varphi) > 0\right) = 0$$

and φ cannot therefore be an arbitrage opportunity.

" $\Rightarrow$" is harder implication to prove, based on convex analysis on the space of random variables on Ω. Since $|\Omega| < \infty$, we can list all the possible scenarios by writing $\Omega = \{\omega_1, \ldots, \omega_n)$. Then the space of $\mathbb{R}$-valued random variables is identified with $\mathbb{R}^n$: any random variable $Y : \Omega \to \mathbb{R}$ can be seen as the point $(Y(\omega_1), \ldots, Y(\omega_n))$ in $\mathbb{R}^n$.

Now consider the set of attainable terminal gains, following a self-financing strategy, with zero initial wealth:

$$L_0 := \left\{\widetilde{V_T(\varphi)} \mid \varphi \text{ is self-financing with } V_0(\varphi) = 0\right\},$$

which is a subset of $\mathbb{R}^n$, in the identification described previously. More than that, it is a non-empty linear subspace ($0 \in L_0$ and the dependence in φ is linear).

It is easy to see that the market has no arbitrage opportunity if and only if L_0 never intersects $(\mathbb{R}_+)^n \setminus \{0\}$. By rescaling that possible intersection, we have that the market is arbitrage-free if and only if

$$L_0 \cap F = \emptyset$$

where

$$F := \left\{Y \in (\mathbb{R}_+)^n \mid \sum_{i=1}^{n} Y_i = 1\right\}$$

The set F is compact and convex, while L_0 is linear. By the separating hyperplane theorem, these two convex bodies can be separated by a hyperplane H containing L_0. If $Z \neq 0$ is a vector normal to H:

$$H := \{Y \in \mathbb{R}^n \mid Y \cdot Z = 0\}$$

Without any loss of generality, one can suppose that F sits on the positive side of H i.e $\forall f \in F, f \cdot Z > 0$. In particular, since the canonical basis of $\mathbb{R}^n$ belongs to F, we have that $Z_i > 0$, $\forall i$. Therefore, the probability measure:

$$\mathbb{Q}(\{\omega_i\}) = \frac{Z_i}{\sum_j Z_j}$$

is a probability measure, equivalent to $\mathbb{P}$. For a trading strategy such that $V_0(\varphi) = 0$, we have:

$$\mathbb{E}^{\mathbb{Q}}\left(\widetilde{V_T(\varphi)}\right) = \sum_{\omega \in \Omega} \widetilde{V_T(\varphi)}(\omega)\mathbb{Q}(\omega)$$

$$
\begin{aligned}
&= \sum_i \frac{\widetilde{V_T(\varphi)}(\omega_i) Z_i}{\sum Z_j} \\
&= \frac{\widetilde{V_T(\varphi)} \cdot Z}{\sum Z_j} \\
&= 0
\end{aligned}
$$

By breaking down the previous equality:

$$0 = \mathbb{E}^{\mathbb{Q}} \left(\widetilde{V_T(\varphi)} \right) = \sum_{t=1}^{T} \mathbb{E}^{\mathbb{Q}} \left(\varphi_t \cdot \Delta \widetilde{S}_{t-1} \right)$$

Take $\varphi^j = 0$ for all $j \neq i$ in order to obtain:

$$\forall i, \forall \eta \text{ predictable }, 0 = \sum_{t=1}^{T} \mathbb{E}^{\mathbb{Q}} \left(\eta_t \Delta \widetilde{S}^i_{t-1} \right)$$

Lemma 4.2 tells us that for every i, $\widetilde{S}^i$ is $\mathbb{Q}$-martingale. $\mathbb{Q}$ is indeed an Equivalent Martingale measure. □

4.3 Pricing

Suppose we are dealing with a market model with no arbitrage. The existence of an Equivalent Martingale Measure $\mathbb{Q}$ proves to be very useful as a *pricing tool*. However there is a little restriction on the options to be priced.

Definition 4.4 (Replicable options and completeness)**.** *A replicable option (or attainable claim) is a random variable X such that there exists a self-financing strategy φ such that:*

$$V_T(\varphi) = X$$

A market is complete is all European options are replicable.

The following subsection discusses completeness and partly describes the set of replicable options, if the market is not complete. For now, we answer the question of pricing, provided that the option is replicable.

Theorem 4.2.
- *For any replicating strategy, the value process $(V_t(\varphi); 0 \leq t \leq T)$ is the same.*
- *Moreover, for any $\mathbb{Q}$, equivalent martingale measure:*

$$\widetilde{V_t(\varphi)} = \mathbb{E}^{\mathbb{Q}} \left(\widetilde{X} | \mathcal{F}_t \right)$$

Proof. For the first point, one can provide two arguments.

- Financial argument: If φ and φ' are two replicating strategies, the associated value processes cannot deviate from each other otherwise one can form an arbitrage by buying the cheapest and selling the more expensive one. $V_t(\varphi)$ gives X at time T no matter the replicating strategy φ.
- Mathematical argument: There is at least one Equivalent Martingale Measure $\mathbb{Q}$ and $\widetilde{V_t(\varphi)}$ is a $\mathbb{Q}$-martingale. Therefore

  $$\widetilde{V_t(\varphi)} = \mathbb{E}^{\mathbb{Q}} \left(\widetilde{V_T(\varphi)} | \mathcal{F}_t \right) = \mathbb{E}^{\mathbb{Q}} \left(\widetilde{X} | \mathcal{F}_t \right),$$

  which does not depend on φ.

The second point is given in the mathematical argument. □

By interpreting the initial value of the portofolio as the value of an option, we have:

Corollary 4.1. *If the market has no arbitrage and is complete, then the price of any option at time t, with European payoff Φ_T is given by:*

$$S_t^0 \mathbb{E}^{\mathbb{Q}} \left(\widetilde{\Phi_T} | \mathcal{F}_t \right) = S_t^0 \mathbb{E}^{\mathbb{Q}} \left(\frac{\Phi_T}{S_T^0} | \mathcal{F}_t \right)$$

4.4 Completeness and the second fundamental theorem of asset pricing

Theorem 4.3. *Consider a finite market model with no arbitrage. The market is complete if and only if the risk neutral measure $\mathbb{Q}$ is unique.*

The theorem is intuitive if one recalls how the first fundamental theorem is proved. We had seen that the risk neutral measure $\mathbb{Q}$ is constructed from the vector orthonormal to the space L_0 of attainable payoffs, with zero initial value. Many choices arise from the fact the orthonormal vector is not unique, unless L_0 is as big as possible ie if it is a hyperplane itself. It is not very hard to see that L_0 being a hyperplane is equivalent to the market being complete.

Proof. " $\Rightarrow$" Consider two risk neutral measures $\mathbb{Q}$ and $\mathbb{Q}'$, which we will now prove to be equal. For any $A \in \mathcal{F}_T$, consider the option with payoff $\Phi_T = \mathbb{1}_A$. By hypothesis, Φ_T is replicable. The associated value process does not depend on the replicating strategy nor on the measure (theorem 4.2), hence:

$$\forall 0 \leq t \leq T, \ \mathbb{E}^{\mathbb{Q}} \left(\widetilde{X} | \mathcal{F}_t \right) = E^{\mathbb{Q}'} \left(\widetilde{X} | \mathcal{F}_t \right)$$

By discarding the discounting (S_T^0 is deterministic), and retricting time to $t = 0$ where the filtration is trivial, we obtain:

$$\mathbb{E}^{\mathbb{Q}} \left(\mathbb{1}_A \right) = \mathbb{E}^{\mathbb{Q}'} \left(\mathbb{1}_A \right),$$

whence uniqueness!

" $\Leftarrow$" We will proceed by contraposition: We will assume that there is a random variable X that is not attainable, making the market incomplete and we will exhiting "a segment" of risk neutral measures. Let L be the set of attainable random variables:

$$\begin{aligned} L &= \left\{ c + \sum_{t=1}^{T} \varphi_t \cdot \Delta \tilde{S}_{t-1} \mid \varphi \text{ predicable } , c \in \mathbb{R} \right\} \\ &= L_0 + \mathbb{R} \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} \end{aligned}$$

where L_0 is the set of attainable random variables, if one starts with zero wealth (as in the proof of the first fundamental theorem). Because X is not attainable, L is a strict subspace of $\mathbb{R}^n$ - and L_0 is smaller than a hyperplane. Therefore, there exists a non-zero vector Z orthogonal to L:

$$\forall Y = \widetilde{V_T(\varphi)} \in L, \ \sum_{\omega \in \Omega} Z(\omega) Y(\omega) = 0$$

Let $\mathbb{Q}$ be a risk-neutral measure. By replacing $Z(\omega)$ by $\frac{Z(\omega)}{\mathbb{Q}(\omega)}$, the previous sum can be rewritten as an expectation under $\mathbb{Q}$:

$$\forall \varphi \text{ predictable, } \mathbb{E}^{\mathbb{Q}}\left[\widetilde{V_T(\varphi)} Z\right] = 0$$

Of course, here we have identified Z with a random variable. Now, one can use Z to "tilt" $\mathbb{Q}$ and create many other equivalent measures:

$$\forall \lambda \in (-1,1),\ \mathbb{Q}^{\lambda}(\omega) := \mathbb{Q}(\omega)\left(1 + \lambda \frac{Z(\omega)}{|Z|_{\infty}}\right)$$

The following facts prove that the family of $\left(\mathbb{Q}^{\lambda}\right)_{\lambda \in (-1,1)}$ is made of Equivalent Martingale measures:

- Each $\mathbb{Q}^{\lambda}$ is a probability:

  $$\mathbb{Q}^{\lambda}(\Omega) = \sum_{\omega \in \Omega} \mathbb{Q}(\omega)\left(1 + \lambda \frac{Z(\omega)}{|Z|_{\infty}}\right) = 1 + \frac{\lambda}{|Z|_{\infty}} \mathbb{E}^{\mathbb{Q}}(Z) = 1$$

  as $\mathbb{E}^{\mathbb{Q}}(Z) = 0$ because the vector with coordinates $Z(\omega)\mathbb{Q}(\omega)$ is orthogonal to the vector $\begin{pmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{pmatrix}$.

- Each $\mathbb{Q}^{\lambda}$ is equivalent to $\mathbb{Q}$: clear.

- Each $\mathbb{Q}^{\lambda}$ is a martingale measure: Thanks to the second point in lemma 4.2, all we have to check is that for all φ predictable with $\widetilde{V_0(\varphi)} = 0$:

  $$\mathbb{E}^{\mathbb{Q}^{\lambda}}\left(\widetilde{V_T(\varphi)}\right) = \mathbb{E}^{\mathbb{Q}^{\lambda}}\left(\widetilde{V_0(\varphi)}\right)$$

  Here we have:

  $$\mathbb{E}^{\mathbb{Q}^{\lambda}}\left(\widetilde{V_T(\varphi)}\right) = \mathbb{E}^{\mathbb{Q}}\left(\left(1 + \lambda \frac{Z(\omega)}{|Z|_{\infty}}\right)\widetilde{V_T(\varphi)}\right) = \mathbb{E}^{\mathbb{Q}}\left(\widetilde{V_T(\varphi)}\right) + \frac{\lambda}{|Z|_{\infty}} \mathbb{E}^{\mathbb{Q}}\left(Z\widetilde{V_T(\varphi)}\right)$$

  Because $\mathbb{Q}$ is an Equivalent martingale measure, the first term is zero. The second term is also zero because every vector with coordinates $\widetilde{V_T(\varphi)}(\omega)$ is orthogonal to the one with coordinates $\mathbb{Q}(\omega)Z(\omega)$.

$\square$

Another take on the second martingale theorem can be "every martingale is a stochastic integral". This is called the martingale representation theorem.

Theorem 4.4. *Consider a no-arbitrage finite market model with $\mathbb{Q}$ a risk neutral measure. The market is complete if and only if every $\mathbb{Q}$-martingale has the representation:*

$$M_t = M_0 + \sum_{s=1}^{t} \widetilde{\varphi}_s \cdot \widetilde{\Delta_{s-1}}$$

with φ a predictable process.

Proof. Seen in exercise class. The idea is that on a finite time horizon, every martingale M is entirely determined by the final value M_T, which itself can be interpreted as a payoff. This payoff being attainable is equivalent to writing it as a discrete stochastic integral. $\square$

Figure 1: A trinomial tree

$$\begin{array}{ccc} & & \widetilde{S_1} = 12 \\ & \nearrow & \\ \widetilde{S_0} = 10 & \rightarrow & \widetilde{S_1} = 10 \\ & \searrow & \\ & & \widetilde{S_1} = 6 \end{array}$$

4.5 A word on incomplete markets

Suppose we have a finite market model with no arbitrage, but that fails to be complete. The simplest example consists in a one-period trinomial tree (figure 4.5).

Let p_1, p_2 and p_3 be risk-neutral probabilities of each branch in the tree. Risk-neutral measures are computed by asking when S is a $\mathbb{Q}$ martingale.

$$\left\{ \begin{array}{rcl} 1 &=& p_1 + p_2 + p_3 \\ 10 &=& \widetilde{S_0} = \mathbb{E}^{\mathbb{Q}}\left(\widetilde{S_1} | \widetilde{S_0} \right) \end{array} \right.$$

$$\iff \left\{ \begin{array}{rcl} 1 &=& p_1 + p_2 + p_3 \\ 1 &=& \frac{3}{5}p_1 + p_2 + \frac{6}{5}p_3 \end{array} \right.$$

The system does not have a unique solution, making the market incomplete thanks to the 2nd fundamental theorem.

5 Panorama of continuous time modeling

This section is mainly here to give you a feeling of continuous time probability. It is not intended to be as thorough and complete as the previous notes. Also, you are not expected to understand it at the same level of detail.

5.1 On continuous time processes

Consider a finite time horizon $T \in \mathbb{R}_+^*$. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a continuous time stochastic process in $\mathbb{R}^d$ will be a family of random variables in $\mathbb{R}^d$ indexed continuously in time:

$$(X_t \ ; \ 0 \leq t \leq T)$$

In this context, a filtration is a continuous family of σ-algebras $(\mathcal{F}_t; \ 0 \leq t \leq T)$ such that $\forall t, s \geq 0$, $\mathcal{F}_t \subset \mathcal{F}_{t+s}$. We will assume as usual that $\mathcal{F}_T = \mathcal{F}$ is the entire σ-algebra and the following usual hypotheses in the theory of stochastic processes:

Hypothesis 5.1 ("Usual hypotheses")**.**

- *$(\Omega, \mathcal{F}, \mathbb{P})$ is complete meaning that if $B \in \mathcal{F}$, with $\mathbb{P}(B) = 0$ then $\mathcal{F}$ contains all subsets of B. Naturally, the probability of $A \subset B$ is zero.*

  Recall from standard measure theory that a measure can always be extended to the completed σ-algebra.

- *$\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets.*

- *$\{\mathcal{F}_t\}_{0 \leq t \leq T}$ is right-continuous:*

$$\mathcal{F}_t = \mathcal{F}_{t^+} := \cap_{s>t} \mathcal{F}_s$$

5.2 Brownian motion

Brownian motion was named after the biologist Brown who described the erratic motion of pollen in water. Its relevance to physics became clear after the works of Einstein who showed that the average displacement of a particle during a period Δt is not $v\Delta t$, v being the speed, but a $\mathcal{O}(\sqrt{\Delta t})$. Mathematically, as we will see, Brownian motion is the most natural path taken at random. Because of that, one can loosely say that most continuous trajectories do not have a speed.

In this section, we denote by $\mathcal{N}$ a generic standard Gaussian random variable.

Definition 5.1. *A standard Brownian motion (on $\mathbb{R}_+$ or $[0,T]$) is a process W such that:*

- $W_0 = 0$ $\mathbb{P}$*-a.s.*
- W *is a.s continuous.*
- W *has independent increments i.e for all* $0 = t_0 < t_1 < \cdots < t_n < \infty$, $W_{t_1} - W_{t_0}$, $W_{t_2} - W_{t_1}$, $\dots$, $W_{t_n} - W_{t_{n-1}}$ *are independent.*
- *For all* $s, t \geq 0$, $W_{t+s} - W_t \overset{\mathcal{L}}{=} \mathcal{N}(0,t) \overset{\mathcal{L}}{=} \sqrt{t}\mathcal{N}$.

Theorem 5.1. *Brownian motion exists and is unique up to indistinguishability.*

In fact, Brownian motion is quite universal and appears as the limit in law of every reasonable random walk properly rescaled.

Theorem 5.2 (Donsker's invariance principle)**.** *Let $(\xi_i)_{i\in\mathbb{N}}$ i.i.d random variables, centered and having finite variance σ^2. Let*

$$B_t^N := \frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nt \rfloor} \xi_i$$

Then the following weak convergence holds:

$$\left(B_t^N; t \geq 0\right) \xrightarrow{\mathcal{L}} (\sigma W_t; t \geq 0)$$

where W is a BM.

Remark 5.1. *Let $0 = t_0 < t_1 < \cdots < t_n = T$ be a regular subdivision of $[0,T]$. Then $t_i - t_{i-1} = \frac{T}{n}$ and:*

$$\begin{aligned}
\sum_{i=1}^n \left|W_{t_i} - W_{t_{i-1}}\right| &\overset{\mathcal{L}}{=} \sum_{i=1}^n \sqrt{t_i - t_{i-1}}\, |\mathcal{N}_i| \\
&= \sqrt{Tn} \frac{\sum_{i=1}^n |\mathcal{N}_i|}{n} \\
&\overset{n\to\infty}{\sim} \sqrt{Tn}\mathbb{E}\left(|\mathcal{N}|\right)
\end{aligned}$$

by the law of large numbers. This quantity is clearly divergent, showing that one cannot sum the increments of Brownian motion in absolute value.

5.3 Stochastic integrals

Classical integration: Recall that a function $A : \mathbb{R}_+ \to \mathbb{R}$ is of finite variation (on compact sets) if for all $T > 0$, the variation

$$\sum_i \left| A_{t_{i+1}} - A_{t_i} \right|$$

is uniformly bounded for all subdivisions $0 = t_0 < t_1 < \cdots < t_n = T$. Then the supremum over all subdivisions is a finite quantity called the total variation of A over $[0, T]$. The Riemann-Stieljes integral $\int_0^t \varphi_s dA_s$ if φ is, say, measurable and bounded while A has finite variation.

This construction is equivalent to the integration against signed measures, seen in measure theory class. There is a correspondence between signed measures and finite variation functions, which we now describe. Every σ-finite signed measure μ on $\mathbb{R}_+^*$ is entirely described by the function A^μ defined as:

$$A_t^\mu = \int_{)0,t]} \mu(ds)$$

The function A^μ is necessarily of finite variation. In fact, an almost tautological statement is that if $|\mu|$ is the total variation of μ i.e the positive measure

$$\forall B \in Bor(\mathbb{R}_+^*),\ |\mu|(B) = \sup_{B = \bigsqcup B_i} \sum_i |\mu(B_i)|$$

then the total variation of A^μ is the increasing function $A^{|\mu|}$.

If we want to make sense of $\int_0^t \varphi_s dW_s$, one needs a different contruction because Brownian motion is not of finite variation (see remark 5.1).

5.3.1 Stochastic integral with respect to Brownian motion

Let W be a standard BM. The goal is to define for a large family of stochastic processes Y, the stochastic integral

$$\left(\int_0^t Y_s dW_s ; 0 \leq t \leq T \right)$$

The construction goes through a very classical approximation scheme. First, we define the stochastic integral against simple processes. The class of simple processes will be denoted $\mathcal{L}_s$. Then, we will extend it to a class of square integrable processes $\mathcal{L}^2$. Of course, $\mathcal{L}_s \subset Lc^2$ and the inclusion is dense for the Lc^2 norm that is to be defined.

Morally, the same goes for the construction of integrals against Radon measures: it defined against linear combinations of indicators of measurable sets, then extended using the fact that every measurable function is an almost sure limit of simple functions.

Of course, one cannot hope to beyond classical integration without assuming somehow that process Y has good measurability properties. The idea is to require that Y does not "look into the future", by being for example adapted. A weaker hypothesis is the notion of *progressive measurability.*

Definition 5.2. *A process Y is progressively measurable with respect to a filtration $\mathbb{F}$ if for all $T > 0$, the map:*

$$\begin{array}{ccc} [0,T] \times \Omega & \to & \mathbb{R} \\ (t, \omega) & \mapsto & X_t(\omega) \end{array}$$

is $Bor\left([0,T]\right) \otimes \mathbb{F}_T$ measurable.

Our class of simple processes will be:

$$\mathcal{L}_s := \left\{ Y_t(\omega) = H_0(\omega)\mathbb{1}_{\{0\}}(t) + \sum_{i=1}^{n-1} H_i(\omega)\mathbb{1}_{)t_i;t_{i+1}]}(t),\ H_i(\omega) \text{ bounded}, \mathcal{F}_{t_i} \text{ measurable} \right\}$$

Basically, simple processes are step functions with random values, measurable with respect to the filtration at the time of the jump. Then, for the purposes of this lecture, the following family is enough for our needs

$$\mathcal{L}^2 := \left\{ Y \text{progressively measurable} \mid \mathbb{E}\left(\int_0^T Y_s^2 ds \right) < \infty \right\}$$

Remark 5.2. *The letter $\mathcal{L}^2$ has a different meaning than $\mathbb{L}^2$.*

$$\mathbb{L}^2 = L^2(\Omega) := \left\{ X\ \mathcal{F} - measurable \mid \mathbb{E}\left(X^2\right) < \infty \right\}$$

whereas $\mathcal{L}^2$ is a subset of

$$\mathbb{L}^2\left([0,T] \times \Omega\right) := \left\{ X\ Bor([0,T]) \otimes \mathcal{F} - measurable \mid \mathbb{E}\left(\int_0^T X_s^2 ds \right) < \infty \right\}$$

Properties 5.1 (Assumed)**.**

- *Every adapted process with left or right-continuous paths is progressively measurable.*
- *$\mathcal{L}_s \subset \mathcal{L}^2$ and the inclusion is dense for the norm:*

$$|Y|_{\mathcal{L}^2} = \mathbb{E}\left[\int_0^T Y_s^2 ds \right]^{\frac{1}{2}}$$

Theorem 5.3 (Construction of the stochastic integral)**.** *For any $Y \in \mathcal{L}^2$, there exists a process:*

$$\left(\int_0^t Y_s ds\ ;\ 0 \le t \le T \right)$$

such that

(i) $\int_0^t Y_s dW_s$ is a continuous $\mathbb{L}^2$-martingale with respect to $\mathbb{F}$.

(ii) Itô isometry:

$$\forall t \in [0,T],\ \mathbb{E}\left(\left(\int_0^t Y_s dW_s \right)^2 \right) = \mathbb{E}\left(\int_0^t Y_s^2 ds \right)$$

The two properties give above are the continuous analogues of the following properties, for every $\mathbb{L}^2$-martingale:

(i') For every bounded φ predicable and bounded,

$$M(\varphi)_t := \sum_{s=1}^{t} \varphi_s \left(M_s - M_{s-1} \right)$$

is an $\mathbb{L}^2$-martingale.

(ii') The $\mathbb{L}^2$-decomposition

$$\mathbb{E}\left(\left(M_t\left(\varphi\right)\right)^2\right) = \sum_{s=1}^{t} \mathbb{E}\left(\varphi_s\left(M_s - M_{s-1}\right)^2\right)$$

Proof of discrete properties (i') and (ii'). (i') The fact that we have a martingale is the first point of lemma 4.2. In order to check it is in $\mathbb{L}^2$, it suffices to notice that every term is in $\mathbb{L}^2$, which is a vector space.

(ii') The $\mathbb{L}^2$-decomposition is proven by expanding the square and keeping only the diagonal terms. The off-diagonal terms vanish using the tower property and the fact that M is a martingale.

□

Proof of theorem "Construction of stochastic integrals". As announced, we proceed by density. Let $Y \in \mathcal{L}_s$ be a simple process. By definition, it is of the form:

$$Y_t = H_0 \mathbb{1}_{\{0\}}(t) + \sum_{i=1}^{n-1} H_i \mathbb{1}_{)t_i;t_{i+1}]}(t)$$

and it is natural to define the stochastic integral as:

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

Clearly, the subdivision in time $t_1 < t_2 < \cdots < t_n$ can be refined at will, and still yields the same definition. Let us now prove that this construction satisfies properties (i) and (ii), when the integrand is a simple process.

(i) for $\mathcal{L}_s$ The stochastic integral is an $\mathbb{L}^2$ martingale. It is easy to see that $\int_0^t Y_s dW_s \in \mathbb{L}^2$ for all t. In order to see it is a martingale, we have to check that

$$\mathbb{E}\left(\int_0^{t+s} Y_u dW_u | \mathcal{F}_t\right) = \int_0^t Y_u dW_u$$

for every fixed t and s. As mentioned before, one can refine the subdivision if necessary and suppose that t and $t+s$ are part of it. Therefore, we only have to prove that the sequence $\int_0^{t_i} Y_u dW_u$, $i = 0, 1, \dots$ is a discrete martingale. This is immediate, because it is a discrete stochastic integral with respect to the discrete martingale W_{t_i}!

(ii) for $\mathcal{L}_s$ Thanks to exactly the same computation as the property (ii') in discrete time

$$\begin{aligned}
&\mathbb{E}\left(\left(\int_0^t Y_s dW_s\right)^2\right)\\
=&\sum_{i=1}^{n-1} \mathbb{E}\left(H_i^2\left(W_{t\wedge t_{i+1}} - W_{t\wedge t_i}\right)^2\right)\\
=&\sum_{i=1}^{n-1} \mathbb{E}\left(H_i^2 \mathbb{E}\left(\left(W_{t\wedge t_{i+1}} - W_{t\wedge t_i}\right)^2 | \mathcal{F}_{t_i}\right)\right)\\
=&\sum_{i=1}^{n-1} \mathbb{E}\left(H_i^2\left(t\wedge t_{i+1} - t\wedge t_i\right)\right)
\end{aligned}$$

$$
\begin{aligned}
&=\mathbb{E}\left(\sum_{i=1}^{n-1}\int_{t\wedge t_{i+1}}^{t\wedge t_i} Y_s^2 ds\right)\\
&=\mathbb{E}\left(\int_0^t Y_s^2 ds\right)
\end{aligned}
$$

Now we use the fact that $\mathcal{L}_s \subset \mathcal{L}^2$ and that this inclusion is dense with respect to the norm $|.|_{\mathcal{L}^2}$. The $\mathbb{L}^2$-isometry tells us that the stochastic integral for fixed t is a continuous map from $\mathcal{L}^2$ to $\mathbb{L}^2$. Therefore, it extends uniquely. However, we will need more.

Consider a sequence $Y^n \in \mathcal{L}_s$ converging to $Y \in \mathcal{L}^2$, in the $|.|_{\mathcal{L}^2}$-norm. As mentioned, the Ito isometry tells us that for every fixed t,

$$
\left(\int_0^t Y_s^n dW_s ; n \in \mathbb{N}\right)
$$

is a Cauchy sequence in $\mathbb{L}^2(\Omega)$. It therefore converges to a unique random variable. This is what we define as $\int_0^t Y_s dW_s$. This however sheds limited information on the process $t \mapsto \int_0^t Y_s dW_s = I(Y)_t$ as t varies. We still have to argue that:

- $t \mapsto I(Y)_t$ is an $\mathbb{L}^2$ martingale as it is an $\mathbb{L}^2$-limit of the $\mathbb{L}^2$-martingales $I(Y^n)$.
- Continuity is a little tricky and is only given for the sake of completeness. It uses tools we have not really got the chance of seeing. The idea is to extract a subsequence, n_k such that the sequence of random functions $\left(t \mapsto \int_0^t Y_s^{n_k} dW_s ; 0 \leq t \leq T\right)$ converge uniformly. The limit $\left(t \mapsto \int_0^t Y_s dW_s ; 0 \leq t \leq T\right)$ is therefore continuous as a uniform limit of continuous functions. These statements are for fixed ω.

  Thanks to Doob's maximal inequality:

$$
\mathbb{E}\left(\sup_{0\leq t\leq T}\left(\int_0^t \left(Y_s^n - Y_s^m\right) dW_s\right)^2\right) \leq 4\mathbb{E}\left(\int_0^T \left(Y_s^n - Y_s^m\right)^2 ds\right)
$$

  which goes to zero as $Y^n \in \mathcal{L}^2$ is a Cauchy sequence. One can extract a subsequence n_k such that $\mathbb{E}\left(\sup_{0\leq t\leq T}\left(\int_0^t \left(Y_s^{n_{k+1}} - Y_s^{n_k}\right) dW_s\right)^2\right)$ is summable. Hence,

$$
\sum_k \sup_{0\leq t\leq T}\left(\int_0^t \left(Y_s^{n_{k+1}} - Y_s^{n_k}\right) dW_s\right)^2 = \sum_k \left(\sup_{0\leq t\leq T} I\left(Y^{n_{k+1}}\right)_t - I\left(Y^{n_k}\right)_t\right)^2 < \infty
$$

  almost surely and therefore, $t \mapsto I\left(Y^{n_k}\right)_t$ is a Cauchy sequence for the supremum norm.

□

The multi-dimensional generalisation of the stochastic integral is straightforward. If Y is a d-dimensional process with:

$$
\mathbb{E}\left(\int_0^T |Y_s|^2 ds\right) < \infty
$$

then we define the integral against a d-dimensional Brownian motion as:

$$
\int_0^t Y_s \cdot dW_s := \sum_{i=1}^d \int_0^t Y^i \cdot dW_s^i
$$

5.3.2 Stochastic integral with respect to Ito processes

Definition 5.3. *An Itô process X driven by a d-dimensional Brownian motion is a continuous an adapted process such that:*

$$X_t = X_0 + \int_0^t Y_s \cdot dW_s + \int_0^t Z_s ds$$

with Y, Z progressively measurable processes and:

$$Y \in \mathcal{L}^2 \iff \mathbb{E}\left(\int_0^T |Y_s|^2 \, ds\right) < \infty$$

$$Z \in \mathcal{L}^1 \iff \mathbb{E}\left(\int_0^T |Z_s| \, ds\right) < \infty$$

Remark 5.3 (Short-hand notation)**.** *One often writes:*

$$dX_t = Y_t dW_t + Z_t dt$$

which is only a short-hand notation. It is reminiscent of the Leibniz notation for the differential and not completely unrelated. Indeed, if $Y = 0$, then

$$X_t = X_0 + \int_0^t Z_s ds$$

Hence X is differentiable and

$$\frac{dX_t}{dt} = X'_t = Z_t$$

which can also be written

$$dX_t = Z_t dt$$

Proposition 5.1. *The quadratic variation process $\langle X, X\rangle_t = \int_0^t Z_s^2 ds$ appears as the limit in probability of*

$$\sum_i \left|X_{t_{i+1}} - X_{t_i}\right|^2$$

along subdivisions $0 = t_0 < t_1 < \cdots < t_n = t$.

It is an important quantity that appears in the Ito formula. Notice that the quadraic variation is a finite variation process and therefore the measure $d\langle X, X\rangle_t$ exists.

Example 5.1 (Brownian motion)**.** *Clearly, Brownian motion itself is an Itô process ($Y = 1$ and $Z = 0$) in the definition. By a simple application of the law of large numbers, one sees that $\langle W, W\rangle_t = t$.*

5.3.3 Ito formula

Although we constructed the stochastic integral, we cannot compute exactly any stochastic integral without some algebraic rules. Notice that the same is true when dealing with classical measure theory: after constructing the integral against the Lebesgue measure, one still cannot compute any integral. One first needs to prove the fundamental theorem of analysis i.e that for every function X that is C^1

$$X_t = X_0 + \int_0^t X'_s ds = X_0 + \int_0^t dX_s \ .$$

This combined with the chain rule:

$$F\left(X_t\right) = F(X_0) + \int_0^t F'(X_s)dX_s$$

allows to compute formulas such as ($X_t = t, F(x) = \frac{x^2}{2}$):

$$\int_0^t s\ ds = \frac{t^2}{2}$$

The Itô formula can be thought of as a generalisation of the chain rule for Ito processes. More formally, it gives the decomposition of $t \mapsto F(X_t)$ as an Ito process, when X is itself an Ito process.

Theorem 5.4 (Itô formula)**.** *Let X be a one dimensional Itô process. If:*

$$\begin{array}{rccc} F: & \mathbb{R}_+ \times \mathbb{R} & \to & \mathbb{R} \\ & (t,x) & \mapsto & f(t,x) \end{array}$$

is $C^{1,2}$ (C^1 in t and C^2 in x) then:

$$\begin{aligned} F(t,X_t) &= F(0,X_0) + \int_0^t \frac{\partial F}{\partial t}(s,X_s)ds + \int_0^t \frac{\partial F}{\partial x}(s,X_s)dX_s + \frac{1}{2}\int_0^t \frac{\partial^2 F}{\partial x^2}(s,X_s)d\langle X,X\rangle_s \\ &= F(0,X_0) + \int_0^t \frac{\partial F}{\partial t}(s,X_s)ds + \frac{\partial F}{\partial x}(s,X_s)Z_s ds + \int_0^t \frac{\partial F}{\partial x}(s,X_s)Y_s dW_s \\ &\quad + \frac{1}{2}\int_0^t \frac{\partial^2 F}{\partial x^2}(s,X_s)Z_s^2 ds \end{aligned}$$

A Conditional expectation (Appendix)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, X an integrable random variable and $\mathcal{G} \subset \mathcal{F}$ a σ-algebra.

Informally, the conditional expectation $\mathbb{E}(X|\mathcal{G})$ is the $\mathcal{G}$-measurable random variable that best approximates X. The following proposition can be thought of as a definition:

Proposition A.1. *There exists an unique (up to $\mathbb{P}$-equivalence) random variable $Y := \mathbb{E}(X|\mathcal{G})$ such that:*

- *Y is $\mathcal{G}$-measurable.*
- *$Y \in \mathcal{L}^1$ i.e $\mathbb{E}(|Y|) < \infty$.*
- *For all $A \in \mathcal{G}$, $\mathbb{E}\left(Y \mathbb{1}_{\{A\}}\right) = \mathbb{E}\left(X \mathbb{1}_{\{A\}}\right)$.*

Proof of proposition A.1. For the uniqueness up to $\mathbb{P}$-equivalence, consider Y' another random variable satisfying the same properties. Then for all $A \in \mathcal{G}$, $\mathbb{E}((Y - Y')\mathbb{1}_A) = 0$. Since $Y - Y'$ is $\mathcal{G}$-measurable, take successively $A = \{Y - Y' \geq 0\} \in \mathcal{G}$ and $A = \{Y - Y' < 0\} \in \mathcal{G}$ in order to obtain $\mathbb{E}\left((Y - Y')\mathbb{1}_{\{Y-Y' \geq 0\}}\right) = 0$ and $\mathbb{E}\left(-(Y - Y')\mathbb{1}_{\{Y-Y'<0\}}\right) = 0$. By forming the sum, we have $\mathbb{E}(|Y - Y'|) = 0$. Hence $Y = Y'$, $\mathbb{P}$-almost surely.

In order to prove existence, we proceed in three steps. First, write $X = X^+ - X^-$. By treating positive and negative parts separately, one can suppose $X \geq 0$. Technically, we have just used the linearity (property 1 in A.1) of conditional expectation, which is a consequence of the uniqueness statement.

Second, if $X \in \mathcal{L}^2$, we perform a classical construction in convex analysis. Consider the inclusion $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P}) \subset \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ of Hilbert spaces for the scalar product $\langle X, Y \rangle = \mathbb{E}(XY)$. This is a closed inclusion. From the general fact that there exists an orthogonal projection on convex subsets of Hilbert spaces, we deduce there is a map $p : \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}) \to \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ such that:

$$\forall (X, Y) \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}) \times \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P}), \ \mathbb{E}((X - p(X))Y) = 0$$

If $Y = \mathbb{1}_A$ then $p(X)$ satisfies the required properties.

Finally, in the general case, consider the increasing approximation $X_n = \min(X, n)$ which are bounded, therefore in $\mathcal{L}^2$. The subsequent properties A.1 are therefore valid for X_n. $Y_n = \mathbb{E}(X_n|\mathcal{G})$ is also increasing as $Y_n + 1 - Y_n = \mathbb{E}(Y_n + 1 - Y_n|\mathcal{G}) \geq 0$, thanks to the positivity property. Therefore, Y_n converges almost surely to a limit Y. Now, for all $A \in \mathcal{G}$:

$$\begin{aligned}
\mathbb{E}(Y\mathbb{1}_A) &= \lim \mathbb{E}(Y_n \mathbb{1}_A) \ \text{ by monotone convergence} \\
&= \lim \mathbb{E}(X_n \mathbb{1}_A) \\
&= \mathbb{E}(X\mathbb{1}_A) \ \text{ by dominated convergence}
\end{aligned}$$

□

Moreover, conditional expectation has the following properties:

Properties A.1. *1. Linearity and positivity:*

$$\forall (a, b) \in \mathbb{R}^2, \ \mathbb{E}(aX + bY|\mathcal{G}) = a\mathbb{E}(X|\mathcal{G}) + b\mathbb{E}(Y|\mathcal{G})$$

$$X \geq 0 a.s \Rightarrow \mathbb{E}(X|\mathcal{G}) \geq 0 a.s$$

along with $\mathbb{E}(|\mathbb{E}(X|\mathcal{G})|) \geq \mathbb{E}(|X|)$.

2. *The tower property: If $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$ is an inclusion of σ-algebras, then*

$$\mathbb{E}\left(\mathbb{E}\left(X|\mathcal{G}\right)|\mathcal{H}\right) = \mathbb{E}\left(X|\mathcal{H}\right)$$

3. *If Z is $\mathcal{G}$-measurable and bounded, then:*

$$\mathbb{E}\left(ZX|\mathcal{G}\right) = Z\mathbb{E}\left(X|\mathcal{G}\right)$$

4. *If X is $\mathcal{G}$-measurable then*

$$\mathbb{E}\left(X|\mathcal{G}\right) = X$$

5. *If X is independent of $\mathcal{G}$:*

$$\mathbb{E}\left(X|\mathcal{G}\right) = \mathbb{E}\left(X\right)$$

Proof. 1. For linearity, let $Z = a\mathbb{E}\left(X|\mathcal{G}\right) + b\mathbb{E}\left(Y|\mathcal{G}\right)$ be the candidate. Because of the linearity of the usual expectation, for all $A \in \mathcal{G}$, $\mathbb{E}\left(\mathbb{1}_A\left(aX + bY\right)\right) = a\mathbb{E}\left(\mathbb{1}_A X\right) + b\mathbb{E}\left(\mathbb{1}_A Y\right) = a\mathbb{E}\left(\mathbb{1}_A \mathbb{E}\left(X|\mathcal{G}\right)\right) + b\mathbb{E}\left(\mathbb{1}_A \mathbb{E}\left(Y|\mathcal{G}\right)\right) = \mathbb{E}\left(\mathbb{1}_A Z\right)$. As Z is integrable and $\mathcal{G}$-measurable, it must be $\mathbb{E}\left(aX + bY|\mathcal{G}\right)$ by uniqueness.

   For positivity, consider $A = \{\mathbb{E}\left(X|\mathcal{G}\right) < 0\} \in \mathcal{G}$. If X is non-negative, then $0 \leq \mathbb{E}\left(\mathbb{1}_A X\right) = \mathbb{E}\left(\mathbb{1}_A \mathbb{E}\left(X|\mathcal{G}\right)\right) = -\mathbb{E}\big(\mathbb{E}\left(X|\mathcal{G}\right)^-\big)$. Therefore $\mathbb{E}\left(X|\mathcal{G}\right)^- = 0$ almost surely and the result holds.

   For the control on the L^1 norm, introduce $B = \{\mathbb{E}\left(X|\mathcal{G}\right) \geq 0\} \in \mathcal{G}$. We have that $\mathbb{E}\left(|\mathbb{E}\left(X|\mathcal{G}\right)|\right) = \mathbb{E}\left(\mathbb{1}_B \mathbb{E}\left(X|\mathcal{G}\right) - \mathbb{1}_A \mathbb{E}\left(X|\mathcal{G}\right)\right) = \mathbb{E}\left(\mathbb{1}_B X - \mathbb{1}_A X\right) \leq \mathbb{E}\left(|X|\right)$.

2. Very good exercise.

3. The statement is easy when Z is an indicator function, for example $\mathbb{1}_B$ with $B \in \mathcal{G}$. Then use a standard approximation argument.

4. Consequence of 3) if X bounded, or simply by checking that the random variable X satisfies the required hypotheses and invoking uniqueness once again.

5. Because of independence, for all $A \in \mathcal{G}$, $\mathbb{E}\left(\mathbb{1}_A X\right) = \mathbb{P}\left(A\right)\mathbb{E}\left(X\right) = \mathbb{E}\left(\mathbb{1}_A \mathbb{E}(X)\right)$. Therefore, the constant random variable $\mathbb{E}(X)$ satisfies all the hypotheses.

$\square$

B Martingales (Appendix)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space along with a filtration:

$$\mathcal{F} = \{\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_T = \mathcal{F}\}$$

Then $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space. Loosely speaking, a martingale is the mathematical formalisation of a "fair gain" process or a process that is constant on average, conditionally. For the purposes of the class, only the definition is required:

Definition B.1. *A martingale $(M_t)_{t\in\mathbb{N}}$ (resp. a super-martingale or sub-martingale) is a process such that:*

- $\forall t \in \mathbb{N}$, *M_t is integrable.*
- $(M_t)_{t\in\mathbb{N}}$ *is $\mathbb{F}$-adapted.*
- $\forall s \geq 0, \mathbb{E}\left(M_{t+s} \mid \mathcal{F}_t\right) = M_t$ *(resp. $\geq, \leq$)*

Because of the tower property, the last statement in the definition of a martingale can be replaced by:

$$\forall t \in \mathbb{N}, \mathbb{E}\left(M_{t+1}|\mathcal{F}_t\right) = M_t$$

An immediate consequence of the definition is that the expectation of a martingale is constant. Also, a simple class of martingales is given by *closed martingales*, which are of the form:

$$M_t = \mathbb{E}\left(X|\mathcal{F}_t\right)$$

for any integrable random variable X. In a finite time horizon, all martingales are closed.

Remark B.1. *If the choice of underlying filtration or measure is ambiguous, one lifts the ambiguity by saying that M is an $\mathbb{F}$-martingale and a $\mathbb{P}$-martingale.*

There are classical theorems concerning martingales such as Doob's inequality or convergence theorems. These are left for a real class on the subject. Nevertheless, the good behavior of martingales regarding stopping times will be needed in the modeling of American options.

Definition B.2. *A discrete stopping time is a random variable $\tau : \Omega \to \{0, 1, 2, \ldots, T\} \sqcup \{\infty\}$ such that one of the following three equivalent conditions hold:*

- $\forall t, \{\tau = t\} \in \mathcal{F}_t$
- $\forall t, \{\tau \leq t\} \in \mathcal{F}_t$
- $\forall t, \{\tau > t\} \in \mathcal{F}_t$

These three conditions are seen to be equivalent using the fact that $\mathbb{F}$ is a filtration and the stability of σ-algebras with respect to union, intersection and complement.

Properties B.1.

- *Deterministic times are stopping times.*
- *If τ and τ' are stopping times, then the same goes for their maximum $\tau \vee \tau'$ and their minimum $\tau \wedge \tau'$.*

Proof. If τ is deterministic, then $\{\tau = t\}$ is either empty or all of Ω, and therefore belongs to any sigma algebra. For the second property, we will consider only the maximum, as the minimum is treated in a similar fashion. For fixed $t \in \mathbb{N}$, we write $\{\tau \wedge \tau' \leq t\} = \{\tau \leq t\} \cup \{\tau' \leq t\} \in \mathcal{F}_t$. $\square$

The σ-algebra generated by a stopping time τ is defined as the set of events whose trace on $\{\tau = t\}$ is in $\mathcal{F}_t$.

Definition B.3.

$$\mathcal{F}_\tau := \{A \in \mathcal{F} \mid \forall t =, A \cap \{\tau = t\} \in \mathcal{F}_t\}$$

Exercise B.1. *Prove this is a σ-algebra.*

For a process X, let us denote by X^τ the process X stopped at τ i.e:

$$\forall t,\ X_t^\tau := X_{\tau \wedge t}$$

The reason why stopping times are important is embodied by the following theorem

Theorem B.1 (Doob's optional stopping theorem).

- *If τ is a stopping time and M is a martingale (resp. a super or sub-martingale), then so is M^τ.*
- *If τ is bounded almost surely ($\exists c > 0, \tau \leq c$ a.s) then*

Proof. For the first statement, notice that for all $t \geq 0$, $M_{t+1}^\tau - M_t^\tau = \mathbb{1}_{\{\tau > t\}} (M_{t+1} - M_t)$. Then, because $\{\tau > t\} \in \mathcal{F}_t$, we have $\mathbb{E}\left(M_{t+1}^\tau - M_t^\tau | \mathcal{F}_t\right) = \mathbb{1}_{\{\tau > t\}} \mathbb{E}\left(M_{t+1} - M_t | \mathcal{F}_t\right) = 0$ (resp. ≥ 0, ≤ 0).

For the second statement, let us suppose for example that M is a martingale. Thanks to Doob's optional stopping theorem, M^τ is also a martingale. Then $\mathbb{E}\left(M^\tau\right) = \mathbb{E}\left(M_{\tau \wedge c}\right) = \mathbb{E}\left(M_c^\tau\right) = \mathbb{E}\left(M_0\right)$. $\square$

Remark B.2. *The typical counter-example to the second statement in Doob's optional stopping theorem consists in doubling strategies. Consider a fair coin-flip game, where player bets a certain amount and wins double. If the player doubles his bet everytime he looses, we will end up winning his initial bet at a certain stopping time τ. Clearly, τ is a geometric random variable, hence not bounded. The gain process $(G_t)_{t \in \mathbb{N}}$ is a martingale. But G_τ is the initial bet.*

References

[Fri07] Christian Fries. *Mathematical finance.* Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2007. Theory, modeling, implementation.

[Hul06] John C. Hull. *Options, futures, and other derivatives.* Pearson Prentice Hall, Upper Saddle River, NJ [u.a.], 6. ed., pearson internat. ed edition, 2006.

[LL08] Damien Lamberton and Bernard Lapeyre. *Introduction to stochastic calculus applied to finance.* Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, second edition, 2008.

[Wil06] R. J. Williams. *Introduction to the mathematics of finance*, volume 72 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2006.