Contents

Preface

The present lecture notes are based on the following literature.

- F. Delbaen and W. Schachermayer. *The mathematics of arbitrage*. Springer Finance. Springer-Verlag, Berlin, 2006.
- H. Föllmer and A. Schied. *Stochastic finance*. Walter de Gruyter & Co., Berlin, extended edition, 2011. An introduction in discrete time.
- S. E. Shreve. *Stochastic calculus for finance. I*. Springer Finance. Springer-Verlag, New York, 2004. The binomial asset pricing model.

Throughout we consider models of financial markets in *discrete time*, i.e., trading is only allowed at discrete time points $0 = t_0 < t_1 < \cdots < t_N = T$. Here, $T > 0$ denotes a finite time horizon. This is in contrast to models in continuous time, where continuous trading during the interval $[0, T]$ is possible.

The following topics of mathematical finance will be covered:

- arbitrage theory;
- completeness of financial markets;
- superhedging;
- pricing of derivatives (European and American options);
- concrete modeling of financial markets via the Binomial asset price model and (its convergence to) the Black Scholes model.

From a mathematical point of view, probability theory and stochastic analysis play a key role in mathematical finance.

Chapter 1

Basic notions from probability theory

We recall here basic notions from probability theory which we will need for modeling financial markets.

1.1 Filtered probability spaces, random variables and stochastic processes

Let us start by recalling the ingredients of a probability space. A probability space consists of three parts:

- a non-empty set Ω (Ergebnismenge), which is the set of possible outcomes;
- a σ-algebra $\mathcal{F}$, i.e., a set consisting of sets of Ω to model all possible events (Ereignisse) (where an event is a set containing zero or more outcomes);
- a probability measure P assigning probabilities to each event.

The precise mathematical definition of these notions are as follows:

Definition 1.1.1. *A set $\mathcal{F} \subset \mathcal{P}(\Omega)$ is called* σ-algebra *if it satisfies*

- $\Omega \in \mathcal{F}$*;*
- $A \in \mathcal{F} \Rightarrow A^c = \Omega \setminus A \in \mathcal{F}$*;*
- $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} \in \mathcal{F}$.

The above definition implies that a σ-algebra is closed under countable intersections.

Definition 1.1.2. *Let $(\Omega, \mathcal{F})$ be a measurable space, i.e. $\mathcal{F}$ is σ-algebra on Ω. Then a probability measure is a function $P : \mathcal{F} \to [0, 1]$ such that*

- $P[\Omega] = 1$;
- *it is σ-additive, i.e. for any sequence of pairwise disjoint sets in $\mathcal{F}$ (i.e., $A_n \cup A_m = \emptyset$ for $n \neq m$), we have $P[\bigcup_{n=1}^{\infty} A_n] = \sum_{n=1}^{\infty} P[A_n]$.*

Definition 1.1.3. • *Two probability measures P, Q are called equivalent, which is denoted by $P \sim Q$ if*

$$P[A] = 0 \Leftrightarrow Q[A] = 0, \quad A \in \mathcal{F}.$$

- *Q is absolutely continuous with respect to P, which is denoted by $Q \ll P$ if*

$$P[A] = 0 \Rightarrow Q[A] = 0, \quad A \in \mathcal{F}.$$

Remark 1.1.4. • From the above definition, we immediately get

$$Q \sim P \Leftrightarrow P \ll Q, Q \ll P.$$

and

$$Q \ll P \Leftrightarrow Q[A] > 0 \Rightarrow P[A] > 0.$$

- In the case when Ω consists of finitely many elements and $P[\{\omega\}] > 0$ for every ω, then for every probability measure Q we have $Q \ll P$. Equivalence means $Q[\{\omega\}] > 0$ for every ω.

Let us recall the notion of an atom:

Definition 1.1.5. *Given a probability space $(\Omega, \mathcal{F}, P)$, then a set A is called atom if $P[A] > 0$ and for any measurable subset $B \subset A$ with $P[B] < P[A]$ we have $P[B] = 0$. In the case of a finite probability space where only the empty set has probability zero, we have the following equivalent definition a set A is called atom if $P[A] > 0$ and for any measurable subset $B \subset A$ with $P[B] < P[A]$ we have $B = \emptyset$.*

Example 1.1.6. Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and $\mathcal{F} = \mathcal{P}(\Omega)$. Consider a probability measure P which satisfies $P[\omega_i] > 0$. Then the atoms are $\{\omega_i\}$, $i \in \{1, \ldots, 4\}$. If the σ-Algebra is given by $\mathcal{F} = \{\emptyset, \Omega, \{\omega_1, \omega_2\}, \{\omega_3, \omega_3\}\}$, then the atoms are $\{\omega_1, \omega_2\}$ and $\{\omega_3, \omega_4\}$.

Definition 1.1.7. *A family of σ-algebras with $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \mathcal{F}_T$ is called* filtration *and $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{\{t \in [0,\ldots,T]\}}, P)$ filtered probability space.*

Remark 1.1.8. $\mathcal{F}_t$ is interpreted as the set of all events which can happen up to time t or equivalently as the information which is available up to time t.

Assumption. *Unless explicitly mentioned, we shall assume that $\mathcal{F}_T = \mathcal{F}$. We do not assume $\mathcal{F}_0$ to be necessarily the trivial σ-algebra $(\emptyset, \Omega)$, although in many applications this is the case.*

For modeling asset prices we consider *stochastic processes* which are families of *random variables*, whose definition we recall subsequently.

Definition 1.1.9. *Let $(\Omega, \mathcal{F})$ and $(E, \mathcal{E})$ be two measurable spaces. A* random variable *X with values in E is a $(\mathcal{F}$-$\mathcal{E})$-measurable function $X : \Omega \to E$, i.e. the preimage of any measurable set $B \in \mathcal{E}$ is in $\mathcal{F}$: $\forall B \in \mathcal{E}$, we have $X^{-1}(B) \in \mathcal{F}$.*

In our setting $(E, \mathcal{E})$ is typically $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, where $\mathcal{B}(\mathbb{R}^n)$ denotes the Borel σ-algebra, defined as the smallest σ-algebra containing the open sets of $\mathbb{R}^n$.

Remark 1.1.10. In the case $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, $(\mathcal{F}$-$\mathcal{B}(\mathbb{R}))$-measurability (or simply $\mathcal{F}$-measurability) is equivalent to

$$\forall a \in \mathbb{R} : \{\omega \in \Omega : X(\omega) \in (-\infty, a]\} \in \mathcal{F}.$$

Definition 1.1.11. *Let Ω be some set and $(E, \mathcal{E})$ be a measurable spaces. Consider a function $X : \Omega \to E$. Then the σ-algebra generated by X, denoted by $\sigma(X)$, is the collection of all inverse images $X^{-1}(B)$ of the sets B in $\mathcal{E}$, i.e.,*

$$\sigma(X) = \{X^{-1}(B) \,|\, B \in \mathcal{E}\}.$$

Definition 1.1.12. *Let $\mathcal{T}$ be an index set, either $\{0, 1, \ldots, T\}$ or $\{1, \ldots, T\}$, and $(\Omega, \mathcal{F})$ and $(E, \mathcal{E})$ two measurable spaces. A stochastic process with values in $(E, \mathcal{E})$ is a family of random variables $X = (X_t)_{t \in \mathcal{T}} = \{X_t \,|\, t \in \mathcal{T}\}$ (i.e. $\mathcal{F}$-measurable).*

Definition 1.1.13. *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \{0,1\ldots,T\}}, P)$ a filtered probability space.*

1. *A stochastic process X is called* adapted *with respect to the filtration $(\mathcal{F}_t)$ if for every $t \in \{0, 1, \ldots, T\}$, X_t is $\mathcal{F}_t$-measurable.*

2. *A stochastic process Y is called* predictable *with respect to the filtration $(\mathcal{F}_t)$ if for every $t \in \{1, \ldots, T\}$, Y_t is $\mathcal{F}_{t-1}$-measurable.*

Example 1.1.14. Let $T = 2$, $\Omega = \{1, 2, 3, 4\}$ and $E = \mathbb{R}$. Consider the following filtration $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_1 = \{\emptyset, \Omega, \{1, 2\}, \{3, 4\}\}$ and $\mathcal{F}_2 = \mathcal{P}(\Omega)$. Question: How do adapted stochastic processes look like? Answer: For $t = 0$, a ($\mathcal{F}_0$-measurable) random variable is constant, for $t = 1$ a ($\mathcal{F}_1$-measurable) random variable is piece-wise constant (constant on $\{1, 2\}$ and $\{3, 4\}$) and for $t = 2$ all functions are ($\mathcal{F}_2$-measurable) random variables.

Course: Theory of Probability I
Term: Fall 2013
Instructor: Gordan Zitkovic

Lecture 10
Conditional Expectation

The definition and existence of conditional expectation

For events A, B with $\mathbb{P}[B] > 0$, we recall the familiar object

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}.$$

We say that $\mathbb{P}[A|B]$ the **conditional probability of** A**, given** B. It is important to note that the condition $\mathbb{P}[B] > 0$ is crucial. When X and Y are random variables defined on the same probability space, we often want to give a meaning to the expression $\mathbb{P}[X \in A|Y = y]$, even though it is usually the case that $\mathbb{P}[Y = y] = 0$. When the random vector (X, Y) admits a joint density $f_{X,Y}(x, y)$, and $f_Y(y) > 0$, the concept of conditional density $f_{X|Y=y}(x) = f_{X,Y}(x, y)/f_Y(y)$ is introduced and the quantity $\mathbb{P}[X \in A|Y = y]$ is given meaning via $\int_A f_{X|Y=y}(x, y)\, dx$. While this procedure works well in the restrictive case of absolutely continuous random vectors, we will see how it is encompassed by a general concept of a conditional expectation. Since probability is simply an expectation of an indicator, and expectations are linear, it will be easier to work with expectations and no generality will be lost.

Two main conceptual leaps here are: 1) we condition with respect to a σ-algebra, and 2) we view the conditional expectation itself as a random variable. Before we illustrate the concept in discrete time, here is the definition.

Definition 10.1. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$, and let $X \in \mathcal{L}^1$ be a random variable. We say that the random variable ξ is (a version of) the **conditional expectation of** X **with respect to** $\mathcal{G}$ - and denote it by $\mathbb{E}[X|\mathcal{G}]$ - if

1. $\xi \in \mathcal{L}^1$.
2. ξ is $\mathcal{G}$-measurable,
3. $\mathbb{E}[\xi \mathbf{1}_A] = \mathbb{E}[X\mathbf{1}_A]$, for all $A \in \mathcal{G}$.

Chapter 2

Models of financial markets on finite probability spaces

We consider a financial market with $1 \leq T \in \mathbb{N}$ periods and $d+1$ financial instruments. More precisely, the modeling framework consists of

- discrete trading times $t = 0, 1, \ldots, T$;
- $d+1$ financial instruments (often a riskless bank account and d risky assets), whose modeling requires a probability space $(\Omega, \mathcal{F}, P)$, a filtration $(\mathcal{F}_t)_{t \in \{0,1,\ldots,T\}}$ and the notion of stochastic processes as introduced in the previous chapter.

2.1 Description of the model

This section is mainly based on [1, Chapter 2].

Adapted stochastic processes are used to model asset price processes. The idea is that $\mathcal{F}_t$ represents the information up to time t and the asset price is measurable with respect to $\mathcal{F}_t$, i.e., its value can be inferred from the knowledge of $\mathcal{F}_t$.

Definition 2.1.1. *A multi-period model of a financial market in discrete time $t \in \{0, 1, \ldots, T\}$, $T \in \mathbb{N}$, consists of an $\mathbb{R}^{d+1}$-valued adapted stochastic process $\widehat{S} = (\widehat{S}^0, \widehat{S}^1, \ldots, \widehat{S}^d)$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$, where*

- *$\widehat{S}^0$ is the so-called numéraire asset used as denomination basis, which is supposed to be strictly positive, i.e. $\widehat{S}^0_t > 0$ for all $t \in \{0, 1, \ldots, T\}$;*
- *$(\widehat{S}^1, \ldots, \widehat{S}^d)$ are $\mathbb{R}^d$-valued adapted stochastic processes for the risky assets.*

The interpretation is as follows: The prices of the assets $0, \ldots, d$ are measured in a fixed money unit, say Euro. The 0^{th} asset plays a special role, it is supposed to be strictly positive and will be used as numéraire. It allows to compare money (Euros) at time 0 to money at time $t > 0$. In many elementary models, $\widehat{S}^0$ is simply the bank account, which is in case of constant interest rates given by $\widehat{S}^0_t = (1+r)^t$.

Definition 2.1.2. • *A trading strategy for the d risky assets $(\widehat{S}^1, \dots, \widehat{S}^d)$ is an $\mathbb{R}^d$-valued predictable process $H_t = (H^1_t, \dots, H^d_t)_{t \in \{1,\dots,T\}}$. The set of all such trading strategies is denoted by $\mathcal{H}$. (In other words $\mathcal{H}$ corresponds to all $\mathbb{R}^d$-valued predictable processes.)*

- *Similarly, a trading strategy for the $d+1$ assets $(\widehat{S}^0, \dots, \widehat{S}^d)$ is an $\mathbb{R}^{d+1}$-valued* predictable *process, which we denote as follows*

$$(\widehat{H}_t)_{t \in \{1,\dots,T\}} = (H^0_t, H^1_t, \dots, H^d_t)_{t \in \{1,\dots,T\}} = (H^0_t, H_t)_{t \in \{1,\dots,T\}}.$$

Remark 2.1.3. The component H^i_t corresponds to the number of shares invested in asset i from period $t-1$ up to t. This means $H^i_t S^i_{t-1}$ is the invested amount at time $t-1$ and $H^i_t S^i_t$ is the resulting wealth at time t. Predictability of $\widehat{H}$ means that an investment can only be made without knowledge of future asset price movements.

Definition 2.1.4. *A trading strategy for the $d+1$ assets $(\widehat{S}^0, \dots, \widehat{S}^d)$ is* self-financing *if for every $t = 1, \dots, T-1$, we have*

$$\widehat{H}_t^\top \widehat{S}_t = \widehat{H}_{t+1}^\top \widehat{S}_t$$

or more explicitly $\sum_{i=0}^d H^i_t \widehat{S}^i_t = \sum_{i=0}^d H^i_{t+1} \widehat{S}^i_t$.

The self-financing condition means that the portfolio is always adjusted in such a way that the current wealth remains the same (one does not remove or add wealth). Accumulated gains or losses are only achieved through changes in the asset prices.

Definition 2.1.5. *The undiscounted wealth process $(\widehat{V}_t)_{\{t \in \{0,1,\dots,T\}\}}$ with respect to a trading strategy $\widehat{H}$ is given by*

$$\begin{aligned} \widehat{V}_0 &= \widehat{H}_1^\top \widehat{S}_0 = \sum_{i=0}^d H^i_1 \widehat{S}^i_0, \\ \widehat{V}_t &= \widehat{H}_t^\top \widehat{S}_t = \sum_{i=0}^d H^i_t \widehat{S}^i_t, \quad t \in \{1, \dots, T\}. \end{aligned} \tag{2.1}$$

The $\mathcal{F}_t$-measurable random variable $\widehat{V}_t$ defined in (2.1) is interpreted as the value of the portfolio at time t defined by the trading strategy $\widehat{H}$.

Remark 2.1.6. Note that if $\widehat{H}$ is self-financing, we have $\widehat{V}_t = \widehat{H}_t^\top \widehat{S}_t = \widehat{H}_{t+1}^\top \widehat{S}_t$.

In the sequel we shall work with discounted price and wealth processes, that means we consider everything in terms of units of the numéraire asset S^0.

Definition 2.1.7. *The* discounted *asset prices are given by*

$$S_t^i := \frac{\widehat{S}_t^i}{\widehat{S}_t^0}, \quad i \in \{1, \dots, d\}, \quad t \in \{0, 1, \dots, T\},$$

and we write $S = (S^1, \dots, S^d)$. The discounted wealth process is given by

$$V_t = \frac{\widehat{V}_t}{\widehat{S}_t^0}, \quad t \in \{0, 1, \dots, T\}.$$

Remark 2.1.8. Note that the discounted numéraire asset $S_t^0 \equiv 1$ for all $t \in \{0, \dots, T\}$.

The self-financing property can be characterized by the following proposition, where we use the notation $\Delta S_u = S_u - S_{u-1}$.

Proposition 2.1.9. *Let $\widehat{S}$ be a model of a financial market as of Definition 2.1.1 and consider an $\mathbb{R}^{d+1}$-valued trading strategy $\widehat{H} = (H^0, H)$ for $\widehat{S}$. Then the following are equivalent:*

1. *$\widehat{H}$ is self-financing.*
2. *The (undiscounted) wealth process satisfies*

$$\widehat{V}_t = \widehat{V}_0 + \sum_{j=1}^{t} \widehat{H}_j^\top \Delta \widehat{S}_j, \quad t = 0, \dots, T.$$

3. *We have*

$$H_t^0 + H_t^\top S_t = H_{t+1}^0 + H_{t+1}^\top S_t, \quad t = 1, \dots, T-1,$$

where S denotes the discounted price process as of Definition 2.1.7.

4. *The discounted wealth process satisfies*

$$V_t = V_0 + \sum_{j=1}^{t} H_j^\top \Delta S_j, \quad t = 0, \dots, T, \tag{2.2}$$

where S denotes the discounted price process as of Definition 2.1.7 and $V_0 = \frac{\widehat{V}_0}{\widehat{S}_0^0} = \frac{\widehat{H}_1^\top \widehat{S}_0}{\widehat{S}_0^0} = H_1^0 + H_1^\top S_0$.

Moreover, there is a bijection between self-financing $\mathbb{R}^{d+1}$-valued trading strategies $\widehat{H} = (H^0, H)$ and pairs (V_0, H), where V_0 is a $\mathcal{F}_0$-measurable random variable and H an $\mathbb{R}^d$-valued trading strategies for the risky assets. Explicitly, $H_t^0 = V_0 + \sum_{u=1}^{t} H_u^\top \Delta S_u - H_t^\top S_t$.

Proof. 1) $\Leftrightarrow$ 2): $\widehat{H}$ is self-financing if and only if

$$\widehat{V}_{j+1} - \widehat{V}_j = \widehat{H}_{j+1}^\top \widehat{S}_{j+1} - \widehat{H}_j^\top \widehat{S}_j = \widehat{H}_{j+1}(\widehat{S}_{j+1} - \widehat{S}_j), \quad j = 0, \dots, T-1$$

which in turn is equivalent to

$$\widehat{V}_t = \widehat{V}_0 + \sum_{j=1}^{t} (\widehat{V}_j - \widehat{V}_{j-1}) = \widehat{V}_0 + \sum_{j=1}^{t} \widehat{H}_j(\widehat{S}_j - \widehat{S}_{j-1}).$$

1) $\Leftrightarrow$ 3) 3) is obtained from 1) by dividing through S_t^0 and conversely 1) is obtained from 3) by multiplying with S_t^0.
3) $\Leftrightarrow$ 4): 3) holds if and only if

$$V_{j+1} - V_j = H_{j+1}^0 + H_{j+1}^\top S_{j+1} - H_j^0 - H_j^\top S_j = H_{j+1}^\top (S_{j+1} - S_j), \quad j = 0, \dots, T-1,$$

which in turn is equivalent to

$$V_t = V_0 + \sum_{j=1}^{t} (V_j - V_{j-1}) = V_0 + \sum_{j=0}^{t-1} H_j^\top (S_j - S_{j-1}).$$

For the last statement let (V_0, H) be given. Since the self-financing property of $\widehat{H}$ is equivalent to (2.2), we can determine H^0 from (V_0, H) via

$$V_0 + \sum_{j=1}^{t} H_j^\top (S_j - S_{j-1}) = V_t = H_t^0 + H_t^\top S_t,$$

where the last equality is simply the definition of the discounted wealth process. Thus

$$H_t^0 = V_0 + \sum_{j=1}^{t} H_j^\top (S_j - S_{j-1}) - H_t^\top S_t = V_0 + \sum_{j=1}^{t-1} H_j^\top (S_j - S_{j-1}) - H_t^\top S_{t-1}$$

which is predictable. Conversely, for a given self-financing $\mathbb{R}^{d+1}$-valued strategy (H^0, H), V_0 is determined via $H_1^0 + H_1^\top S_0$.

□

Definition 2.1.10. *Let $S = (S^1, \dots, S^d)$ be a model of a financial market in discounted terms (as of Definition 2.1.7) and consider an $\mathbb{R}^d$-valued trading strategy $H \in \mathcal{H}$. The discounted* gains process *with respect to H is defined through the* stochastic integral *(in discrete time)*

$$G_t := (H \bullet S)_t := \sum_{j=1}^{t} H_j^\top (S_j - S_{j-1}) =: \sum_{j=1}^{t} H_j^\top \Delta S_j$$

and corresponds to the gains or losses accumulated up to time t in discounted terms.

Remark 2.1.11. Note that by Proposition 2.1.9 the discounted wealth process V of a self-financing strategy is given as the sum of the discounted initial wealth V_0 and the discounted gains process. Moreover due to the second part of 2.1.9, for any $\mathbb{R}^d$-valued trading strategy $H \in \mathcal{H}$ and initial wealth V_0 we can define $V_t := V_0 + (H \bullet S)_t$ which then corresponds to the discounted wealth processes of a self-financing $\mathbb{R}^{d+1}$-valued trading strategy $\widehat{H} = (H^0, H)$ where $H_t^0 = V_0 + \sum_{u=1}^t H_u^\top \Delta S_u - H_t^\top S_t$.

From now on we shall work in terms of the *discounted* $\mathbb{R}^d$*-valued process denoted by* S and *discounted wealth process* V.

2.2 No-arbitrage and the fundamental theorem of asset pricing

This section is mainly based on [1, Chapter 2].

Definition 2.2.1. *Let* $S = (S^1, \dots, S^d)$ *be a model of a financial market in discounted terms.*

- *An* $\mathbb{R}^d$*-valued trading strategy* $H \in \mathcal{H}$ *is called* arbitrage opportunity *if*

$$(H \bullet S)_T \geq 0 \quad P\text{-a.s.} \quad \text{and} \quad P[(H \bullet S)_T > 0] > 0.$$

- *We call a model* arbitrage-free *or satisfies the* no-arbitrage *condition (NA) if there exists no arbitrage strategy.*

Remark 2.2.2. The notion of arbitrage can equivalently be formulated as follows: A self-financing $\mathbb{R}^{d+1}$-valued strategy $\widehat{H}$ is called arbitrage opportunity if the associated wealth process $\widehat{V}$ satisfies $\widehat{V}_0 = 0$ and $\widehat{V}_T \geq 0$ P-a.s and $P[\widehat{V}_T > 0] > 0$.

Assumption 1. *From now on we assume that the probability space* Ω *underlying our model is finite.*

$$\Omega = \{\omega_1, \dots, \omega_N\}$$

for some $N \in \mathbb{N}$ *and a probability measure* P *such that*

$$P[\omega_n] = p_n > 0, \text{ for } n = \{1, \dots, N\}$$

and that $\mathcal{F} = \mathcal{F}_T = \mathcal{P}(\Omega)$.

Recall the notation $L(\Omega, \mathcal{F}, P)$ from (1.1) which denotes in the present case (as $p_n > 0$ for all n) the space of random variables (which are under the above assumption on $\mathcal{F}$ *all* functions from $\Omega \to \mathbb{R}$).

Definition 2.2.3. *A (discounted)* European contingent claim *(derivative/option)* f *is an element of* $L(\Omega, \mathcal{F}, P)$.