STAT GU4265 – Stochastic methods in finance

Johannes Wiesel

February 7, 2022

Contents

Notation

$\mathbb{R}$	the set of real numbers
$\mathbb{R}_+$	the set of non-negative real numbers $[0,\infty)$
$\mathbb{N}$	the set of natural numbers $\{1,2,\ldots\}$
$\mathbb{C}$	the set of complex numbers
$\mathbb{Z}$	the set of integers $\{\ldots,-2,-1,0,1,2,\ldots\}$
$\mathbb{Z}_+$	the set of non-negative integers $\{0,1,2,\ldots\}$
A^c	the complement of a set A, $A^c=\{\omega\in\Omega,\omega\notin A\}$
F_X	the distribution function of a random variable X
p_X	the mass function of a discrete random variable X
f_X	the density function of an absolutely continuous random variable X
ϕ_X	the characteristic function of X
$\mathbb{E}[X]$	the expected value of the random variable X
$\operatorname{Var}(X)$	the variance of X
$\operatorname{Cov}(X,Y)$	the covariance of X and Y
$\mathbb{E}[X \mid B]$	the conditional expectation of X given the event B
$a\wedge b$	$\min\{a,b\}$
$a\vee b$	$\max\{a,b\}$
a^+	$\max\{a,0\}$
$\limsup_{n\to\infty} x_n$	the limit superior of the sequence $x_1,x_2,\ldots$
$\liminf_{n\to\infty} x_n$	the limit inferior of the sequence $x_1,x_2,\ldots$
$a\cdot b$	Euclidean inner (or dot) product in $\mathbb{R}^n$, $\quad a\cdot b=\sum_{i=1}^n a_ib_i$
$\lvert a\rvert$	Euclidean norm in $\mathbb{R}^n$, $\quad \lvert a\rvert=(a\cdot a)^{1/2}$
$X\sim\nu$	the random variable X is distributed as the probability measure ν
$\mathbb{1}_A$	the indicator function of the event A
$N\left(\mu,\sigma^2\right)$	the normal distribution with mean μ and variance σ^2
$N_n(\mu,V)$	the n-dimensional normal distribution with mean μ and variance V
$\operatorname{Bin}(n,p)$	the binomial distribution with parameters n and p
Unif (a,b)	the uniform distribution on the interval (a,b)
L^p	the set of random variables X with $\mathbb{E}[\lvert X\rvert^p]<\infty$

Chapter 1

Crashcourse on probability theory

These notes are a list of many of the definitions and results of probability theory needed to follow the Advanced Financial Models course. Since they are free from any motivating exposition or examples, and since no proofs are given for any of the theorems, these notes should be used only as a reference.

1.1 Measures

Definition 1.1.1. Let Ω be a set. A σ-algebra (speak: sigma-algebra) on Ω is a non-empty set $\mathcal{F}$ of subsets of Ω such that

1. $\Omega \in \mathcal{F}$,
2. if $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$,
3. if $A_1, A_2, \ldots \in \mathcal{F}$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

The terms σ-algebra and sigma-field are interchangeable. The Borel σ-algebra $\mathcal{B}$ on $\mathbb{R}$ is the smallest σ-algebra containing every open interval. More generally, if Ω is a topological space, for instance $\mathbb{R}^n$, the Borel σ-algebra on Ω is the smallest σ-algebra containing every open set.

Definition 1.1.2. Let Ω be a set and let $\mathcal{F}$ be a σ-algebra on Ω. A measure μ on the measurable space $(\Omega, \mathcal{F})$ is a $\mu : \mathcal{F} \to [0, \infty]$ such that

1. $\mu(\emptyset) = 0$,
2. if $A_1, A_2, \ldots \in \mathcal{F}$ are disjoint then $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu\left(A_i\right)$.

Theorem 1.1.1. *There exists a unique measure Leb on $(\mathbb{R}, \mathcal{B})$ such that*

$$\mathrm{Leb}(a, b] = b - a$$

for every $b > a$. This measure is called Lebesgue measure.

Definition 1.1.3. A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a measure such that $\mathbb{P}(\Omega) = 1$. Let Ω be a set, $\mathcal{F}$ a σ-algebra on Ω, and $\mathbb{P}$ a probability measure on $(\Omega, \mathcal{F})$. The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space.
The set Ω is called the sample space, and an element of Ω is called an outcome. A subset of Ω which is an element of $\mathcal{F}$ is called an event.
Let $A \in \mathcal{F}$ be an event. If $\mathbb{P}(A) = 1$ then A is called an almost sure event, and if $\mathbb{P}(A) = 0$ then A is called a null event. The phrase 'almost surely' is often abbreviated a.s. A σ-algebra is called trivial if each of its elements is either almost sure or null.

1.2 Random variables

Definition 1.2.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A random variable is a function $X : \Omega \to \mathbb{R}$ such that the set $\{\omega \in \Omega : X(\omega) \leq t\}$ is an element of $\mathcal{F}$ for all $t \in \mathbb{R}$. Let A be a subset of $\mathbb{R}$, and let X be a random variable. We use the notation $\{X \in A\}$ to denote the set $\{\omega \in \Omega : X(\omega) \in A\}$. For instance, the event $\{X \leq t\}$ denotes $\{\omega \in \Omega : X(\omega) \leq t\}$. The distribution function of X is the function $F_X : \mathbb{R} \to [0, 1]$ defined by

$$F_X(t) = \mathbb{P}(X \leq t)$$

for all $t \in \mathbb{R}$. We also use the term random variable to refer to measurable functions X from Ω to more general spaces. In particular, we call a function $X : \Omega \to \mathbb{R}^n$ a random variable or random vector if $X(\omega) = (X_1(\omega), \ldots, X_n(\omega))$ and X_i is a random variable for each $i \in \{1, \ldots, n\}$.

Definition 1.2.2. Let A be an event in Ω. The indicator function of the event A is the random variable $\mathbb{1}_A : \Omega \to \{0, 1\}$ defined by

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \in A^c \end{cases}$$

for all $\omega \in \Omega$.

1.3 Expectations and variances

Definition 1.3.1. Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. The expected value of X is denoted by $\mathbb{E}[X]$ and is defined as follows

- X is simple, i.e. takes only a finite number of values $x_1, \ldots, x_n$.

$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i \mathbb{P}\left(X = x_i\right).$$

- $X \geq 0$ almost surely.

$$\mathbb{E}[X] = \sup\{\mathbb{E}(Y) : Y \text{ simple and } 0 \leq Y \leq X \text{ a.s. }\}$$

  Note that the expected value of a non-negative random variable may take the value ∞.

- Either $\mathbb{E}\left[X^+\right]$ or $\mathbb{E}\left[X^-\right]$ is finite.

$$\mathbb{E}[X] = \mathbb{E}\left[X^+\right] - \mathbb{E}\left[X^-\right]$$

- X is vector valued and $\mathbb{E}[|X|] < \infty$.

$$\mathbb{E}\left[(X_1, \ldots, X_d)\right] = \left(\mathbb{E}\left[X_1\right], \ldots, \mathbb{E}\left[X_d\right]\right)$$

A random variable X is integrable iff $\mathbb{E}[|X|] < \infty$ and is square-integrable iff $\mathbb{E}\left[X^2\right] < \infty$. The terms expected value, expectation, and mean are interchangeable. The variance of an integrable random variable X, written $\mathrm{Var}(X)$, is

$$\mathrm{Var}(X) = \mathbb{E}\left\{[X - \mathbb{E}(X)]^2\right\} = \mathbb{E}\left[X^2\right] - \mathbb{E}]X]^2.$$

The covariance of square-integrable random variable X and Y, written $\mathrm{Cov}(X, Y)$, is

$$\mathrm{Cov}(X, Y) = \mathbb{E}\{[X - \mathbb{E}[X]][Y - \mathbb{E}[Y]]\} = \mathbb{E}(XY) - \mathbb{E}[X]\mathbb{E}[Y]$$

If neither X or Y is almost surely constant, then their correlation, written $\rho(X, Y)$, is

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(X)^{1/2}\,\mathrm{Var}(Y)^{1/2}}.$$

Random variables X and Y are called uncorrelated if $\mathrm{Cov}(X, Y) = 0$.

Theorem 1.3.1. *Let X and Y be integrable random variables. The following properties hold for the expectation:*

1. *linearity: $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$ for constants a, b.*
2. *positivity: Suppose $X \geq 0$ almost surely. Then $\mathbb{E}[X] \geq 0$ with equality if and only if $X = 0$ almost surely.*

Definition 1.3.2. For $p \geq 1$, the space L^p is the collection of random variables such that $\mathbb{E}\left[|X|^p\right] < \infty$. The space L^∞ is the collection of random variables which are bounded almost surely.

Theorem 1.3.2 (Jensen's inequality)**.** *Let X be a random variable and $g : \mathbb{R} \to \mathbb{R}$ be a convex function. Then*

$$\mathbb{E}[g(X)] \geq g(\mathbb{E}[X])$$

whenever the expectations exist. If g is strictly convex, the above inequality is strict unless X is constant.

Theorem 1.3.3 (Hölder's inequality). *Let X and Y be random variables and let $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$. If $X \in L^p$ and $Y \in L^q$ then*

$$\mathbb{E}[XY] \leq \mathbb{E}\left[|X|^p\right]^{1/p} \mathbb{E}\left[|Y|^q\right]^{1/q}$$

with equality if and only if either $X = 0$ almost surely or X and Y have the same sign and $|Y| = a|X|^{p-1}$ almost surely for some constant $a \geq 0$. The case when $p = q = 2$ is called the Cauchy-Schwarz inequality.

Definition 1.3.3. A random variable X is called discrete if X takes values in a countable set; i.e. there is a countable set S such that $X \in S$ almost surely. If X is discrete, the function $p_X : \mathbb{R} \to [0, 1]$ defined by $p_X(t) = \mathbb{P}(X = t)$ is called the mass function of X.
The random variable X is absolutely continuous (with respect to Lebesgue measure) if and only if there exists a function $f_X : \mathbb{R} \to [0, \infty)$ such that

$$\mathbb{P}(X \leq t) = \int_{-\infty}^{t} f_X(x)dx$$

for all $t \in \mathbb{R}$, in which case the function f_X is called the density function of X. If X is a random vector taking values in $\mathbb{R}^n$, then the density of X, if it exists, is the function $f_X : \mathbb{R}^n \to [0, \infty)$ such that

$$\mathbb{P}(X \in A) = \int_A f_X(x)dx$$

for all Borel subsets $A \subseteq \mathbb{R}^n$.

Theorem 1.3.4. *Let the function $g : \mathbb{R} \to \mathbb{R}$ be such that $g(X)$ is integrable. If X is a discrete random variable with probability mass function p_X taking values in a countable set S then*

$$\mathbb{E}[g(X)] = \sum_{t \in S} g(t)p_X(t).$$

If X is an absolutely continuous random variable with density function f_X then

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx$$

More generally, if X is a random vector valued in $\mathbb{R}^n$ with density f_X and $g : \mathbb{R}^n \to \mathbb{R}$ then

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}^n} g(x)f_X(x)dx.$$

1.4 Special distributions

Definition 1.4.1. Let X be a discrete random variable taking values in $\mathbb{Z}_+$with mass function p_X. The random variable X is called

- **Bernoulli** with parameter p if

$$p_X(0) = 1 - p \text{ and } p_X(1) = p.$$

  where $0 < p < 1$. Then $\mathbb{E}[X] = p$ and $\mathrm{Var}(X) = p(1-p)$.

- **Binomial** with parameters n and p, written $X \sim \mathrm{Bin}(n, p)$, if

$$p_X(k) = \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k} \text{ for all } k \in \{0, 1, \ldots, n\}$$

  where $n \in \mathbb{N}$ and $0 < p < 1$. Then $\mathbb{E}[X] = np$ and $\mathrm{Var}(X) = np(1-p)$.

- **Poisson** with parameter λ if

$$p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda} \text{ for all } k = 0, 1, 2, \ldots$$

  where $\lambda > 0$. Then $\mathbb{E}[X] = \lambda$.

- **geometric** with parameter p if

$$p_X(k) = p(1-p)^{k-1} \text{ for all } k = 1, 2, 3, \ldots$$

  where $0 < p < 1$. Then $\mathbb{E}[X] = 1/p$.

Definition 1.4.2. Let X be a continuous random variable with density function f_X. The random variable X is called

- **uniform** on the interval (a, b), written $X \sim \mathrm{Unif}(a, b)$, if

$$f_X(t) = \frac{1}{b-a} \text{ for all } a < t < b$$

  for some $a < b$. Then $\mathbb{E}[X] = \frac{a+b}{2}$.

- **normal or Gaussian** with mean μ and variance σ^2, written $X \sim N\left(\mu, \sigma^2\right)$, if

$$f_X(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \text{ for all } t \in \mathbb{R}$$

  for some $\mu \in \mathbb{R}$ and $\sigma^2 > 0$. Then $\mathbb{E}[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$.

- **exponential** with rate λ, if

$$f_X(t) = \lambda e^{-\lambda t} \text{ for all } t \geq 0$$

for some $\lambda > 0$. Then $\mathbb{E}[X] = 1/\lambda$.

Definition 1.4.3. If X is a random vector valued in $\mathbb{R}^n$ with density

$$f_X(x) = (2\pi)^{-n/2} \det(V)^{-1/2} \exp\left(-\frac{1}{2}(x-\mu) \cdot V^{-1}(x-\mu)\right)$$

for a positive definite $n \times n$ matrix V and vector $\mu \in \mathbb{R}^n$, then X is said to have the n dimensional normal (or Gaussian) distribution with mean μ and variance V, written $X \sim N_n(\mu, V)$. Then $\mathbb{E}\left[X_i\right] = \mu_i$ and $\operatorname{Cov}\left(X_i, X_j\right) = V_{ij}$.

1.5 Conditional probability and expectation, independence

Definition 1.5.1. Let B be an event with $\mathbb{P}(B) > 0$. The conditional probability of an event A given B, written $\mathbb{P}(A \mid B)$, is

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

The conditional expectation of X given B, written $\mathbb{E}[X \mid B]$, is

$$\mathbb{E}[X \mid B] = \frac{\mathbb{E}\left[X \mathbb{1}_B\right]}{\mathbb{P}(B)}.$$

Theorem 1.5.1 (The law of total probability)**.** *Let $B_1, B_2, \ldots$ be disjoint, non-null events such that $\bigcup_{i=1}^{\infty} B_i = \Omega$. Then*

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}\left(A \mid B_i\right) \mathbb{P}\left(B_i\right)$$

for all events A.

Definition 1.5.2. Let $A_1, A_2, \ldots$ be events. If

$$\mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbb{P}\left(A_i\right)$$

for every finite subset $I \subset \mathbb{N}$ then the events are said to be independent.
Random variables $X_1, X_2, \ldots$ are called independent if the events $\left\{X_1 \leq t_1\right\}, \left\{X_2 \leq t_2\right\}, \ldots$ are independent. The phrase 'independent and identically distributed' is often abbreviated i.i.d.

Theorem 1.5.2. *If X and Y are independent and integrable, then*

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y].$$

1.6 Probability inequalities

Theorem 1.6.1 (Markov's inequality)**.** *Let X be a positive random variable. Then*

$$\mathbb{P}(X \geq \varepsilon) \leq \frac{\mathbb{E}[X]}{\varepsilon}$$

for all $\varepsilon > 0$.

Theorem 1.6.2 (Tschebycheff's inequality)**.** *Let X be a random variable with $\mathbb{E}[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$. Then*

$$\mathbb{P}(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

for all $\varepsilon > 0$.

1.7 Characteristic functions

Definition 1.7.1. The characteristic function of a real-valued random variable X is the function $\phi_X : \mathbb{R} \to \mathbb{C}$ defined by

$$\phi_X(t) = \mathbb{E}\left[e^{itX}\right]$$

for all $t \in \mathbb{R}$, where $i = \sqrt{-1}$. More generally, if X is a random vector valued in $\mathbb{R}^n$ then $\phi_X : \mathbb{R}^n \to \mathbb{C}$ defined by

$$\phi_X(t) = \mathbb{E}\left[e^{it \cdot X}\right]$$

is the characteristic function of X.

Theorem 1.7.1 (Uniqueness of characteristic function)**.** *Let X and Y be real-valued random variables with distribution functions F_X and F_Y. Let ϕ_X and ϕ_Y be the characteristic functions of X and Y. Then*

$$\phi_X(t) = \phi_Y(t) \textit{ for all } t \in \mathbb{R}$$

if and only if

$$F_X(t) = F_Y(t) \textit{ for all } t \in \mathbb{R}$$

1.8 Fundamental probability results

Definition 1.8.1 (Modes of convergence)**.** Let $X_1, X_2, \ldots$ and X be random variables.

- $X_n \to X$ almost surely if $\mathbb{P}\left(X_n \to X\right) = 1$,
- $X_n \to X$ in L^p, for $p \geq 1$, if $\mathbb{E}[|X|^p] < \infty$ and $\mathbb{E}[|X_n - X|^p] \to 0$,
- $X_n \to X$ in probability if $\mathbb{P}\left(|X_n - X| > \varepsilon\right) \to 0$ for all $\varepsilon > 0$,

- $X_n \to X$ in distribution if $F_{X_n}(t) \to F_X(t)$ for all points $t \in \mathbb{R}$ of continuity of F_X.

Theorem 1.8.1. *The following implications hold:*

$$\left.\begin{array}{c} X_n \to X \text{ almost surely} \\ \text{or} \\ X_n \to X \text{ in } L^p, p \geq 1 \end{array}\right\} \Rightarrow X_n \to X \text{ in probability } \Rightarrow X_n \to X \text{ in distribution}$$

Furthermore, if $r \geq p \geq 1$ *then* $X_n \to X$ *in* $L_r \Rightarrow X_n \to X$ *in* L^p.

Definition 1.8.2. Let $A_1, A_2, \ldots$ be events. The term eventually is defined by

$$\{A_n \text{ eventually }\} = \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} A_n$$

and infinitely often by

$$\{A_n \text{ infinitely often }\} = \bigcap_{N\in\mathbb{N}} \bigcup_{n\geq N} A_n.$$

[The phrase 'infinitely often' is often abbreviated $i.o.$]

Theorem 1.8.2 (The first Borel-Cantelli lemma)**.** *Let* $A_1, A_2, \ldots$ *be a sequence of events. If*

$$\sum_{n=1}^{\infty} \mathbb{P}\left(A_n\right) < \infty$$

then $\mathbb{P}\left(A_n \text{ infinitely often }\right) = 0$.

Theorem 1.8.3 (The second Borel-Cantelli lemma)**.** *Let* $A_1, A_2, \ldots$ *be a sequence of independent events. If*

$$\sum_{n=1}^{\infty} \mathbb{P}\left(A_n\right) = \infty$$

then $\mathbb{P}\left(A_n \text{ infinitely often }\right) = 1$.

Theorem 1.8.4 (Monotone convergence theorem)**.** *Let* $X_1, X_2, \ldots$ *be positive random variables with* $X_n \leq X_{n+1}$ *almost surely for all* $n \geq 1$, *and let* $X = \sup_{n\in\mathbb{N}} X_n$. *Then* $X_n \to X$ *almost surely and*

$$\mathbb{E}\left[X_n\right] \to \mathbb{E}[X].$$

Theorem 1.8.5 (Fatou's lemma)**.** *Let* $X_1, X_2, \ldots$ *be positive random variables. Then*

$$\mathbb{E}\left[\liminf_{n\to\infty} X_n\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[X_n\right].$$

Theorem 1.8.6 (Dominated convergence theorem)**.** *Let $X_1, X_2, \ldots$ and X be random variables such that $X_n \to X$ almost surely. If $\mathbb{E}\left[\sup_{n \geq 1} |X_n|\right] < \infty$ then*

$$\mathbb{E}\left[X_n\right] \to \mathbb{E}[X].$$

Theorem 1.8.7 (A strong law of large numbers)**.** *Let $X_1, X_2, \ldots$ be independent and identically distributed integrable random variables with common mean $\mathbb{E}\left[X_i\right] = \mu$. Then*

$$\frac{X_1 + \ldots + X_n}{n} \to \mu \text{ almost surely.}$$

Theorem 1.8.8 (Central limit theorem)**.** *Let $X_1, X_2, \ldots$ be independent and identically distributed with $\mathbb{E}\left[X_i\right] = \mu$ and $\operatorname{Var}\left(X_i\right) = \sigma^2$ for each $i = 1, 2, \ldots$, and let*

$$Z_n = \frac{X_1 + \ldots + X_n - n\mu}{\sigma\sqrt{n}}$$

Then $Z_n \to Z$ in distribution, where $Z \sim N(0, 1)$.

Chapter 2

Introduction: standing assumptions – complications we ignore

Unfortunately, actual financial markets are very complicated. In order to develop a systematic financial theory, we thus concentrate on the essential features of these markets and ignore the less essential complications. Therefore, the theory that will be presented here is concerned with the analysis of market models that have plenty of simplifying assumptions.

That is not to say that these complications are not important. Indeed, there is active ongoing research attempting to remove these simplifying assumptions from the canonical theory. Below is a list of these assumptions.

2.1 Dividends

The total stock of a publicly traded firm is divided into a fixed number N of shares. The owner of each share is then entitled to the fraction $1/N$ of the total profit of the firm. A portion of the firm's profit is usually reinvested by management, for instance by building new factories, but the rest of the profit is paid out to the shareholders. In particular, the owner of each share of stock will receive periodically a dividend payment. However, in this course,

we will usually assume that there are no dividend payments.

Actually, this assumption is not as terrible as it sounds. We will see see shortly how to adapt the theory developed for assets that pay no dividends to incorporate assets that have non-zero dividend payments.

2.2 Tick size

Financial markets usually have a smallest increment of price, the tick (the tick refers back to the days when prices were quoted on ticker tape.) Indeed, the tick size can vary from market to market, and even for assets traded in the same market. However, in this course,

we will assume that the tick size is zero.

This is a convenient assumption for those who prefer continuous mathematics to discrete. It is usually a harmless assumption, unless the prices of interest are very close to zero.

2.3 Transactions costs

Financial transactions are processed by a string of middle men, each of whom charge a fee for their services. Usually the fee is nearly proportional to the size of the transaction. However, in this course,

we will assume that there are no transactions costs.

This assumption is justified by by the fact that transactions costs are often very small relative to the size of typical transactions. But one must always remember that in some applications, it might not be wise to neglect these costs.

2.4 Short-selling constraints

In the real world, it is actually possible for someone to sell an asset that he does not own. The essential mechanism is to borrow a share of that asset from a broker, and then immediately to sell it to the market. This procedure is called short selling.

Brokers, however, place constraints on this behaviour. Indeed, they usually require collateral and charge a fee for their service. Furthermore, if the market price of the asset increases, or if the price of the collateral decreases, the broker may ask the short seller to put up even more collateral. However, in this course,

we will assume that there are no short-selling constraints.

Indeed, the theory of discrete-time trading is cleaner without additional assumptions on the sizes of trades. But we will see that to overcome some technical problems in the theory of continuous-time trading, it will be natural to restrict trading to what are called admissible strategies.

2.5 Divisibility of assets

There is another real-world trading constraint of a rather technical nature. The smallest unit of stock is the share. A share cannot be further divided - it is generally impossible to buy half a share of a particular stock. However, in this course,

we will assume that assets are infinitely divisible.

2.6 Bid-ask spread

Real-world trading is asymmetrical since the price to buy a share is usually higher than the price to sell it. The reason is that are two different ways to buy or sell an asset listed on an exchange: the limit order and the market order.
A limit buy order is an offer to buy a certain number of shares of the asset at a certain price. A limit sell order is defined similarly. The collection of unfilled limit orders is called the limit order book.
At any time, there is the highest price for which there is an order to buy the asset. This is called the bid price. The lowest price for which there is an order to sell is called the ask price. The bid/ask spread is the difference.
A market order are instructions to execute a transaction at the best available price. In particular, if the market order is to buy, then the lowest limit sell order is filled first. Therefore, for small market buy orders, the per share price paid is the ask price. Similarly, if a market sell order arrives, then the highest limit buy order is filled first, and hence the per share price received is the bid price. However, in this course,

we will assume that there are no bid-ask spreads.

This assumption is justified by the observation that in many markets, the spread is very small. However, in times of crisis, this assumption is not usually applicable, and hence the theory breaks down dramatically.

2.7 Market depth

As described above, there are only a finite number of limit orders on the book at one time. If a large market buy order arrives, for instance, then the lowest limit sell order is filled first. But if the market order is bigger than the total shares available to buy at the ask price, then the limit orders at the next-to-lowest price are filled, and progresses up the book until the market order is finally filled. In this way, the ask price increases. The market depth is the number of shares available to buy or sell at the ask or bid price respectively. Equivalently, the depth of a market is a measure of the size of a market order necessary to move quoted prices. However, in this course,

we will assume that there is infinite market depth.

Equivalently, we will assume that investors are small relative to the limit order book, so they are price takers, not price makers. However, the most recent financial crisis shows that this assumption does not always approximate reality - just ask the traders at Lehman Brothers!

2.8 Further modelling complications

Microeconomic models usually involve the interaction of hypothetical agents who are endowed with preferences over some set of economic variables. The observed outcome of the system is

then described by an equilibrium in which the competing preferences of the various agents are balanced through some economic mechanism, such as trade.
We will not deal much with such equilibrium models, but note here that equilibrium models make predictions about the structure of prices in a financial market (after we have made all of the simplifying assumptions listed above). In particular, in an equilibrium model, there cannot be an arbitrage. This will be explained in Chapter 2, but for the sake of this preface, we discuss some of the standard assumptions of equilibrium models and why they may fail in real life.

2.8.1 The expected utility hypothesis

In many standard microeconomic models, agents have preferences over random variables. For instance, suppose that the agent is young now but is planning for retirement. The amount of money that the agent will have in his pension fund when he retires can be modelled as a random variable. Of course, the particular random variable depends on the investment policy the agent chooses now.
The agent much choose their favourite investment strategy. Therefore, we must model their preferences over random variables. The expected utility hypothesis says that the agent prefers the random variable X to the random variable Y if and only if

$$\mathbb{E}[U(X)] > \mathbb{E}[U(Y)]$$

where the function $U : \mathbb{R} \to \mathbb{R}$, called the agent's utility function, models the agents aversion risk.
The expected utility hypothesis seems to have a certain intuitive appeal, and practically, it does make the modelling problem more tractable. Furthermore, von Neumann and Morgenstern showed that the expected utility hypothesis is equivalent to a sensible seeming axiomisation of preferences.

2.8.2 The Allais paradox

The expected utility hypothesis can be tested, and it seems that real human beings do not always behave as though their preferences are consistent with it. Consider two games.

Game A. You must choose between a payment of either X or Y dollars, where

$$X = \begin{cases} 101 & \text{with prob. } 0.33 \\ 100 & \text{with prob. } 0.66 \\ 0 & \text{with prob. } 0.01 \end{cases} \quad \text{and} \quad Y = 100 \ \text{ with prob.1}$$

Game B. Again you must choose between a payment of either X or Y dollars, but now

$$X = \begin{cases} 100 & \text{with prob. } 0.34 \\ 0 & \text{with prob. } 0.66 \end{cases}$$

and

$$Y = \begin{cases} 101 & \text{with prob. } 0.33 \\ 0 & \text{with prob. } 0.67 \end{cases}$$

Apparently, in real experiments, a significant number of people prefer Y in both games. For these people, their preferences are not compatible with the expected utility hypothesis. To see why not, suppose for the sake of finding a contradiction that the agent has a utility function U. Then

$$\text{Game A: } Y \text{ preferred } \Leftrightarrow 0.33U(101) + 0.66U(100) + 0.01U(0) < U(100)$$

$$\text{Game B: } Y \text{ preferred } \Leftrightarrow 0.34U(100) + 0.66U(0) < 0.33U(101) + 0.67U(0)$$

But the above inequalities cannot both hold true!

2.9 The Ellsberg paradox

Underlying the expected utility hypothesis is the assumption that economic agents are perfect statisticians. In reality, of course, when faced with a random outcome, there is risk associated with the realisation of the randomness, but also uncertainty in the unknown distribution of the randomness. Here is an example.

Consider an urn with 30 balls, coloured red, yellow and black. You know that there are 10 red balls in the urn. However, you do not know the number of yellow balls or the number of black balls (but, of course, their sum to 20). A single ball is drawn from the urn. Consider two games:

Game A.

$$X = \begin{cases} 100 & \text{if red} \\ 0 & \text{if yellow or black} \end{cases} \quad \text{and} \quad Y = \begin{cases} 100 & \text{if yellow} \\ 0 & \text{if red or black} \end{cases}$$

Game B.

$$X = \begin{cases} 100 & \text{if red or black} \\ 0 & \text{if yellow} \end{cases} \quad \text{and} \quad Y = \begin{cases} 100 & \text{if yellow or black} \\ 0 & \text{if red} \end{cases}$$

In experiments, people tend to prefer X in game A and Y in game B. This also contradicts the expected utility hypothesis. Suppose the agent estimates the probability of drawing yellow as p where $0 < p < 2/3$.

$$\text{Game A: } X \text{ preferred } \Leftrightarrow \frac{1}{3}U(100) + \frac{2}{3}U(0) > pU(100) + (1-p)U(0)$$

$$\text{Game B: } Y \text{ preferred } \Leftrightarrow (1-p)U(100) + pU(0) < \frac{2}{3}U(100) + \frac{1}{3}U(0),$$

a contradiction! One way to avoid this paradox is to let the agent have preferences not only of risk but also uncertainty. For example, suppose that there is a set of probability models $\mathcal{P}$ consistent with the agent's beliefs. An agent who is very averse uncertainty would prefer X to Y if and only if

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[U(X)] > \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[U(Y)].$$

This point of view is pursued in an area called robust finance, but we will not discuss it in these notes.

2.10 Prerequisite knowledge

The emphasis of this course is on some of the mathematical aspects of financial market models. Very little is assumed of the reader's knowledge of the workings of financial markets. However, some mathematical background is needed.

Our starting point is the famous observation (sometimes attributed to Niels Bohr) that it is difficult to make predictions, especially about the future. Indeed, anyone with even a passing acquaintance with finance knows that most of us cannot predict with absolute certainty how the the price of an asset will fluctuate - otherwise we would be much richer! Therefore, the proper language to formulate the models that we will study is the language of probability theory. An attempt is made to keep this course self-contained, but you should be familiar with the basics of the theory, including knowing the definition and key properties of the following concepts: random variable, expected value, variance, conditional probability/expectation, independence, Gaussian (normal) distribution, etc. Familiarity with measure theoretical probability is helpful, though a crashcourse on probability theory is given in Chapter 1.

Chapter 3

Discrete-time models

We consider a market with n assets. The identity of the assets is not important as long as the standing assumptions (zero dividends, zero tick size, zero transaction costs, no short-selling constraints, infinite divisibility, zero bid-ask spread, infinite market depth) are fulfilled. We usually think of the assets as being stocks and bonds, but they also can be more exotic things like pork belly futures.

We will use the notation P_t^i for the price of asset i at time t. In this section, the time index set is the non-negative integers, so the notation $t \geq 0$ should be interpreted as $t \in \{0, 1, 2, \ldots\}$.

3.1 Measurability and conditional expectations

A modelling assumption that we will use throughout is that the collection of prices

$$P = \left(P_t^1, \ldots, P_t^n\right)_{t \geq 0}$$

is an n-dimensional stochastic process adapted to a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$. We now briefly describe what this means.

A stochastic process $(Z_t)_{t \geq 0}$ is just a collection of random variables (or vectors) indexed by the parameter t. In our case, the parameter is interpreted as time, either discrete or continuous. (Recall that when we speak of a random variable Z, we secretly have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ in the background such that the map $Z : \Omega \to \mathbb{R}$ is $\mathcal{F}$-measurable. See the crashcourse if this sounds unfamiliar to you.)

We now formalise the concept of information being revealed as time marches forward. The correct notions are that of a filtration and adaptedness.

Motivation: You are probably already familiar with the notion of the measurability of a set. Measurability is a hugely important (though technical) idea in the theory of Lebesgue integration: for instance, Vitali showed that it is impossible to define the Lebesgue measure of every subset of $\mathbb{R}$.

But on top of its technical importance, measurability has is a way to model information. Give an probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let $\mathcal{G}$ be some set of information. For the moment, we will be vague about what $\mathcal{G}$ is, but it intuitively should have the property that an event $A \in \mathcal{F}$ is $\mathcal{G}$-measurable iff

$$\mathbb{P}(A \mid \mathcal{G}) \in \{0, 1\}.$$

Note we have not yet defined the notation $\mathbb{P}(A \mid \mathcal{G})$, but it should thought of as the probability of the event A given knowledge of the information set $\mathcal{G}$.
For instance, consider the experiment of tossing a fair coin two times. We can model this experiment on the sample space $\Omega = \{HH, HT, TH, TT\}$. The set of all events is the set

$$\mathcal{F} = \{\emptyset, \{HH\}, \ldots, \{HH, HT\}, \ldots, \{HH, HT, TH\}, \ldots, \{HH, HT, TH, TT\}\}$$

of all $2^4 = 16$ subsets of Ω. The probability measure is just the one that assigns $\mathbb{P}(\{\omega\}) = 1/4$ equal probability to each elementary event. Suppose $\mathcal{G}$ is information revealed by the first toss of the coin. The set event

$$A = \{ \text{ first toss is heads } \} = \{HH, HT\}$$

is $\mathcal{G}$-measurable, since for any sensible definition of the conditional probability we must have

$$\mathbb{P}(A \mid \mathcal{G}) = \begin{cases} 1 & \text{if the first toss is heads} \\ 0 & \text{if the first toss is tails.} \end{cases}$$

On the other hand, the event

$$B = \{ \text{ both tosses are heads } \} = \{HH\}$$

is not $\mathcal{G}$-measurable. Indeed, we must have

$$\mathbb{P}(B \mid \mathcal{G}) = \left\{ \begin{array}{cl} 1/2 & \text{if the first toss is heads} \\ 0 & \text{if the first toss is tails.} \end{array} \right.$$

Now returning to the general case, rather than modelling the set $\mathcal{G}$ of information as a new mathematical structure, we simply identify $\mathcal{G}$ with the collection of all $\mathcal{G}$-measurable events. Notice that, assuming that the conditional probability $\mathbb{P}(\cdot \mid \mathcal{G})$ somehow behaves like an unconditional probability, then $\mathcal{G}$ is a σ-algebra. The lesson of all of this is that it makes sense to model information as a sub-σ-algebra of the σ-algebra of all events $\mathcal{F}$. It remains to properly define the conditional probability and then check that it has the correct properties. We briefly recall some notions from probability.

Definition 3.1.1. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra of events. A random variable $X : \Omega \to \mathbb{R}$ is measurable with respect to $\mathcal{G}$ (or briefly, $\mathcal{G}$ measurable) if and only if the event $\{X \leq x\}$ is an element of $\mathcal{G}$ for all $x \in \mathbb{R}$.

You know what that the conditional expectation of an integrable random variable X given a non-null event G means

$$\mathbb{E}[X \mid G] = \frac{\mathbb{E}\left[X \mathbb{1}_G\right]}{\mathbb{P}(G)}$$

The next theorem leads to a definition of conditional expectation given a σ-algebra:

Theorem 3.1.1 (Existence and uniqueness of conditional expectations)**.** *Let X be an integrable random variable defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra of $\mathcal{F}$. Then there exists an integrable $\mathcal{G}$-measurable random variable Y such that*

$$\mathbb{E}\left[\mathbb{1}_G Y\right] = \mathbb{E}\left[\mathbb{1}_G X\right]$$

for all $G \in \mathcal{G}$. Furthermore, if there exists another $\mathcal{G}$-measurable random variable Y' such that $\mathbb{E}\left[\mathbb{1}_G Y'\right] = \mathbb{E}\left[\mathbb{1}_G X\right]$ for all $G \in \mathcal{G}$, then $Y = Y'$ almost surely.

Definition 3.1.2. Let X be an integrable random variable and let $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra. The conditional expectation of X given $\mathcal{G}$, written $\mathbb{E}[X \mid \mathcal{G}]$, is a $\mathcal{G}$-measurable random variable with the property that

$$\mathbb{E}\left[\mathbb{1}_G \mathbb{E}[X \mid \mathcal{G}]\right] = \mathbb{E}\left[\mathbb{1}_G X\right]$$

Example 3.1.1 (σ-algebra generated by a countable partition)**.** Let X be a non-negative random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $G_1, G_2, \ldots$ be a sequence of disjoint events with $\mathbb{P}\left(G_n\right) > 0$ for all n and $\bigcup_{n \in \mathbb{N}} G_n = \Omega$.

Let $\mathcal{G}$ be the smallest σ-algebra containing $\{G_1, G_2, \ldots, \ldots\}$. That is, every element of $\mathcal{G}$ is of the form $\bigcup_{n \in I} G_n$ where $I \subseteq \mathbb{N}$. Then

$$\mathbb{E}[X \mid \mathcal{G}](\omega) = \mathbb{E}\left[X \mid G_n\right] = \frac{\mathbb{E}\left[X \mathbb{1}_{G_n}\right]}{\mathbb{P}\left(G_n\right)} \qquad \text{if } \omega \in G_n$$

where the right-hand side denotes conditional expectation given the event G_n. More concretely, suppose $\Omega = \{HH, HT, TH, TT\}$ consists of two tosses of a fair coin, and let

$$\mathcal{G} = \{\emptyset, \{HH, HT\}, \{TH, TT\}, \Omega\}$$

be the σ-algebra containing the information revealed by the first toss. Consider the random variable

$$X(\omega) = \begin{cases} a & \text{if } \omega = HH \\ b & \text{if } \omega = HT \\ c & \text{if } \omega = TH \\ d & \text{if } \omega = TT \end{cases}.$$

Then

$$\mathbb{E}[X \mid \mathcal{G}](\omega) = \begin{cases} (a+b)/2 & \text{if } \omega \in \{HH, HT\} \\ (c+d)/2 & \text{if } \omega \in \{TH, TT\} \end{cases}$$

The important properties of conditional expectations are collected below:

Theorem 3.1.2. *Let all random variables appearing below be such that the relevant conditional expectations are defined, and let $\mathcal{G}$ be a sub-σ-algebra of the σ-algebra $\mathcal{F}$ of all events.*

- *linearity:* $\mathbb{E}[aX + bY \mid \mathcal{G}] = a\mathbb{E}[X \mid \mathcal{G}] + b\mathbb{E}[Y \mid \mathcal{G}]$ *for all constants a and b*
- *positivity: If $X \geq 0$ almost surely, then $\mathbb{E}[X \mid \mathcal{G}] \geq 0$ almost surely.*
- *Jensen's inequality: If f is convex, then $\mathbb{E}[f(X) \mid \mathcal{G}] \geq f[\mathbb{E}(X \mid \mathcal{G})]$*
- *monotone convergence theorem: If $0 \leq X_n \uparrow X$ a.s. then $\mathbb{E}\left[X_n \mid \mathcal{G}\right] \uparrow \mathbb{E}[X \mid \mathcal{G}]$ a.s.*
- *Fatou's lemma: If $X_n \geq 0$ a.s. for all n, then $\mathbb{E}\left[\liminf_n X_n \mid \mathcal{G}\right] \leq \liminf_n \mathbb{E}\left[X_n \mid \mathcal{G}\right]$*
- *dominated convergence theorem: If $\sup_n |X_n|$ is integrable and $X_n \to X$ a.s. then $\mathbb{E}\left[X_n \mid \mathcal{G}\right] \to \mathbb{E}[X \mid \mathcal{G}]$ a.s.*
- *If X is independent of $\mathcal{G}$ (the events $\{X \leq x\}$ and G are independent for each $x \in \mathbb{R}$ and $G \in \mathcal{G}$) then $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$. In particular, $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$ if $\mathcal{G}$ is trivial.*
- *"take out what's known property": If X is $\mathcal{G}$-measurable, then $\mathbb{E}[XY \mid \mathcal{G}] = X\mathbb{E}[Y \mid \mathcal{G}]$. In particular, if X is $\mathcal{G}$-measurable, then $\mathbb{E}[X \mid \mathcal{G}] = X$.*
- *tower property or law of iterated expectations: If $\mathcal{H} \subseteq \mathcal{G}$ then*

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{H}] \mid \mathcal{G}] = \mathbb{E}[X \mid \mathcal{H}].$$

Definition 3.1.3. The conditional probability of an event $A \in \mathcal{F}$ given a sub-σ-algebra $\mathcal{G}$ is defined by

$$\mathbb{P}(A \mid \mathcal{G}) = \mathbb{E}\left[\mathbb{1}_A \mid \mathcal{G}\right].$$

We now come full circle to show that the motivation for defining measurability is compatible with the definitions we have chosen:

Proposition 3.1.1. *If $\mathbb{P}(A \mid \mathcal{G}) \in \{0,1\}$ almost surely, then there exists a $\mathcal{G}$-measurable event A' such that $\mathbb{P}\left(A \backslash A'\right) = 0 = \mathbb{P}\left(A' \backslash A\right)$.*

Proof. Since the conditional probability takes values in $\{0,1\}$ there exists a $\mathcal{G}$-measurable event A' such that

$$\mathbb{P}(A \mid \mathcal{G}) = \mathbb{1}_{A'}$$

Note that

$$\begin{aligned}
\mathbb{P}(A) &= \mathbb{E}\left[\mathbb{1}_A\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_A \mid \mathcal{G}\right]\right] \\
&= \mathbb{E}\left[\mathbb{1}_{A'}\right] \\
&= \mathbb{P}\left(A'\right)
\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}(A \cap A') &= \mathbb{E}[\mathbb{1}_A \mathbb{1}_{A'}] \\ &= \mathbb{E}[\mathbb{E}[\mathbb{1}_A \mathbb{1}_{A'} \mid \mathcal{G}]] \\ &= \mathbb{E}[\mathbb{1}_{A'} \mathbb{E}[\mathbb{1}_A \mid \mathcal{G}]] \\ &= \mathbb{E}\left[\mathbb{1}_{A'}^2\right] \\ &= \mathbb{P}(A')\end{aligned}$$

and hence

$$\mathbb{E}\left[(\mathbb{1}_A - \mathbb{1}_{A'})^2\right] = \mathbb{P}(A) + \mathbb{P}\left(A'\right) - 2\mathbb{P}\left(A \cap A'\right) = 0.$$

□

Continuing with the theme of measurability, we introduce a few more terms:

Definition 3.1.4. A filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a collection of σ-algebras such that $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for all $0 \leq s \leq t$.

Definition 3.1.5. A process $X = (X_t)_{t\geq 0}$ is adapted to $\mathbb{F}$ iff the random variable X_t is $\mathcal{F}_t$–measurable for all $t \geq 0$.

When discussing an adapted stochastic process but a filtration is not explicitly mentioned, then we are implicitly working with the natural filtration of the process.

Definition 3.1.6. Given a stochastic process $X = (X_t)_{t\geq 0}$, the natural filtration of X, (or the filtration generated by X) is the smallest filtration for which X is adapted. That is, it is the filtration $(\mathcal{F}_t)_{t\geq 0}$ where

$$\mathcal{F}_t = \sigma\left(X_s, 0 \leq s \leq t\right)$$

To gain some intuition about these definitions, consider this example.

Example 3.1.2. Return to the experiment of tossing a fair coin two times. The flow of information is modelled by the following σ-algebras

- $\mathcal{F}_0 = \{\emptyset, \Omega\}$,
- $\mathcal{F}_1 = \{\emptyset, \{HH, HT\}, \{TH, TT\}, \Omega\}$,
- $\mathcal{F}_2 = \mathcal{F}$.

Now consider a stochastic process $(X_t)_{t\in\{0,1,2\}}$ that is adapted to the filtration $(\mathcal{F}_t)_{t\in\{0,1,2\}}$. Intuitively, the value of the random variable X_t is known once after t tosses of the coin. For instance, X_0 must be a constant,

$$X_0(\omega) = a \text{ for all } \omega \in \Omega,$$

since there is no information before the experiment. On the other hand, the random variable X_1 must be of the form

$$X_1(\omega) = \begin{cases} b & \text{if } \omega \in \{HH, HT\} \\ c & \text{if } \omega \in \{TH, TT\} \end{cases}$$

since the only information known at time 1 is whether or not the first coin came up heads. Finally, X_2 can be any function on Ω, that is, of the form

$$X_2(\omega) = \begin{cases} d & \text{if } \omega = HH \\ e & \text{if } \omega = HT \\ f & \text{if } \omega = TH \\ g & \text{if } \omega = TT \end{cases}$$

Alternatively, on this particular filtered probability space, the adapted process X can be visualised by the tree diagram: Notice that for all $t \in \{0, 1, 2\}$ the event $\{X_t \leq x\}$ is in $\mathcal{F}_t$ for every real x.

For this course, it will be convenient to assume that there is no randomness at time 0 . This can be made formal by assuming

the σ-algebra $\mathcal{F}_0$ is trivial.

This means that if A is an element $\mathcal{F}_0$ then either $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$. In particular, every $\mathcal{F}_0$-measurable random variable is almost surely constant. In the discrete-time theory, there nothing lost by further assuming $\mathcal{F}_0 = \{\emptyset, \Omega\}$. However, it turns out that this further assumption is technically inconvenient in the continuous-time theory. Before continuing to the financial models, we list one final definition in this section.

Definition 3.1.7. A discrete-time process $X = (X_t)_{t \geq 1}$ is previsible (or predictable) iff the random variable X_t is $\mathcal{F}_{t-1}$-measurable for all $t \geq 1$.

Remark 3.1.3. Note that the time index set for a previsible process $(X_t)_{t \geq 1}$ is (usually) $\{1, 2, \ldots\}$, not $\{0, 1, \ldots\}$. In particular, X_0 is not necessarily defined.

Remark 3.1.4. In discrete time, a process X is previsible if and only if the process Y is adapted, where $X_t = Y_{t-1}$. That is to say, the notion of previsibility can be dispensed with by simply changing notation. However, in continuous time, there is a more subtle difference between the notions of previsibility and adaptedness. Therefore, for the sake of a unified treatment of the discrete and continuous time cases, we keep it in.

3.2 The set-up

Returning to our financial modelling, we assume that the market prices are given by a n-dimensional adapted process $P = (P_t)_{t \geq 0}$.

To the market described by the adapted process P, we now introduce an investor. At time t, the investor comes into the market with an initial amount of money X_t. The investor receives some income I_t and consumes some amount C_t.

With the remaining money $X_t + I_t - C_t$, the investor buys a portfolio $H_{t+1} \in \mathbb{R}^n$ of the assets, where the real number H^i_{t+1} denotes the number of shares held in asset i. (If $H^i_{t+1} > 0$ then the position is said to be long, and if $H^i_{t+1} < 0$ then the position is said to be short.) There are two important accounting relationships between the variables. First, the budget constraint is

$$X_t + I_t - C_t = H_{t+1} \cdot P_t,$$

where we are using the notation

$$a \cdot b = \sum_{i=1}^{n} a^i b^i$$

for the usual Euclidean inner (or dot) product in $\mathbb{R}^n$. Second, at time $t+1$, the investor's portfolio is worth

$$X_{t+1} = H_{t+1} \cdot P_{t+1},$$

which says that the amount of money the investor has in the market at time $t+1$ is equal to the liquidation value of the assets.

We will model the price process P and external income process I as exogenously given adapted processes. We consider the investor's initial wealth $X_0 = x$ as a given constraint, and the investor's portfolio process $(H_t)_{t \geq 1}$ as their control. In order to eliminate clairvoyant investors, we insist that the control H is previsible. Note that given x and H, the investor's wealth is

$$X_t^{x,H} = \begin{cases} x & \text{if } t = 0 \\ H_t \cdot P_t & \text{if } t \geq 1 \end{cases}$$

and the consumption is then

$$C_t^{x,H} = X_t^{x,H} + I_t - H_{t+1} \cdot P_t.$$

Now that we have our market model and we've introduced an investor into this market, our first challenge is to find out how to invest optimally. We consider one such optimal investment problem. The main motivation for studying this problem is to introduce the very important notion of a martingale deflator.

Let $T > 0$ be some non-random time horizon and prefers a consumption stream $c = (c_t)_{0 \leq t \leq T}$ to $c' = (c'_t)_{0 \leq t \leq T}$ iff and only if

$$\mathbb{E}[U(c)] > \mathbb{E}\left[U\left(c'\right)\right]$$

where $U : \mathbb{R}^{1+T} \to \mathbb{R} \cup \{-\infty\}$ is a given utility function. We will assume that $c_t \mapsto U(c_0, \ldots, c_T)$ is strictly increasing for each t, modelling the assumption that the investor strictly prefers more to less. (Usually we also assume that U is strictly concave, so that the investor is risk-averse, strictly preferring to consume the non-random quantity $\mathbb{E}(c)$ to the random quantity c, for any non-constant random vector c.)

We suppose that investor's initial wealth is x given. We also suppose that they will live exactly to age T, and since they derive no utility from wealth in the afterlife, choose to consume his remaining wealth at time T. Summing up, the investor faces the problem

$$\text{maximise } \mathbb{E}[U(c)] \quad \text{subject to } \begin{cases} c_t = C_t^{x,H} & \text{for } 0 \le t \le T \\ c_T = H_T \cdot P_T + I_T \end{cases}$$

With this problem in mind, we introduce an important definition:

Definition 3.2.1. An arbitrage is an n-dimensional previsible process H such that there exists a non-random time $T > 0$ with the properties, that the consumption stream

$$\begin{aligned} c_0 &= -H_1 \cdot P_0 \\ c_t &= H_t \cdot P_t - H_{t+1} \cdot P_t, \text{ for } 1 \le t \le T-1, \\ c_T &= H_T \cdot P_T \end{aligned}$$

satisfies

- $c_t \ge 0$ almost surely for all $0 \le t \le T$,
- $\mathbb{P}\left(c_t > 0 \text{ for some } 0 \le t \le T\right) > 0$.

Note that if H^{f} is a feasible investment strategy for the above investment problem and if H^{a} is an arbitrage, then $H^{\mathrm{f}} + H^{\mathrm{a}}$ is also feasible (since it can be funded with same initial wealth x). However, the new strategy has strictly higher expected utility

$$\mathbb{E}\left[U\left(c^{\mathrm{f}} + c^{\mathrm{a}}\right)\right] > \mathbb{E}\left[U\left(c^{\mathrm{f}}\right)\right].$$

Inductively, the strategy $H^{\mathrm{f}} + kH^{\mathrm{a}}$ is feasible for every $k \ge 0$. In particular, if there is an arbitrage then there cannot be an optimal investment strategy to the utility maximisation problem.

Remark 3.2.1. And why do we care about the existence of optimal investment strategies? To explain, we take a moment to ask where do prices come from? We consider a one-period model. We model P_1 as a random variable, as it is not known at time 0. What determines the randomness? One could argue that all that matters is the beliefs of the market participants, not the underlying mechanism that causes the apparent randomness. So, we assume that there are J investors, and each investor j has a probability measure $\mathbb{P}_j$, where $j = 1, \ldots, J$ modelling the distribution of P_1. We also assume that each agent j comes to the market with initial capital x_j. The market already has the n assets, with total supply of asset i given by S^i and $S = \left(S^1, \ldots, S^n\right)$. The agents trade with each other until each arrive at an optimal allocation H_j^* and collectively determine an initial price P_0^*.

To formalise this, we have the following definition:

Definition 3.2.2. Given initial wealth x_j, utility functions U_j and probability measures $\mathbb{P}_j$, for $j = 1, \dots J$, which determine the distribution of the random vector P_1, let

$$H_j(p) = \arg\max_{H \in \mathbb{R}^d} \{\mathbb{E}_j [U(c_0, c_1)] : \quad c_0 = x_j - H \cdot p, \ c_1 = H \cdot P_1\}$$

be the optimal portfolio for agent i assuming the initial price is $P_0 = p$. An equilibrium price P_0^* is a solution to the equation

$$\sum_{j=1}^{J} H_j (P_0^*) = S$$

where the notation $\mathbb{E}_j$ denotes expectation with respect to $\mathbb{P}_j$.

Note that the above condition says that for the equilibrium initial price P_0^*, the agents portfolios H_j^* solve their version of the optimal investment problem $(*)$. A consequence is the following motivating result:

Proposition 3.2.1. *If the market is in equilibrium then no agent can believe there is an arbitrage.*

Before proceeding, we consider a simple consequence of the assumption of no arbitrage in a market model.

Proposition 3.2.2. *Consider a market with $n = 1$ asset. If there exists a non-random $T > 0$ such that $P_T \geq 0$ almost surely, then $P_t \geq 0$ almost surely for all $0 \leq t \leq T$.*

Proof. Let $\tau = \inf\{t \geq 0 : P_t < 0\}$ and let

$$H_t = \mathbb{1}_{\{\tau < t \leq T\}}$$

Note that H is previsible since for $1 \leq t \leq T$ we have

$$\{\tau < t\} = \cup_{s=0}^{t-1} \{P_s < 0\} \in \mathcal{F}_{t-1}.$$

Hence $c_t = -P_t > 0$ on $\{t = \tau < T\}$ and $c_T = H_T \cdot P_T \geq 0$. Since there is no arbitrage, we must have $\mathbb{P}(\tau < T) = 0$, or equivalently, $\mathbb{P}(P_s \geq 0 \text{ for all } 0 \leq s \leq T-1) = 1$ as claimed. □

3.3 The first fundamental theorem and martingales

We have tried to argue above that it is natural to insist that our market model is free of arbitrage strategies. But how can we check that a given price process P is arbitrage free? The answer is contained in a famous theorem:

Theorem 3.3.1 (First fundamental theorem of asset pricing, 1st FTAP)**.** *A market model has no arbitrage if and only if there exists a martingale deflator.*

We spend the next few pages unpacking this theorem. First, we need a definition to get started:

Definition 3.3.1. A martingale deflator is an adapted (real-valued) process Y such that $Y_t > 0$ for all $t \geq 0$ almost surely, and such that the n-dimensional process

$$M_t = P_t Y_t$$

is a martingale.

Now we come to one of the most important concepts in financial mathematics, the martingale. A martingale is simply an adapted stochastic process that is constant on average in the following sense:

Definition 3.3.2. A martingale relative to a filtration $\mathbb{F}$ is an adapted stochastic process $M = (M_t)_{t\geq 0}$ with the following properties:

- $\mathbb{E}\left[|M_t|\right] < \infty$ for all $t \geq 0$,
- $\mathbb{E}\left[M_t \mid \mathcal{F}_s\right] = M_s$ for all $0 \leq s \leq t$.

Remark 3.3.2. The above definition of martingale is the same both discrete- and continuous-time processes. However, if the time index set is discrete $\mathbb{T} = \mathbb{Z}_+$, it is an exercise to show that an integrable process M is a martingale only if $\mathbb{E}\left[M_{t+1} \mid \mathcal{F}_t\right] = M_t$ for all $t \geq 0$. That is, it is sufficient to verify the conditional expectations of the process one period ahead.

Below are some examples of martingales.

Example 3.3.1. Let $\xi_1, \xi_2, \xi_3, \ldots$ be independent integrable random variables such that $\mathbb{E}\left[\xi_i\right] = 0$ for all i. The process $(S_t)_{t\geq 0}$ given by $S_0 = 0$ and

$$S_t = \xi_1 + \ldots + \xi_t$$

is a martingale relative to its natural filtration. Indeed, the random variable S_t is integrable since

$$\mathbb{E}\left[|S_t|\right] \leq \mathbb{E}\left[|\xi_1|\right] + \ldots + \mathbb{E}\left[|\xi_t|\right]$$

by the triangular inequality and all the terms in this finite sum are finite by assumption. Also,

$$\begin{aligned}
\mathbb{E}\left[S_{t+1} \mid \mathcal{F}_t\right] &= \mathbb{E}\left[S_t + \xi_{t+1} \mid \mathcal{F}_t\right] \\
&= \mathbb{E}\left[S_t \mid \mathcal{F}_t\right] + \mathbb{E}\left[\xi_{t+1} \mid \mathcal{F}_t\right] \\
&= S_t + \mathbb{E}\left[\xi_{t+1}\right] = S_t
\end{aligned}$$

where the conditional expectation $\mathbb{E}\left[\xi_{t+1} \mid \mathcal{F}_t\right]$ is replaced by the unconditional expectation $\mathbb{E}\left[\xi_{t+1}\right]$ by the assumption that ξ_{t+1} is independent of $\mathcal{F}_t = \sigma\left(S_1, \ldots, S_t\right) = \sigma\left(\xi_1, \ldots, \xi_t\right)$.

Example 3.3.2. We now construct one of the most important examples of a martingale. Let X be an integrable random variable, and let

$$M_t = \mathbb{E}\left[X \mid \mathcal{F}_t\right].$$

Then $M = (M_t)_{t\geq 0}$ is a martingale. Integrability follows from the definition of conditional expectation. Now, for every $0 \leq s \leq t$ we have

$$\begin{aligned} \mathbb{E}\left[M_t \mid \mathcal{F}_s\right] &= \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{F}_t\right] \mid \mathcal{F}_s\right] \\ &= \mathbb{E}\left[X \mid \mathcal{F}_s\right] = M_s \end{aligned}$$

by the tower property. Notice that this example also works in continuous time.

Sometimes we are given a process $(M_t)_{0\leq t\leq T}$ where $T > 0$ is a fixed, non-random time horizon. To check that this process is a martingale, we need only check that

$$M_t = \mathbb{E}\left[M_T \mid \mathcal{F}_t\right] \text{ for all } 0 \leq t \leq T,$$

because this corresponds to the construction above with $X = M_T$.
This last example is theorem shows how to take one martingale and build another one.

Proposition 3.3.1. *Let M be a martingale and let K be a bounded predictable process. Then the process X defined by*

$$X_t = \sum_{s=1}^{t} K_s\left(M_s - M_{s-1}\right)$$

is a martingale.

Proof. First, note X_t is $\mathcal{F}_t$-measurable by construction. Also, by assumption, we have $\mathbb{E}\left[|M_t|\right] < \infty$ for all t since M is a martingale and that there exist a constant $C > 0$ such that $|K_t| \leq C$ almost surely for all $t \geq 0$. Hence

$$\begin{aligned} \mathbb{E}\left[|X_t|\right] &\leq \sum_{s=1}^{t} \mathbb{E}\left[|K_s|\,|M_s - M_{s-1}|\right] \\ &\leq \sum_{s=1}^{t} C\left[\mathbb{E}\left[|M_s|\right] + \mathbb{E}\left[|M_{s-1}|\right]\right] < \infty \end{aligned}$$

Using the predictability of K and the slot property of conditional expectation, we have

$$\begin{aligned} \mathbb{E}\left[X_{t+1} - X_t \mid \mathcal{F}_t\right] &= \mathbb{E}\left[K_{t+1}\left(M_{t+1} - M_t\right) \mid \mathcal{F}_t\right] \\ &= K_{t+1}\mathbb{E}\left[M_{t+1} - M_t \mid \mathcal{F}_t\right] \\ &= 0 \end{aligned}$$

implying the martingale property $\mathbb{E}\left[X_{t+1} \mid \mathcal{F}_t\right] = X_t$, since X_t is $\mathcal{F}_t$-measurable. □

Remark 3.3.3. The martingale X above is often called a martingale transform or a discrete time stochastic integral. As we will see, it is one of the key building blocks for the continuous time theory to come.

3.4 Local martingales

It turns out that to prove the 1st FTAP, even in the easy direction, we need a little more technology. With that introduction, we begin our study of local martingales. First we start with a definition.

Definition 3.4.1. A stopping time for a filtration $(\mathcal{F}_t)_{t\in\mathbb{T}}$ is a random variable τ taking values in $\mathbb{T}\cup\{\infty\}$ such that the event $\{\tau\le t\}$ is $\mathcal{F}_t$-measurable for all $t\in\mathbb{T}$.

Example 3.4.1. Obviously, non-random times are stopping times. That is, if $\tau = t_0$ for some fixed $t_0\ge 0$, then $\{\tau\le t\}=\Omega$ if $t_0\le t$ and $\emptyset$ otherwise.

Example 3.4.2. Here is a typical example of a stopping time. Let $(Y_t)_{t\ge 0}$ be a discrete-time adapted process and let A be a Borel set (for instance, an interval). Then the random variable

$$\tau = \inf\{t\ge 0 : Y_t\in A\}$$

(with the usual convention that $\inf\emptyset = +\infty$) corresponding to the first time the process enters the set A is a stopping time. Indeed,

$$\{\tau\le t\} = \bigcup_{s=0}^{t}\{Y_s\in A\}$$

is $\mathcal{F}_t$-measurable because each $\{Y_s\in A\}$ is $\mathcal{F}_s$-measurable by the adaptedness of Y, and $\mathcal{F}_s\subseteq\mathcal{F}_t$ by the definition of filtration.

Stopping times can be used to stop processes.

Definition 3.4.2. For an adapted process X (in discrete or continuous time) and a stopping time τ, the process X^τ defined by $X^\tau_t = X_{t\wedge\tau}$ is said to be X stopped at τ.

Stopping times interact well with martingales: stopped martingales are still martingales.

Proposition 3.4.1. *Let X be a discrete-time martingale and let τ be a stopping time. Then X^τ is a martingale.*

Remark 3.4.1. A version of this theorem also holds for continuous-time martingales with continuous sample paths.

Proof. Note that

$$X^\tau_t = X_0 + \sum_{s=1}^{t}\mathbb{1}_{\{s\le\tau\}}\left(X_s - X_{s-1}\right).$$

Since the event $\{t\le\tau\} = \{\tau\le t-1\}^c$ is $\mathcal{F}_{t-1}$-measurable by the definition of stopping time, the process $K_t = \mathbb{1}_{\{t\le\tau\}}$ is predictable. Since X^τ is the martingale transform of the bounded predictable process K with respect to the martingale X, it is a martingale. $\square$

The above result says that the martingale property is stable under stopping. We use this property as motivation for the following definition.

Definition 3.4.3. A local martingale is an adapted process $X = (X_t)_{t \geq 0}$, in either discrete or continuous time, such that there exists an increasing sequence of stopping times (τ_N) with $\tau_N \uparrow \infty$ such that the stopped process X^{τ_N} is a martingale for each N.

Remark 3.4.2. Note that martingales are local martingales. Indeed, given a martingale X and any sequence of stopping times $\tau_N \uparrow \infty$, the stopped process X^{τ_N} is a martingale.

Remark 3.4.3. Note that the local martingale property is also stable under stopping. Indeed, let X be a local martingale and τ a stopping time. Then by definition, there exists a sequence of stopping times $\sigma_N \uparrow \infty$ such that X^{σ_N} is a martingale. Hence $(X^{\sigma_N})^\tau = X^{\sigma_N \wedge \tau}$ is again a martingale since $\sigma_N \wedge \tau$ is a stopping time. But note that $X^{\sigma_N \wedge \tau} = (X^\tau)^{\sigma_N}$, implying that the sequence of stopping times $\sigma_N \uparrow \infty$ is such that $(X^\tau)^{\sigma_N}$ is a martingale. This means X^τ is a local martingale.

Theorem 3.4.4. *Suppose M is a discrete-time local martingale and K is a predictable process. Let*

$$X_t = \sum_{s=1}^{t} K_s \left(M_s - M_{s-1}\right)$$

for $t \geq 1$. Then X is a local martingale.

Remark 3.4.5. This is the martingale transform as before, but now do not insist that K is bounded or that M is a true martingale. As a consequence, we cannot assert that X is a true martingale, merely a local martingale. The idea is that by localising, we can study the algebraic and measurability structure of the martingale transform without worrying about integrability issues.

Proof. Since M is a local martingale by assumption, there exists a sequence of stopping times $(\tau_n)_n$ with $\tau_n \uparrow \infty$ a.s. such that M^{τ_n} is a martingale.
Let $u_n = \inf \{t \geq 0 : |K_{t+1}| > n\}$ with the convention $\inf \emptyset = +\infty$. Note that since K is predictable we have

$$\{u_n \leq t-1\}^c = \{u_n \geq t\} = \{|K_s| \leq n \text{ for all } 0 \leq s \leq t\} \in \mathcal{F}_{t-1}$$

and hence that u_n is a stopping time with $u_n \uparrow \infty$. Finally, let $v_n = \tau_n \wedge u_n$. Note $v_n \uparrow \infty$ and v_n is a stopping time since $\{v_n \leq t\} = \{\tau_n \leq t\} \cup \{u_n \leq t\}$.
Now $M^{v_n} = (M^{\tau_n})^{u_n}$ is a stopped martingale, and hence a martingale. Also $\left(K_t \mathbb{1}_{\{t \leq v_n\}}\right)_{t \geq 1}$ is a predictable process, bounded by n. Writing

$$X_t^{v_n} = \sum_{s=1}^{t} K_s \mathbb{1}_{\{s \leq v_n\}} \left(M_s^{v_n} - M_{s-1}^{v_n}\right)$$

we see that the stopped process is the martingale transform of a bounded predictable process with respect to the martingale, and hence is a martingale. ☐

The next theorem gives a sufficient condition that a local martingale is a true martingale.

Theorem 3.4.6. *Let X be a local martingale in either discrete or continuous time. Let Y_t be a process such that $|X_s| \leq Y_t$ almost surely for all $0 \leq s \leq t$. If $\mathbb{E}[Y_t] < \infty$ for all $t \geq 0$, then X is a true martingale.*

Proof. Let $(\tau_N)_N$ be a localising sequence of stopping times for X. Note that $X_{t \wedge \tau_N} \to X_t$ a.s. since $\tau_N \uparrow \infty$. Furthermore, by assumption $|X_{t \wedge \tau_N}| \leq Y_t$ which is integrable, so we may apply the conditional version of the dominated convergence theorem to conclude

$$\begin{aligned}
\mathbb{E}[X_t \mid \mathcal{F}_s] &= \mathbb{E}\left[\lim_N X_{t \wedge \tau_N} \mid \mathcal{F}_s\right] \\
&= \lim_N \mathbb{E}[X_{t \wedge \tau_N} \mid \mathcal{F}_s] \\
&= \lim_N X_{s \wedge \tau_N} \\
&= X_s
\end{aligned}$$

for $0 \leq s \leq t$, where we have used the fact that the stopped process $(X_{t \wedge \tau_N})_{t \geq 0}$ is a martingale. □

The following corollary is useful:

Corollary 3.4.1. *Suppose X is a DISCRETE-TIME local martingale such that $\mathbb{E}[|X_t|] < \infty$ for all $t \geq 0$. Then X is a true martingale.*

Proof. Let $Y_t = |X_0| + \ldots + |X_t|$. The process Y is integrable by assumption and $|X_s| \leq Y_t$ for all $0 \leq s \leq t$. The conclusion follows from the previous theorem. □

In the absence of integrability, the next best property is non-negativity. First we need some definitions.

Definition 3.4.4. A supermartingale relative to a filtration $(\mathcal{F}_t)_{t \geq 0}$ is an adapted stochastic process $(U_t)_{t \geq 0}$ with the following properties:

- $\mathbb{E}[|U_t|] < \infty$ for all $t \geq 0$,
- $\mathbb{E}[U_t \mid \mathcal{F}_s] \leq U_s$ for all $0 \leq s \leq t$.

A submartingale is an adapted process $(V_t)_{t \geq 0}$ with the following properties:

- $\mathbb{E}[|V_t|] < \infty$ for all $t \geq 0$,
- $\mathbb{E}[V_t \mid \mathcal{F}_s] \geq V_s$ for all $0 \leq s \leq t$.

Remark 3.4.7. Hence a supermartingale decreases on average, while a submartingale increases on average. A martingale is a stochastic process that is both a supermartingale and a submartingale.

As in the case of the definition of martingale, to show that an adapted, integrable process U is a supermartingale in discrete time, it is enough to show that $\mathbb{E}\left[U_{t+1} \mid \mathcal{F}_t\right] \leq U_t$ for all $t \geq 0$.

Theorem 3.4.8. *Suppose X is a local martingale in either continuous or discrete time. If $X_t \geq 0$ for all $t \geq 0$, then X is a supermartingale.*

Proof. In the general case, let $(\tau_N)_N$ be the localising sequence for X. First we show that X_t is integrable for each $t \geq 0$. Fatou's lemma yields

$$\begin{aligned}\mathbb{E}\left[|X_t|\right] &= \mathbb{E}\left[X_t\right] \\ &= \mathbb{E}\left[\lim_N X_{t\wedge\tau_N}\right] \\ &\leq \liminf_N \mathbb{E}\left[X_{t\wedge\tau_N}\right] \\ &= X_0 < \infty.\end{aligned}$$

Now that we have established integrability, we can discuss conditional expectations. The conditional version of Fatou's lemma yields

$$\begin{aligned}\mathbb{E}\left[X_t \mid \mathcal{F}_s\right] &= \mathbb{E}\left[\lim_N X_{t\wedge\tau_N} \mid \mathcal{F}_s\right] \\ &\leq \liminf_N \mathbb{E}\left[X_{t\wedge\tau_N} \mid \mathcal{F}_s\right] \\ &= \liminf_N X_{s\wedge\tau_N} \\ &= X_s\end{aligned}$$

for $0 \leq s \leq t$, as claimed. ☐

As before, discrete time local martingales are particularly nice:

Corollary 3.4.2. *If X is a DISCRETE-TIME local martingale such that $X_t \geq 0$ a.s. for all $t \geq 0$, then X is a martingale.*

Proof. By the above theorem, we have that $\mathbb{E}\left[|X_t|\right] = \mathbb{E}\left[X_t\right] \leq X_0 < \infty$. Since X is integrable, the previous corollary implies X is a martingale. ☐

Theorem 3.4.9. *Suppose that*

$$X_t = X_0 + \sum_{s=1}^{t} K_s\left(M_s - M_{s-1}\right)$$

where K is predictable, M is a local martingale and X_0 is a constant. If $X_T \geq 0$ a.s. for some non-random $T > 0$, then $(X_t)_{0\leq t\leq T}$ is a true martingale.

Proof. Just as before, let $v_N = \inf\{t \geq 0 : |K_{t+1}| > N\} \wedge \tau_N$, where $(\tau_N)_N$ is a localising sequence for X. Note $X_s \mathbb{1}_{\{t \leq \tau_N\}}$ is integrable on $\{t \leq \tau_N\}$ for all $0 \leq s \leq t$, since M^{v_N} is integrable by definition of martingale, and K_s is bounded on $\{t \leq \tau_N\}$. Hence we have

$$\begin{aligned}
0 &\leq \mathbb{E}\left[X_T \mathbb{1}_{\{T \leq v_N N\}} \mid \mathcal{F}_{T-1}\right] \\
&= \mathbb{E}\left[X_{T-1}\mathbb{1}_{\{T \leq v_N\}} + K_T \mathbb{1}_{\{T \leq v_N\}}\left(M_T^{v_N} - M_{T-1}^{v_N}\right) \mid \mathcal{F}_{T-1}\right] \\
&= X_{T-1}\mathbb{1}_{\{T \leq v_N\}} + K_T \mathbb{1}_{\{T \leq v_N\}}\mathbb{E}\left[M_T^{v_N} - M_{T-1}^{v_N} \mid \mathcal{F}_{T-1}\right] \\
&= X_{T-1}\mathbb{1}_{\{T \leq v_N\}},
\end{aligned}$$

where we have used the fact that $X_{T-1}, \mathbb{1}_{\{T \leq v_N\}}$ and K_T are $\mathcal{F}_{T-1}$-measurable and M^{v_N} is a true martingale. Taking $N \to \infty$ shows $X_{T-1} \geq 0$ a.s., induction shows that $X_t \geq 0$ for all $0 \leq t \leq T$. Therefore $(X_t)_{0 \leq t \leq T}$ is a non-negative local martingale in discrete time and hence a true martingale. □

3.5 Proof of the 1FTAP, easier direction

Recall our framework. There exist n-dimensional price $(P_t)_{t \geq 0}$, adapted to a given filtration $(\mathcal{F}_t)_{t \geq 0}$. A martingale deflator is a positive (real-valued) adapted process $(Y_t)_{t \geq 0}$ such that YP is integrable and

$$\mathbb{E}\left[Y_{t+1} \cdot P_{t+1} \mid \mathcal{F}_t\right] = Y_t P_t \text{ for all } t \geq 0$$

Our aim is to show that if there exists a martingale deflator, then there is no arbitrage. We first prove a useful lemma.

Lemma 3.5.1. *Given a real constant x and a n-dimensional previsible process $(H_t)_{t \geq 1}$, let*

$$\begin{aligned}
X_0 &= x \\
X_t &= H_t \cdot P_t \textit{ for } t \geq 1
\end{aligned}$$

and

$$c_t = X_t - H_{t+1} \cdot P_t \textit{ for } t \geq 0$$

Suppose there exists a martingale deflator Y. Set

$$Z_t = X_t Y_t + \sum_{s=0}^{t-1} c_s Y_s.$$

Then Z is a local martingale. Furthermore, if $c_t \geq 0$ a.s for $0 \leq t \leq T-1$ and $X_T \geq 0$ a.s for some non-random $T > 0$, then $(M_t)_{0 \leq t \leq T}$ is a martingale.

Proof. Recall that M is a martingale by the definition of martingale deflators. Note that

$$\begin{aligned} Z_t - Z_{t-1} &= X_t Y_t + Y_{t-1}\left(c_{t-1} - X_{t-1}\right) \\ &= H_t \cdot \left[Y_t P_t - Y_{t-1} P_{t-1}\right] \\ &= H_t \cdot \left(M_t - M_{t-1}\right). \end{aligned}$$

Note that Z is martingale transform of the previsible process H with respect to the martingale M, and hence Z is a local martingale. Furthermore, if $Z_T \geq 0$ a.s., from the last theorem of the previous section, we have $(Z_t)_{0 \leq t \leq T}$ is a true martingale. □

Proof of 1st FTAP, easy direction. Suppose that there is a martingale deflator Y. Let H be an n-dimensional previsible process and let $c_0 = -H_1 \cdot P_0$ and

$$c_t = H_t \cdot P_t - H_{t+1} \cdot P_t \qquad \text{for } t \geq 1.$$

Suppose there is some non-random $T > 0$ such that $c_T = H_T \cdot P_T$ and $c_t \geq 0$ almost surely for all $0 \leq t \leq T$. To show that H is not an arbitrage, must show that $c_t = 0$ almost surely for all $0 \leq t \leq T$. To this end, let

$$Z_T = \sum_{s=0}^{T} c_s Y_s.$$

Since $Y_s > 0$ and $c_s \geq 0$ for all $0 \leq s \leq T$, we need only show that $Z_T = 0$ almost surely. By the pigeon-hole principle, it is sufficient to show

$$\mathbb{E}\left[Z_T\right] = 0.$$

To finish the proof, let

$$Z_t = X_t Y_t + \sum_{s=0}^{t-1} c_s Y_s \qquad \text{for } 0 \leq t \leq T$$

where we have set $X_0 = 0$ and $X_t = H_t \cdot P_t$ for $1 \leq t \leq T$. By the previous lemma, the process $(Z_t)_{0 \leq t \leq T}$ is a martingale. Since $Z_0 = 0$, we are done. □

3.6 Proof of harder direction of the 1FTAP

In this section we will present elements of the proof of the first fundamental theorem of asset pricing. We will give a complete proof of the one-period case, and sketch the main steps to prove the full multi-period case. But first, we take a moment to reflect on the economic motivation.

3.6.1 Motivation: Langrangian duality

Consider the investor's utility maximisation problem to find an investment strategy H to

$$\text{maximise } \mathbb{E}[U(c)] \qquad \text{subject to } \begin{cases} c_0 = x - H_1 \cdot P_0, \\ c_t = (H_t - H_{t+1}) \cdot P_t & \text{for } 1 \le t \le T-1 \\ c_T = H_T \cdot P_T \end{cases}$$

where the utility function u is increasing in each argument. We have discussed previously that the existence of a maximiser implies that there does not exist an arbitrage. We now explain why the existence of a maximiser also points to the existence of a martingale deflator. As usual in a constrained optimisation problem, we apply the Lagrangian method. Recall that this involves replacing our given objective function with the so-called Lagrangian which encodes the constraints. In this case the Lagrangian is

$$\begin{aligned} L(c, H, Y) = & \mathbb{E}\left[u\left(c_0, \ldots, c_T\right)\right] + Y_0\left(x - c_0 - H_1 \cdot P_0\right) \\ & + \mathbb{E}\left[\sum_{t=1}^{T-1} Y_t\left[H_t \cdot P_t - H_{t+1} \cdot P_t - c_t\right] + Y_T\left[H_T \cdot P_T - c_T\right]\right] \end{aligned}$$

where the real-valued adapted process $(Y_t)_{0 \le t \le T}$ is the family of Lagrange multipliers. To identify the dual feasibility condition, we seek to find conditions on the Lagrange multiplier process Y implied by the existence of a maximiser of the Lagrangian $L(c, H, Y)$ over adapted c and previsible H. To this end, we rewrite the Lagrangian as

$$\begin{aligned} L(c, H, Y) = & \mathbb{E}\left[u\left(c_0, \ldots, c_T\right) - \sum_{t=0}^{T} Y_t c_t\right] + x Y_0 \\ & + \sum_{t=1}^{T} \mathbb{E}\left[H_t \cdot \left[P_t Y_t - P_{t-1} Y_{t-1}\right]\right] \end{aligned}$$

By formally differentiating with respect to the t-th consumption variable, we find the maximised consumption c^* satisfies

$$\mathbb{E}\left[\frac{\partial u}{\partial c_t}\left(c^*\right) \mid \mathcal{F}_t\right] = Y_t.$$

Differentiating with respect to H_t yields

$$\mathbb{E}\left[P_t Y_t \mid \mathcal{F}_{t-1}\right] = P_{t-1} Y_{t-1}$$

These two conditions suggest that Y is an adapted, positive process such that the process

$$M_t = Y_t P_t$$

is a martingale - i.e. Y is a martingale deflator.

3.6.2 Proof when $T = 1$

We now proceed to turn those fuzzy heuristics into a proper proof. We consider the one-period case. So our market data are the prices P_0 and P_1.
We suppose that the market $(P_t)_{t\in\{0,1\}}$ has no arbitrage, so that for any vector $H \in \mathbb{R}^n$ that has the property that $H \cdot P_0 \leq 0 \leq H \cdot P_1$ almost surely, it must be the case that $H \cdot P_0 = 0 = H \cdot P_1$ almost surely. We will show that, given any positive random variable Z, there exists a martingale deflator Y_0, Y_1 such that the product $Y_1 Z$ is bounded by a constant. This extra boundedness assumption is much stronger than what we need, but it comes for free from the proof and we will find it useful later in the course. Let

$$\zeta = \frac{e^{-\|P_1\|^2/2}}{1+Z}$$

Define a function $F : \mathbb{R}^n \to \mathbb{R}$ by

$$F(H) = e^{H\cdot P_0} + \mathbb{E}\left[e^{-H\cdot P_1}\zeta\right]$$

The positive random variable ζ is introduced to ensure integrability. Indeed note that the integrand $e^{-H\cdot P_1}\zeta \leq e^{\|H\|^2/2}$ is bounded for each choice of H. In particular, the function F is finite everywhere and (by the dominated convergence theorem) smooth.
We will show that no investment-consumption arbitrage implies that the function F has a minimiser H^*. By the first order condition for a minimum, we have

$$0 = \nabla F\left(H^*\right) = e^{H^*\cdot P_0} P_0 - \mathbb{E}\left[e^{-H^*\cdot P_1}\zeta P_1\right]$$

and hence we may take

$$Y_0 = e^{H^*\cdot P_0} \text{ and } Y_1 = e^{-H^*\cdot P_1}\zeta$$

Note that $Y_1 Z < C$ for some constant $C > 0$ (which depends on H^* in general). So let $(H_k)_k$ be a sequence such that $F(H_k) \to \inf_H F(H)$. If $(H_k)_k$ is bounded, we can pass to a convergent subsequence, by the Bolzano-Weierstrass theorem, such that $H_k \to H^*$. By the smoothness of F we have

$$\inf_H F(H) = \lim_k F(H_k) = F\left(\lim_k H_k\right) = F\left(H^*\right)$$

so H^* is our desired minimiser. It remains to show that no arbitrage implies that there exists a bounded minimising sequence $(H_k)_k$. We show that if every minimising sequence $(H_k)_k$ is unbounded, then there would be a contradiction. Now we arrive at a little technicality. Let

$$\mathcal{U} = \{u \in \mathbb{R}^n : u \cdot P_0 = 0 = u \cdot P_1 \text{ a.s. }\} \subseteq \mathbb{R}^n$$

and let

$$\mathcal{V} = \mathcal{U}^\perp$$

Notice that if $u \in \mathcal{U}$ and $v \in \mathcal{V}$ then $F(u+v) = F(v)$. Hence, by projecting a given minimising sequence onto the subspace $\mathcal{V}$, it is sufficient to consider minimising sequence $(H_k)_k$ taking values in $\mathcal{V}$.
So suppose, for the sake of finding a contradiction, that all minimising sequence $(H_k)_k$ taking values in $\mathcal{V}$ are unbounded. We can pass to a subsequence such that $\|H_k\| \uparrow \infty$. Now let

$$\hat{H}_k = \frac{H_k}{\|H_k\|}.$$

Note that $\left\|\hat{H}_k\right\| = 1$ and that $\hat{H}_k \in \mathcal{V}$. Since $\left(\hat{H}_k\right)_k$ is bounded, we can again pass to a convergent subsequence such that $\hat{H}_k \to \hat{H}$. Notice once more that $\|\hat{H}\| = 1$ and that $\hat{H} \in \mathcal{V}$.
We know that the sequence $F(H_k)$ is bounded (since it is convergent) but we also have

$$F(H_k) = \left(e^{\hat{H}_k \cdot P_0}\right)^{\|H_k\|} + \mathbb{E}\left[\left(e^{-\hat{H}_k \cdot P_1}\right)^{\|H_k\|}\zeta\right]$$

so we must conclude that $\hat{H} \cdot P_0 \leq 0 \leq \hat{H} \cdot P_1$ a.s. (since otherwise the right-hand side would blow up). By no-arbitrage, we conclude that the candidate arbitrage $\hat{H}$ is not actually an arbitrage, so $\hat{H} \cdot P_0 = 0 = \hat{H} \cdot P_1$ a.s.
We have shown that $\hat{H}$ is in $\mathcal{U}$. But since $\hat{H}$ is also in $\mathcal{V} = \mathcal{U}^\perp$, so we have $\hat{H} = 0$. And, finally, this contradicts $\|\hat{H}\| = 1$.

3.6.3 Elements of the proof of the harder direction of the multi-period 1FTAP

We have already seen the one period case. The full multi-period proof is a little more difficult because of some technicalities involving measurability.
We begin with two propositions that show that two of the existential-type results we needed in one-period proof have measurable versions.

Proposition 3.6.1. *Let $f : \mathbb{R}^n \times \Omega \to \mathbb{R}$ be such that $f(x, \cdot)$ is measurable for all x, and that $f(\cdot, \omega)$ is continuous and has a unique minimiser $X^*(\omega)$ for each ω. Then X^* is measurable.*

Let us pause to think about what it means for the unique minimiser $x^* \in \mathbb{R}^n$ of a continuous function $g : \mathbb{R}^n \to \mathbb{R}$ to be in a closed rectangle of the form $A = [a_1, b_1] \times \cdots \times [a_n, b_n]$.
First note that that for all $q \neq x^*$, we have $g(x^*) < g(q)$. Let $Q \subset \mathbb{R}^n$ be countable and dense - for instance, let Q be the set of points with rational coordinates. By the continuity of g and the density of Q, we have that for every $q \neq x^*$, there exists a $p \in Q$ such that $g(p) < g(q)$.
Now if $x^* \in A$, then for any $q \in A^c \cap Q$, there exists a $p \in A \cap Q$ such that $g(p) < g(q)$, since $q \neq x^*$ and $A \cap Q$ is dense in A.
Conversely, suppose that for any $q \in A^c \cap Q$, there exists a $p \in A \cap Q$ such that $g(p) < g(q)$. This means

$$\inf_{x \in A \cap Q} g(x) \leq \inf_{x \in A^c \cap Q} g(x).$$

By the continuity of g the above inequality implies

$$\inf_{x\in A} g(x) = \inf_{x\in A\cup A^c} g(x) = g(x^*)$$

and in particular, we have $x^* \in A$ since A is closed.

Proof. For any closed rectangle $A \subset \mathbb{R}^n$ we have

$$\{\omega : X^*(\omega) \in A\} = \bigcap_{q\in A^c\cap Q} \bigcup_{p\in A\cap Q} \{\omega : f(p,\omega) < f(q,\omega)\}$$

where Q is a countable dense subset of $\mathbb{R}^n$. Since the Borel σ-algebra is generated by such rectangles, this implies the measurability of X^*. □

We also need a useful measurable version of the Bolzano-Weierstrass theorem.

Proposition 3.6.2. *Let $(\xi_i)_{i\geq 1}$ be a sequence of measurable functions $\xi_i : \Omega \to \mathbb{R}^n$ such that $\sup_i \|\xi_i(\omega)\| < \infty$ for all $\omega \in \Omega$. Then there exists an increasing sequence of integer-valued measurable functions I_j and an $\mathbb{R}^n$-valued measurable function ξ^* such that*

$$\xi_{I_j(\omega)}(\omega) \to \xi^*(\omega) \text{ as } j \to \infty$$

for all $\omega \in \Omega$.

Proof. First we consider the $n = 1$ case. Let $\xi^*(\omega) = \limsup_i \xi_i(\omega)$. Note ξ^* is finite-valued and measurable, and that for every $j > 0$ there exists an infinite number of i 's such that $\xi_i(\omega) \geq \xi^*(\omega) - 1/j$. Now let

$$I_j = \inf\{i \geq j : \xi_i \geq \xi^* - 1/j\}$$

Since we have the representation of the event

$$\{I_j \leq h\} = \cup_{i=j}^h \{\xi_i \geq \xi^* - 1/j\}$$

for each $h \geq j$, the function I_j is measurable and $\xi_{I_j} \to \xi^*$ as desired. Now we prove the claim for any dimension $n \geq 1$ by induction. Suppose that the claim is true for dimension $n = N$. Let $(\xi_i)_i$ be a sequence of measurable function valued in $\mathbb{R}^{N+1}$ such that $\sup_i \|\xi_i(\omega)\| < \infty$. Writing $\xi_i = (\zeta_i, \eta_i)$ where ζ_i takes values in $\mathbb{R}^N$ and η_i takes values in $\mathbb{R}$, we have by assumption the existence of a measurable sequence I_j and a measurable ζ^* such that

$$\zeta_{I_j} \to \zeta^*.$$

Notice that $\left(\eta_{I_j(\omega)}(\omega)\right)_j$ is bounded for each ω, and hence by the $n = 1$ case, there exists an increasing measurable sequence J_k and a measurable η^* such that $\eta_{I_{J_k}} \to \eta^*$. In particular,

$$\xi_{I_{J_k}} \to (\zeta^*, \eta^*) = \xi^*$$

as desired. □

3.7 Numéraires and equivalent martingale measures

It is sometimes useful (for example, for the example sheet) to introduce some vocabulary:

Definition 3.7.1. In a market with no dividends, a pricing kernel (or stochastic discount factor or state price density) between times s and t, where $0 \leq s < t$, is a positive $\mathcal{F}_t$-measurable random variable $\rho_{s,t}$ such that

$$P_s = \mathbb{E}\left[\rho_{s,t} P_t \mid \mathcal{F}_s\right].$$

Let Y be a martingale deflator, so that $\mathbb{E}\left[Y_t P_t \mid \mathcal{F}_s\right] = Y_s P_s$ for all $0 \leq s < t$, and let $\rho_{s,t} = Y_t / Y_s$. If $\rho_{s,t} P_t$ is integrable, then $\rho_{s,t}$ is a pricing kernel between times s and t. Conversely, suppose $\rho_{s,s+1}$ is a pricing kernel between times s and $s+1$ for $s \geq 0$, and let $Y_t = \rho_{0,1} \cdots \rho_{t-1,t}$. Then assuming YP is integrable, then Y is a martingale deflator.
The point of this section is to consider the effect of adding the further assumption that there exists a portfolio with a strictly positive price.

Definition 3.7.2. A numraire portfolio is n-dimensional previsible process $\eta = (\eta_t)_{t \geq 0}$ such that $\eta_t \cdot P_t > 0$ almost surely for each $t \geq 0$, and satisfying the pure-investment self-financing condition

$$(\eta_t - \eta_{t+1}) \cdot P_t = 0$$

A numéraire asset is an asset with a strictly positive price.

Remark 3.7.1. If asset i is a numéraire asset, then the constant portfolio η defined for all $t \geq 0$ by

$$\eta_t^j = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}$$

is a numéraire portfolio.

In the context of a market with a numéraire, we consider a different type of arbitrage:

Definition 3.7.3. A terminal consumption arbitrage is a n-dimensional previsible process H such that there exists a non-random $T > 0$ such that almost surely we have

$$\begin{aligned} c_0 &= -H_1 \cdot P_0 = 0 \\ c_t &= (H_t - H_{t+1}) \cdot P_t = 0 \text{ for } 1 \leq t \leq T-1 \\ c_T &= H_T \cdot P_T \geq 0 \end{aligned}$$

and

$$\mathbb{P}\left(c_T > 0\right) > 0.$$

Note that our first definition of arbitrage allowed for consumption at intermediate times. Clearly, a terminal consumption arbitrage is an arbitrage according to our earlier definition. We now ask when the existence of an arbitrage (possibly with intermediate consumption) implies the existence of a terminal consumption arbitrage.

Proposition 3.7.1. *Consider a market with a numéraire η and let $N_t = \eta_t \cdot P_t$ for $t \geq 0$. Let H be an investment-consumption strategy with consumption stream*

$$\begin{aligned} c_0 &= x - H_1 \cdot P_0 \\ c_t &= (H_t - H_{t+1}) \cdot P_t, \end{aligned}$$

where x is the initial wealth. Let

$$K_t = H_t + \eta_t \sum_{s=0}^{t-1} \frac{c_s}{N_s}$$

for $t \geq 1$. Then K is a pure-investment strategy from the same initial wealth x. (That is, K be the strategy that consists of holding at time t the portfolio H_t but of instead of consuming the amount c_t, instead invest this money into the numéraire portfolio.)
In particular, H is an arbitrage if and only if K is a terminal consumption arbitrage.

Proof. Note that

$$\begin{aligned} (K_t - K_{t+1}) \cdot P_t = &(H_t - H_{t+1}) \cdot P_t - \eta_{t+1} \cdot P_t \frac{c_t}{N_t} \\ &+ (\eta_t - \eta_{t+1}) \cdot P_t \sum_{s=1}^{t-1} \frac{c_s}{N_s} \\ =&0 \end{aligned}$$

so K is a pure investment strategy by the assumption that η is pure-investment. Suppose $c_T = H_T \cdot P_T$ for some non-random time T. Then

$$K_T \cdot P_T = N_T \sum_{s=0}^{T} \frac{c_s}{N_s} \geq 0$$

The left-hand side is positive if and only if c_t is positive for some $0 \leq t \leq T$. □

Now we move to a definition that only makes sense in a market with a numéraire.

Definition 3.7.4. Let P be a market model defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The measure $\mathbb{P}$ is called the objective (or historical or statistical) measure for the model. Suppose that there exists a numéraire portfolio η and let $N = \eta \cdot P$. An equivalent martingale measure relative to this numéraire is any probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that the discounted price processes

$$\left(\frac{P_t}{N_t} \right)_{t \geq 0}$$

is a martingale under $\mathbb{Q}$.

Before proceeding, it might be useful to recall some facts from probability theory.

Definition 3.7.5. Let $(\Omega, \mathcal{F})$ be a measurable space and let $\mathbb{P}$ and $\mathbb{Q}$ be two probability measures on $(\Omega, \mathcal{F})$. The measures $\mathbb{P}$ and $\mathbb{Q}$ are equivalent, written $\mathbb{P} \sim \mathbb{Q}$, iff

$$\{A \in \mathcal{F} : \mathbb{P}(A) = 1\} = \{A \in \mathcal{F} : \mathbb{Q}(A) = 1\}$$

The above definition says that equivalent probability measures have the same almost sure events. Complementarily, equivalent probability measures have the same null sets: that is, $\mathbb{P} \sim \mathbb{Q}$ iff

$$\{A \in \mathcal{F} : \mathbb{P}(A) = 0\} = \{A \in \mathcal{F} : \mathbb{Q}(A) = 0\}$$

It turns out that equivalent measures can be characterised by the following theorem. When there are more than one probability measure floating around, we use the notation $\mathbb{E}^{\mathbb{P}}$ to denote expected value with respect to $\mathbb{P}$, etc.

Theorem 3.7.2 (Radon-Nikodym theorem)**.** *The probability measure $\mathbb{Q}$ is equivalent to the probability measure $\mathbb{P}$ if and only if there exists a $\mathbb{P}$-a.s. positive random variable Z such that*

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}\left[Z\mathbb{1}_A\right]$$

for each $A \in \mathcal{F}$.

The random variable Z is called the density, or the Radon-Nikodym derivative, of $\mathbb{Q}$ with respect to $\mathbb{P}$, and is often denoted

$$Z = \frac{d\mathbb{Q}}{d\mathbb{P}}$$

Note that the Radon-Nikodym derivative satisfies the identity

$$\mathbb{Q}(A) = \int_A \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P}.$$

Also, note that $\mathbb{P}$ has a density with respect to $\mathbb{Q}$ given by

$$\frac{d\mathbb{P}}{d\mathbb{Q}} = \frac{1}{Z}$$

We only need the easy direction of the theorem, that the existence of a positive density implies equivalence, for this course. Here is a proof. The proof of the harder direction is omitted since we do not need it.

Proof. Suppose $\mathbb{P}(Z > 0) = 1$ and that $\mathbb{E}^{\mathbb{P}}[Z] = 1$. Define a set function $\mathbb{Q}$ by $\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}\left[Z\mathbb{1}_A\right]$. Note that $\mathbb{Q}$ is countably additive by the monotone convergence theorem. Also, $\mathbb{Q}(\Omega) = \mathbb{E}^{\mathbb{P}}[Z] = 1$, so $\mathbb{Q}$ is a probability measure. If $\mathbb{P}(A) = 0$, then the event $\{\mathbb{1}_A = 0\}$ is $\mathbb{P}$-almost sure and hence

$$\mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}\left[Z\mathbb{1}_A\right] = 0$$

Conversely, if $\mathbb{Q}(A) = 0$ we can conclude that $\{Z\mathbb{1}_A = 0\}$ is $\mathbb{P}$-a.s. by the pigeon-hole principle since $\{Z\mathbb{1}_A \geq 0\}$ is $\mathbb{P}$-a.s. But since $\{Z > 0\}$ is $\mathbb{P}$-a.s., we must conclude that $\{\mathbb{1}_A = 0\}$ is $\mathbb{P}$-a.s., i.e. $\mathbb{P}(A) = 0$. Thus $\mathbb{Q}$ and $\mathbb{P}$ are equivalent. $\square$

Example 3.7.1. Consider the sample space $\Omega = \{1, 2, 3\}$ with the set $\mathcal{F}$ of events all subsets of Ω. Consider probability measures $\mathbb{P}$ and $\mathbb{Q}$ defined by

- $\mathbb{P}\{1\} = \frac{1}{2}, \mathbb{P}\{2\} = \frac{1}{2}$, and $\mathbb{P}\{3\} = 0$
- $\mathbb{Q}\{1\} = \frac{1}{1000}, \mathbb{Q}\{2\} = \frac{999}{1000}$, and $\mathbb{Q}\{3\} = 0$.

Then $\mathbb{P}$ and $\mathbb{Q}$ are equivalent. We may take their density $Z = \frac{d\mathbb{Q}}{d\mathbb{P}}$ to be

$$Z(1) = \frac{1}{500}, \quad Z(2) = \frac{999}{500}, \quad Z(3) = 0.$$

(Since both measures don't 'see' the event $\{3\}$, we can let $Z(3)$ be any value.)

Now, returning to our financial model, we have a result that says, that in a market with a numéraire, the notion of an equivalent martingale measure is morally the same as the notion of a martingale deflator.

Proposition 3.7.2. *Suppose the market has a numéraire, and fix a non-random time horizon $T > 0$. The market model $(P_t)_{0\leq t\leq T}$ has an equivalent martingale measure relative to the numéraire if and only if there is a martingale deflator.*

Proof. Let Y be a process such that $\{Y_T > 0\}$ is $\mathbb{P}$-a.s. and such that $Y_T P_T$ is $\mathbb{P}$ integrable. Define a new measure $\mathbb{Q}$ by the density

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{Y_T N_T}{\mathbb{E}^{\mathbb{P}}\left[Y_T N_T\right]}$$

Our analysis turns on the Bayes formula

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{P_T}{N_T} \mid \mathcal{F}_t\right] = \frac{\mathbb{E}^{\mathbb{P}}\left[P_T Y_T \mid \mathcal{F}_t\right]}{\mathbb{E}^{\mathbb{P}}\left[N_T Y_T \mid \mathcal{F}_t\right]}$$

Suppose Y is a martingale deflator. We have

$$\mathbb{E}^{\mathbb{P}}\left[P_T Y_T \mid \mathcal{F}_t\right] = P_t Y_t$$

by definition. Also note that

$$Y_t N_t - Y_{t-1} N_{t-1} = \eta_t \cdot (Y_t P_t - Y_{t-1} P_{t-1})$$

and hence YN is a local martingale. However, since YN is non-negative, we know from last section that YN is a true martingale. In particular

$$\mathbb{E}^{\mathbb{P}}\left[N_T Y_T \mid \mathcal{F}_t\right] = N_t Y_t$$

By the Bayes formula we have

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{P_T}{N_T} \mid \mathcal{F}_t\right] = \frac{P_t}{N_t}$$

and hence P/N is a $\mathbb{Q}$-martingale, i.e. $\mathbb{Q}$ is an equivalent martingale measure. Conversely, suppose $\mathbb{Q}$ is an equivalent martingale measure. Let

$$Z_t = \mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_t\right].$$

Note that Z is a positive $\mathbb{P}$-martingale. Let

$$Y_t = Z_t/N_t$$

Since the random variable P_T/N_T is $\mathbb{Q}$-integrable by the definition of martingale, we can conclude that $P_T Y_T$ is $\mathbb{P}$-integrable. Furthermore, the process Y is positive and satisfies

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[N_T Y_T \mid \mathcal{F}_t\right] &= \mathbb{E}^{\mathbb{P}}\left[Z_T \mid \mathcal{F}_t\right] \\
&= Z_t \\
&= N_t Y_t.
\end{aligned}$$

Hence by the Bayes formula

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[P_T Y_T \mid \mathcal{F}_t\right] &= \mathbb{E}^{\mathbb{Q}}\left[\frac{P_T}{N_T} \mid \mathcal{F}_t\right] \mathbb{E}^{\mathbb{P}}\left[N_T Y_T \mid \mathcal{F}_t\right] \\
&= \frac{P_t}{N_t}\left(N_t Y_t\right) \\
&= P_t Y_t
\end{aligned}$$

so that PY is a $\mathbb{P}$-martingale and hence Y is a martingale deflator. □

Combining the two results of this section, we have the usual formulation of the first fundamental theorem of asset pricing:

Theorem 3.7.3 (First Fundamental Theorem of Asset Pricing when there is a numéraire)**.** *Suppose the market has is a numéraire, and fix a non-random time horizon $T > 0$. The market model $(P_t)_{0 \le t \le T}$ has no terminal consumption arbitrage if and only if there exists an equivalent martingale measure relative to the numéraire.*

3.8 Special numéraires and equivalent martingale measures

In this section we consider two classes of numéraire assets (and their respective equivalent martingale measures) that arise very frequently in applications.

Definition 3.8.1. A (risk-free zero-coupon) bond is an asset such that there exists a nonrandom time $T > 0$ (called its maturity date) and such that its price at time T is a nonrandom positive constant (called its face value or principal value). Unless otherwise specified, we shall assume that the face value of a bond is 1. We will denote the time t price of the bond of maturity T by P_t^T for $0 \le t \le T$.

Proposition 3.8.1. *Suppose the market contains a bond. If the market has no arbitrage, then the bond is a numéraire.*

Proof. There are at least two ways to prove this. It is important to understand both methods.

- *A 'primal' argument.* The idea is that if the bond price drops to zero or less, then an investor could lock in a risk-less profit. In particular, if there is no arbitrage, the price must stay positive. In mathematical notation, let

$$\tau = \inf\left\{0 \le t \le T : P_t^T \le 0\right\}.$$

  with the convention that $\tau = +\infty$ if $P_t^T > 0$ for all $0 \le t \le T$. Note that τ is a stopping time. Consider the predictable process $H_t = \mathbb{1}_{\{\tau < t \le T\}}$. This corresponds to a portfolio of buying the bond immediately after the price drops to zero or below and hold it until maturity. Note that

$$\begin{aligned} c_0 &= -H_1 P_0^T = -\mathbb{1}_{\{\tau=0\}} P_0^T \ge 0 \\ c_t &= (H_t - H_{t+1})\, P_t^T = -\mathbb{1}_{\{\tau=t\}} P_t^T \ge 0 \quad \text{for } 1 \le t \le T \\ c_T &= \mathbb{1}_{\{\tau<T\}} P_T^T = \mathbb{1}_{\{\tau<T\}} \ge 0, \end{aligned}$$

  since $P_T^T = 1$ by definition. If there is no arbitrage, then $c_t = 0$ a.s. for all $0 \le t \le T$. Hence $\tau = +\infty$ a.s.

- *A 'dual' argument.* Since there is no arbitrage, there exists a martingale deflator Y. Note YP^T is a martingale, and hence

$$P_t^T = \frac{1}{Y_t}\mathbb{E}\left[Y_T P_T^T \mid \mathcal{F}_t\right] = \frac{1}{Y_t}\mathbb{E}\left[Y_T \mid \mathcal{F}_t\right]$$

  since $P_T^T = 1$ by definition. Since $Y_t > 0$ a.s. for all $t \ge 0$, the conclusion follows.

□

Since we now know that no arbitrage implies that bonds are numéraires, we can discuss equivalent martingale measures:

Definition 3.8.2. An equivalent martingale measure with respect to a bond of maturity $T > 0$ is called a T-forward measure.

For an application of forward measures, we consider a forward contract:

Definition 3.8.3. A forward initiated at time t with maturity date T is a contract where, at time t, no money is exchanged and at time T, an asset with price S_T is swapped for F_t^T units of money, where F_t^T is known at time t. The quantity F_t^T is called the forward price of the asset.

Proposition 3.8.2. *Consider a market model with a bond of maturity T and forward contract initiated at time t and maturing at time T. If the market has no arbitrage, the bond is a numeraire and then*

$$F_t^T = \mathbb{E}^{\mathbb{Q}^T}\left[S_T \mid \mathcal{F}_t\right]$$

where $\mathbb{Q}^T$ is a T-forward measure.

Proof. Let D_s be the price of the forward contract at time s, for $t \le s \le T$. Since there is no arbitrage, we have

$$\frac{D_t}{P_t^T} = \mathbb{E}^{\mathbb{Q}^T}\left[D_T \mid \mathcal{F}_t\right]$$

Noting that $D_t = 0$ and $D_T = S_T - F_t^T$, and that F_t^T is $\mathcal{F}_t$-measurable yields the result, upon rearrangement. □

We now suppose there is a whole family of bonds, indexed by their maturities.

Definition 3.8.4. Consider a market with bonds of all maturities $T \in \{0, 1, 2, \ldots\}$. The spot interest rate at time t is

$$r_t = \frac{1}{P_{t-1}^t} - 1.$$

The value of the bank account (or money market account) is given by $B_0 = 1$ and

$$B_t = \prod_{s=1}^{t} (1 + r_s) \text{ for } t \ge 1.$$

for all t.

Remark 3.8.1. Note that the spot interest rate and bank account processes are previsible.

The connection between bonds and the bank account is the following:

Proposition 3.8.3. *Suppose the arbitrage-free market model has bonds of all maturities. Then there exists a pure-investment (that is, no consumption) strategy η such that*

$$B_t = \eta_t \cdot \left(P_t^0, P_t^1, P_t^2, \ldots\right)$$

for all $t \ge 0$.

Proof. Let $\eta_t = B_t \delta^t$, where δ^t is the portfolio of holding exactly one bond of maturity t. That is to say, η is the strategy of investing all of the accumulated wealth in the bond of maturity t during the interval $(t-1, t]$. Note that η is previsible and that

$$\begin{aligned}
\eta_{t+1} \cdot \left(P_t^0, P_t^1, P_t^2, \ldots\right) &= B_{t+1} \delta^{t+1} \cdot \left(P_t^0, P_t^1, P_t^2, \ldots\right) \\
&= B_t \left(1 + r_{t+1}\right) P_t^{t+1} \\
&= B_t \\
&= B_t \delta^t \cdot \left(P_t^0, P_t^1, P_t^2, \ldots\right) \\
&= \eta_t \cdot \left(P_t^0, P_t^1, P_t^2, \ldots\right)
\end{aligned}$$

□

Definition 3.8.5. An equivalent martingale measure with respect to the bank account is called a risk-neutral measure.

Now we consider a natural application of a risk-neutral measure.

Definition 3.8.6. A futures contract initiated at time t with maturity date T is a contract where, at time t, no money is exchanged; at every time $t < u < T$, the sum of f_u^T units of money is swapped for f_{u-1}^T units of money; and finally at time T, the asset with price S_T is swapped for f_{T-1}^T units of money, where f_s^T is known at time s for $t \leq s \leq T$. The quantity f_t^T is called the futures price of the asset.

A futures contract is essentially a portfolio of forward contracts of different maturities, with the payout of each forward dependent on the futures prices at the different times before maturity.

Proposition 3.8.4. *Consider a market model with a bank account and futures contracts for maturity T initiated at each time u for $t \leq u < T$. If the market has no arbitrage, then*

$$f_t^T = \mathbb{E}^{\mathbb{Q}}\left[S_T \mid \mathcal{F}_t\right]$$

where $\mathbb{Q}$ is a risk-neutral measure.

Proof. By an extension of the 1st FTAP a market (with a dividend paying asset) is free of arbitrage if and only if there exists a risk-neutral measure $\mathbb{Q}$, i.e. a measure under which the process M defined by

$$M_t = \frac{P_t}{B_t} + \sum_{s=1}^{t} \frac{\delta_s}{B_s}$$

is a martingale. In our example, we set $\delta_u = f_u^T - f_{u-1}^T$ and $P_u = 0$ for $t \leq u \leq T$, where $f_T^T = S_T$. Note

$$\begin{aligned} 0 = \mathbb{E}^{\mathbb{Q}}\left[M_{u+1} - M_u \mid \mathcal{F}_u\right] &= \mathbb{E}^{\mathbb{Q}}\left[\left(f_{u+1}^T - f_u^T\right)/B_{u+1} \mid \mathcal{F}_u\right] \\ &= \left(\mathbb{E}^{\mathbb{Q}}\left[f_{u+1}^T \mid \mathcal{F}_u\right] - f_u^T\right)/B_{u+1} \end{aligned}$$

by the $\mathcal{F}_u$-measurability of B_{u+1}. Hence $\left(f_u^T\right)_{t \leq u \leq T}$ is a $\mathbb{Q}$-martingale. □

Remark 3.8.2. Let f_t^T be the futures price of a stock at time t for maturity T. If $T \mapsto f_t^T$ is non-decreasing, the market for that stock is said to be in contango at time t. And if $T \mapsto f_t^T$ is non-increasing, the market is said to be in normal backwardation.

Forward measures and risk-neutral measures are in general different. In particular, forward prices and futures prices usually disagree. But there is an important example where they agree:

Proposition 3.8.5. *Suppose that the spot interest rate process is not random. A probability measure is a T-forward measure if and only if it is risk-neutral.*

Proof. The fundamental result is that if B is a non-random process then

$$P_t^T = \frac{B_t}{B_T}.$$

Assume that the market has no arbitrage. By the first FTAP both YP^T and YB are martingales and thus we have

$$\begin{aligned} B_t &= \frac{1}{Y_t}\mathbb{E}\left[B_T Y_T \mid \mathcal{F}_t\right] \\ &= \frac{B_T}{Y_t}\mathbb{E}\left[P_T^T Y_T \mid \mathcal{F}_t\right] \\ &= B_T P_t^T \end{aligned}$$

Now, since B_T is assumed to be a constant, the process $(S_t/B_t)_{0\leq t\leq T}$ is a martingale with respect to a certain measure if and only if $\left(S_t/P_t^T\right)_{0\leq t\leq T}$ is a martingale with respect to the same measure. □

3.9 Contingent claim pricing and hedging

The setting of this section is as follows. We find ourselves in a market with prices $(P_t)_{t\geq 0}$. A contingent claim is any cash payment where the size of the payment is contingent on the prices of other assets or any other variable. There are two major types of contingent claims that we will study in these notes: European and American.

- European: specified by a time horizon $T > 0$ and $\mathcal{F}_T$-measurable random variable ξ_T modelling the payout at the maturity date T.
- American: specified by a time horizon $T > 0$ and an adapted process $(\xi_t)_{0\leq t\leq T}$ where ξ_t models the payout of the claim if the owner of the claim chooses to exercise at time t.

We put ourselves in the shoes of an investment bank that would like to market a new contingent claim. The question are these: what is a 'good' initial price for this claim? How can the seller hedge against the liability of owing the buyer the payout of the claim? We first consider European options.

Example 3.9.1 (Forward contract)**.** Given a market for a traded asset with prices $(S_t)_{t\geq 0}$, a forward contract initiated at a fixed time t for maturity T is a European claim with payout $\xi_T = S_T - F_t^T$, where F_t^T is the forward price at time t for maturity T. We have discussed this example in the previous section.

Example 3.9.2 (Call option)**.** A European call option gives the owner of the option the right, but not the obligation, to buy a given stock at a fixed time T (called the maturity date) at some fixed price K (called the strike). There are two cases: If $K \geq S_T$, then the option is worthless to the owner since there is no point paying a price above the market price for the underlying stock. On the other hand, if $K < S_T$, then the owner of the option can buy the stock for the price K from the counterparty and immediately sell the stock for the price S_T to the market, realising a profit of $S_T - K$. Hence, the payout of the call option is $\xi_T = (S_T - K)^+$, where $a^+ = \max\{a, 0\}$ as usual. The 'hockey-stick' graph of the function $g(x) = (x - K)^+$is below.

We will assume that the original market has no arbitrage, since otherwise it is difficult to formulate a reasonable answer to the pricing and hedging questions. Therefore, we will assume that there is at least one martingale deflator.

Proposition 3.9.1. *Consider an arbitrage-free market with prices P. Introduce to this market a European contingent claim with maturity T and payout ξ_T. Suppose that for $t < T$, the price of the claim at time t is ξ_t. If the augmented market with prices $(P_t, \xi_t)_{0 \leq t \leq T}$ has no arbitrage, then*

$$\underline{\xi}_t \leq \xi_t \leq \bar{\xi}_t \textit{ for all } 0 \leq t \leq T$$

where

$$\underline{\xi}_t = \operatorname{ess\,inf} \left\{ \frac{1}{Y_t} \mathbb{E} \left[Y_T \xi_T \mid \mathcal{F}_t \right] : Y \;\; \textit{a martingale deflator for the original market} \right.$$
$$\left. \textit{such that} \;\; \xi_T Y_T \;\; \textit{is integrable} \right\}$$

and $\bar{\xi}_t$ is defined similarly in terms of the essential supremum.

Proof. This is just the 1st FTAP. Indeed, if the augmented market is arbitrage free there must be a martingale deflator for the augmented market. Such a martingale deflator is necessarily a martingale deflator for the original market. In particular, if Y is a martingale deflator for the augmented market, then $\xi_t = \frac{1}{Y_t} \mathbb{E} \left(Y_T \xi_T \mid \mathcal{F}_t \right)$. That ξ_t is in the interval $\left[\underline{\xi}_t, \bar{\xi}_t \right]$ follows directly from the definition of essential infimum and supremum. □

Remark 3.9.1. We pause briefly to discuss the notions of essential supremum and essential infimum. Given a collection of random variables $(X_k)_{k \in K}$ indexed by an arbitrary (possibly uncountable) set K, a random variable Y is called the essential supremum of the collection, denoted

$$Y = \operatorname*{ess\,sup}_{k \in K} X_k$$

iff it satisfies

- $Y \geq X_k$ almost surely for all $k \in K$, and

- if another random variable Z is such that $Z \geq X_k$ almost surely for all $k \in K$, then $Z \geq Y$ almost surely.

The essential infimum of the collection is defined similarly. The proof of the existence of the essential supremum of a family of random variables is on example sheet 2 .

Recall, that in contrast, the function $\hat{Y}$ defined by $\hat{Y}(\omega) = \sup_{k \in K} X_k(\omega)$ has the property that $\hat{Y} \geq X_k$ for all $k \in K$ everywhere, and if Z is another function such that $Z \geq X_k$ for all $k \in K$ everywhere, then $Z \geq \hat{Y}$ everywhere.
Here is an example to show that the ordinary notion of supremum is not the correct notion in certain probabilistic settings. Let the set of outcomes Ω be the interval $[0,1]$, the set of events $\mathcal{F}$ be the Borel σ-algebra and the probability measures $\mathbb{P}$ be the Lebesgue measure. Fix a subset $K \subseteq [0,1]$ and let $X_k = \mathbb{1}_{\{k\}}$. Note that $\hat{Y} = \sup_k X_k = \mathbb{1}_K$. There are a couple of reasons why $\hat{Y}$ is not very useful probabilistically.
Firstly, note that $\hat{Y}$ is a measurable map from Ω to $\mathbb{R}$ if and only if K is a Borel set. Since K was arbitrary, it is not always the case that $\hat{Y}$ is a random variable.
Secondly, even if K is measurable and $\hat{Y}$ is a random variable, it is 'too big'. Indeed, the smaller random variable $Y = 0$ has the property that $Y \geq X_k$ almost surely for all $k \in K$. Returning to our application for bounding no-arbitrage prices, we don't know a priori whether the set of martingale deflators is countable or not. So the ordinary supremum may not be measurable. Furthermore, we only care about almost sure inequality for each martingale deflator (not inequality simultaneously for all martingale deflators for every outcome), the essential supremum and infimum appearing the statement of the result are appropriate. The above theorem says that the principle of no-arbitrage usually is not enough to uniquely price a contingent claim. At best, it gives an interval where the no-arbitrage price may lie. However, there is a special class of contingent claims that can be priced uniquely.

Definition 3.9.1. A European contingent claim with payout ξ_T is replicable or attainable iff there exists an initial wealth x and pure investment strategy H such that $X_T^{x,H} = \xi_T$ almost surely. (Recall that $X_0 = x, X_t = H_t \cdot P_t$ for $t \geq 1$ and pure investment means $H_{t+1} \cdot P_t = X_t$ for $t \geq 0$.

One of the reasons to single out attainable claims is that there is an unambiguous way to price them according to the no-arbitrage principle:

Theorem 3.9.2 (Characterisation of attainable claims)**.** *Suppose that the market model with n-dimensional price process P has no arbitrage. Let ξ_T be the payout of a European contingent claim with maturity date $T > 0$. The following are equivalent:*

1. *The claim is atttainable.*

2. *There exists a unique process $(\xi_t)_{0 \leq t \leq T}$ such that the augmented market (P, ξ) has no arbitrage.*

3. *There exists a number ξ_0 such that $\mathbb{E}\left[Y_T\xi_T\right] = Y_0\xi_0$ for all martingale deflators (of the original market) such that $Y_T\xi_T$ is integrable.*

Proof. (1) $\Rightarrow$ (2) This is the law of one price from the ??? example sheet. Indeed, let ξ_t be the price of the claim and X_t be the price of the replicating portfolio at time t. Let τ be the first time $\xi_t \neq X_t$. One the event $\{\tau < T\}$ do the following: at time τ buy the cheaper one, sell the expensive one, and consume the difference; and at time T unwind both positions for zero cost since $\xi_T = X_T$ by assumption. In notation, let H be the replicating strategy and let consider the strategy $(\tilde{H}, h)$ in the augmented market given by

$$\left(\tilde{H}_t, h_t\right) = \mathbb{1}_{\{\tau \leq t-1\}}\left(H_t, -1\right)\operatorname{sign}\left(\xi_\tau - X_\tau\right) \text{ for } 1 \leq t \leq T$$

and $\tilde{H}_{T+1} = 0$ and $h_{T+1} = 0$. The corresponding consumption stream is

$$c_t = \mathbb{1}_{\{\tau = t\}}\left|X_t - \xi_t\right|$$

This strategy would be an arbitrage unless $\mathbb{P}(\tau < T) = 0$. Hence, no arbitrage in the augmented market implies $\xi_t = X_t$ a.s. for all $0 \leq t \leq T$ and hence the price process is uniquely determined by the replicating strategy. (By the way, the same argument shows that if $X_T^{x,H} = X_T^{y,K}$ almost surely for possibly different strategies H and K initial wealths x and y, then $X_t^{x,H} = X_t^{y,K}$ almost surely for all $0 \leq t \leq T$. In particular, two replicating strategies of an attainable claim yield the same no arbitrage price.)
(2) $\Rightarrow$ (3) Suppose that there is no arbitrage in the augmented market. Then there exists a martingale deflator Y for that market. In particular, this martingale deflator is a martingale deflator for the original market and also for the new asset. That is, $Y\xi$ is a martingale and in particular, $\mathbb{E}\left[Y_T\xi_T\right] = Y_0\xi_0$. Now, assuming that the initial price ξ_0 of the claim is uniquely identified yields the result. ☐

It remains to prove the implication $(3) \Rightarrow (1)$. We will only prove the $T = 1$ version. Recall that in this setting, the martingale deflator: $Y_0 > 0, Y_1 > 0$ almost surely and $\mathbb{E}\left[Y_1P_1\right] = Y_0P_0$, can be replaced by the pricing kernel: $\rho > 0$ almost surely and $\mathbb{E}\left[\rho P_1\right] = P_0$, by setting $\rho = Y_1/Y_0$. In what follows, we will assume that the one-period market with prices P_0, P_1 has no arbitrage. By the fundamental theorem of asset pricing, there exists a pricing kernel. But, given an arbitrary random variable ξ_1, one may worry whether there must exist a pricing kernel ρ such that $\rho\xi_1$ is integrable. Fortunately, it turns out that if there is no arbitrage, there does indeed exist at least one pricing kernel ρ such that $\rho\xi_1$ is integrable. We will proceed by a series of lemmas.

Lemma 3.9.1. *Suppose that there exists a number ξ_0 such that $\mathbb{E}\left[\rho\xi_1\right] < \xi_0$ for all pricing kernels such that $\rho\xi_1$ is integrable. Then there exists a portfolio $H \in \mathbb{R}^n$ such that*

$$H \cdot P_0 \leq \xi_0 \text{ and } H \cdot P_1 \geq \xi_1 \text{ a.s.}$$

and there is positive probability that at least one of the above inequalities is strict.

Proof. By assumption, there does not exist a pricing kernel for the augmented market (P, ξ). By the fundamental theorem of asset pricing, there exists an arbitrage in the augmented market, i.e. a portfolio $(\tilde{H}, h) \in \mathbb{R}^{n+1}$ such that

$$\tilde{H} \cdot P_0 + h\xi_0 \leq 0 \text{ and } \tilde{H} \cdot P_1 + h\xi_1 \geq 0 \text{ a.s.}$$

where there is positive probability that at least one of the above inequalities is strict. Let ρ be a pricing kernel for the original market such that ρP_1 is integrable. Note that we have

$$\begin{aligned} 0 &\leq \mathbb{E}\left[\rho\left(\tilde{H} \cdot P_1 + h\xi_1\right)\right] \\ &= \tilde{H} \cdot P_0 + h\mathbb{E}\left[\rho\xi_1\right] \\ &\leq h\left[\mathbb{E}\left[\rho\xi_1\right] - \xi_0\right] \end{aligned}$$

Since $\mathbb{E}\left[\rho\xi_1\right] - \xi_0 < 0$, we conclude that $h \leq 0$. We can rule out the case that $h = 0$. For instance, if $h = 0$ we would have $\tilde{H}$ being an arbitrage in the original market - a contradiction. [6] We are left with $h < 0$. This shows that $H = -\tilde{H}/h$ satisfies the conclusion of the lemma. □

Lemma 3.9.2. *Suppose that there exists a number ξ_0 such that $\mathbb{E}\left[\rho\xi_1\right] \leq \xi_0$ for all pricing kernels such that $\rho\xi_1$ is integrable. Then there exists a portfolio $H \in \mathbb{R}^n$ such that*

$$H \cdot P_0 \leq \xi_0 \quad \textit{and} \quad H \cdot P_1 \geq \xi_1 \textit{ a.s.}$$

Proof. For all $k > 0$, we have $\mathbb{E}\left[\rho\xi_1\right] < \xi_0 + 1/k$ and hence by the previous lemma there exists a portfolio $H_k \in \mathbb{R}^n$ such that $H_k \cdot P_0 \leq \xi_0 + 1/k$ and $H_k \cdot P_1 \geq \xi_1$ a.s. We consider two cases now:

- Case: $(H_k)_k$ is bounded. In this case, we can pass to a convergent subsequence such that $H_k \to H^*$. Note
$$H^* \cdot P_0 \leq \xi_0 \quad \text{and} \quad H^* \cdot P_1 \geq \xi_1 \text{ a.s.}$$
as desired.

- Case: $(H_k)_k$ is unbounded. Recall from the proof of the 1st FTAP the notation
$$\mathcal{U} = \{u \in \mathbb{R}^n : u \cdot P_0 = 0 = u \cdot P_1 \text{ a.s. }\}$$
and
$$\mathcal{V} = \mathcal{U}^\perp$$
By projecting our given sequence onto $\mathcal{V}$, we can assume that $H_k \in \mathcal{V}$ for all k and that $(H_k)_k$ is unbounded (otherwise, we are back to the previous case). We can pass to a subsequence such that $\|H_k\| \to \infty$ and to a further subsequence such that $\hat{H}_k = H_k / \|H_k\|$ converges to a non-zero limit $\hat{H} \in \mathcal{V}$. Now dividing the inequalities by $\|H_k\|$ and taking the limit yields
$$\hat{H} \cdot P_0 \leq 0 \leq \hat{H} \cdot P_1 \text{ a.s.}$$

By no arbitrage, we have $\hat{H} \cdot P_0 = 0 = \hat{H} \cdot P_1$ a.s., or in other notation $\hat{H} \in \mathcal{U}$. Since $\hat{H}$ is in $\mathcal{V}$, we have $\hat{H} = 0$, a contradiction. This shows that this second case is impossible.

□

Proof of (3) ⇒ (1) in one period. Suppose that there exists a number ξ_0 such that $\mathbb{E}\left[\rho\xi_1\right] = \xi_0$ for all pricing kernels such that $\rho\xi_1$ is integrable. Note that

$$\mathbb{E}\left[\rho\xi_1\right] \leq \xi_0 \quad \text{for all } \rho$$

so that there exists a portfolio $H^+ \in \mathbb{R}^n$ such that

$$H^+ \cdot P_0 \leq \xi_0 \quad \text{and} \quad H^+ \cdot P_1 \geq \xi_1 \text{ a.s.}$$

Similarly,

$$\mathbb{E}\left[\rho\left(-\xi_1\right)\right] \leq -\xi_0 \quad \text{for all } \rho$$

so that there exists a portfolio $H^- \in \mathbb{R}^n$ such that

$$H^- \cdot P_0 \leq -\xi_0 \quad \text{and} \quad H^- \cdot P_1 \geq -\xi_1 \text{ a.s.}$$

Adding this together yields

$$\left(H^+ + H^-\right) \cdot P_0 \leq 0 \leq \left(H^+ + H^-\right) \cdot P_1 \text{ a.s.}$$

By no arbitrage in the original market, we have

$$\left(H^+ + H^-\right) \cdot P_0 = 0 = \left(H^+ + H^-\right) \cdot P_1 \text{ a.s.}$$

Hence the portfolios $H = H^+$and $H = -H^-$satisfy the desired conclusion. □

Example 3.9.3 (Put-call parity formula)**.** Suppose we start with a market with three assets with prices $\left(B_t^T, S_t, C_t\right)_{0 \leq t \leq T}$. The first asset is a bond with maturity date T and unit principal value, so that in particular, $B_T^T = 1$ almost surely. The next asset is a stock. The last asset is a call option on that stock with strike K and maturity T, so that $C_T = (S_T - K)^+$. Suppose that this market is free of arbitrage.

Now we introduce another claim, called a put option. A put option gives the owner of the option the right, but not the obligation, to sell the stock for a fixed strike price at a fixed maturity date. If the strike is K and maturity date is T, then a similar argument as we used for the call option, the payout of a put option is $P_T = (K - S_T)^+$.

It turns out that the put option is replicable in the market $\left(B^T, S, C\right)$. Indeed, we have the identity

$$\begin{aligned}
P_T &= (K - S_T)^+ \\
&= K - S_T + (S_T - K)^+ \\
&= (K, -1, +1) \cdot \left(B_T^T, S_T, C_T\right).
\end{aligned}$$

Hence $H_t = (K, -1, +1)$ for all $1 \le t \le T$ is a replicating strategy. Now, suppose we want to assign prices P_t to the put for $0 \le t < T$. The above theorem says there is no arbitrage in the augmented market $\left(B^T, S, C, P\right)$ if and only if

$$P_t - C_t = KB_t^T - S_t$$

This is the famous put-call parity formula.

Since attainable claims have unique no-arbitrage prices, we single out the markets for which every claim is attainable:

Definition 3.9.2. A market is complete if and only if every European contingent claim is attainable. A market is incomplete otherwise.

In discrete time models complete, markets have a lot (probably too much) structure:

Theorem 3.9.3. *If the market model P with n assets is complete, then for each $t \ge 0$ the probability space Ω can be partitioned into no more than n^t $\mathcal{F}_t$-measurable events of positive probability, and in particular, the n-dimensional random vector P_t takes values in a set of at most n^t elements.*

Proof. We will proceed by induction. First suppose $A_1, \ldots, A_k$ are a collection of disjoint $\mathcal{F}_t$-measurable events with $\mathbb{P}(A_i) > 0$ for all i.

Claim: the set $\{\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_k}\}$ is linearly independent, and in particular, the dimension of the span of $\{\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_k}\}$ is exactly k.

Indeed, we must show that if

$$a_1\mathbb{1}_{A_1} + \ldots a_k\mathbb{1}_{A_k} = 0 \text{ a.s.}$$

for some constants $a_1, \ldots, a_k$, then $a_1 = \cdots = a_k = 0$. To this end, note that if $i \ne j$ the sets A_i and A_j are disjoint and hence $\mathbb{1}_{A_i}\mathbb{1}_{A_j} = 0$. By multiplying both sides of the equation by $\mathbb{1}_{A_i}$ we get $a_i\mathbb{1}_{A_i} = 0$. But since $\mathbb{P}(A_i) > 0$ it must be the case that $a_i = 0$, proving the claim.
Since the market is complete, each of the $\mathbb{1}_{A_i}$ is replicable. Hence

$$\operatorname{span}\{\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_k}\} \subseteq \{H_t \cdot P_t : H_t \text{ is } \mathcal{F}_{t-1}\text{-meas.}\}$$

Now let $B_1, \ldots, B_N$ be a maximal partition of Ω into disjoint $\mathcal{F}_{t-1}$-measurable sets of positive measure, where by the induction hypothesis $N \le n^{t-1}$. If a random vector H_t is $\mathcal{F}_{t-1}$-measurable, then it takes exactly one value on each of the B_j 's for a total of at most N values $h_1, \ldots, h_N$. Hence

$$\begin{aligned}\{H_t \cdot P_t : H_t \text{ is } \mathcal{F}_{t-1}\text{-meas. }\} &= \{h_1 \cdot P_t\mathbb{1}_{B_1} + \ldots + h_N \cdot P_t\mathbb{1}_{B_N} : h_1, \ldots, h_N \in \mathbb{R}^n\} \\ &= \operatorname{span}\left\{P_t^i\mathbb{1}_{B_j} : 1 \le i \le n, 1 \le j \le N\right\}\end{aligned}$$

and the dimension of the space above is at most nN. Therefore, we have shown $k \le nN \le n^t$, completing the induction. $\square$

We can characterise complete markets:

Theorem 3.9.4 (Second Fundamental Theorem of Asset Pricing)**.** *An arbitrage-free market model is complete if and only if there exists a unique martingale deflator Y such that $Y_0 = 1$.*

Proof. ('if' direction) Let Y and Y' be martingale deflators with $Y_0 = 1 = Y'_0$. Suppose the market is complete, fix a non-random time $T > 0$ and consider the claim with payout $\xi_T = Y_T - Y'_T$. By completeness, there exists (x, H) such that $X_T^{x,H} = \xi_T$. By completeness, every $\mathcal{F}_T$-measurable random variable is bounded (since it can take at most n^T different values) so both $Y_T \xi_T$ and $Y'_T \xi_T$ are integrable. In particular, we have

$$\mathbb{E}\left[Y_T \xi_T\right] = x = \mathbb{E}\left[Y'_T \xi_T\right].$$

Subtracting the left and right side of the above equation and using $\xi_T = Y_T - Y'_T$ yields

$$\mathbb{E}\left[\left(Y_T - Y'_T\right)^2\right] = 0$$

from which the uniqueness of the martingale deflator follows.
('only if' direction) Suppose there is a unique martingale deflator such that $Y_0 = 1$. Then for every contingent claim with payout ξ_T there exists a unique number ξ_0 such that $\mathbb{E}\left[Y_T \xi_T\right] = \xi_0$ for every (that is, the unique) martingale deflator. By the characterisation of attainability, we have ξ_T is attainable, as desired. □

This box summarises the fundamental theorems:

1FTAP:	No arbitrage	⇔	Existence of martingale deflator
2FTAP:	Completeness + No arbitrage	⇔	Uniqueness of martingale deflator

Finally, we close this section with another useful consequence of completeness:

Proposition 3.9.2. *Suppose the arbitrage-free market model is complete. Then there exists a bank account.*

Proof. By completeness, bonds of all maturities can be replicated. Hence a bank account can be constructing by holding all the wealth during the period $(t-1, t]$ in the bond with maturity t. □

3.10 Replication with calls and puts

We consider a market consisting of a bond, a stock with time-T price $S_T \geq 0$, and a family of European calls and puts with strikes in a finite set $\mathcal{K} = \{K_1, \ldots, K_N\} \subseteq (0, \infty)$ all with maturity T.

Theorem 3.10.1. *Suppose g is piece-wise linear with kinks precisely at the points $\mathcal{K}$. Then the European claim with time T payout $\xi_T = g(S_T)$ is attainable.*

Proof. Note that g is differentiable everywhere except $\mathcal{K}$. For every $a \notin \mathcal{K}$, note that the following identity holds

$$g(s) = g(a) + g'(a)(s-a) + \sum_{K \in \mathcal{K}, K<a} \Delta_K (K-s)^+ + \sum_{K \in \mathcal{K}, K>a} \Delta_K (s-K)^+$$

where $\Delta_K = g'(K+) - g'(K-)$. Hence the replicating strategy is to hold $g(a) - ag'(a)$ shares of the bond, to hold $g'(a)$ shares of stock, and holding Δ_K puts of strike $K < a$ and Δ_K calls of strike $K > a$ for all $K \in \mathcal{K}$. □

Remark 3.10.2. To prove the identity, note that

$$g(x) = g(0) + g'(0)x + \sum_{K \in \mathcal{K}} \Delta_K (x-K)^+$$

for all $x \geq 0$ and

$$g'(x) = g'(0) + \sum_{K \in \mathcal{K}, K<x} \Delta_K$$

for $x \notin K$, so that

$$\begin{aligned} g(s) - g(a) - g'(a)(s-a) &= \sum_{K \in \mathcal{K}} \Delta_K \left[(s-K)^+ - (a-K)^+ - (s-a)\mathbb{1}_{\{K<a\}} \right] \\ &= \sum_{K \in \mathcal{K}, K<a} \Delta_K (K-s)^+ + \sum_{K \in \mathcal{K}, K>a} \Delta_K (s-K)^+. \end{aligned}$$

Remark 3.10.3. In fact, when g is twice continuously differentiable, the continuous analogue of the above identity holds

$$g(s) = g(a) + g'(a)(S_1 - a) + \int_0^a g''(K)(K-s)^+ dK + \int_a^\infty g''(K)(s-K)^+ dK$$

for any $a > 0$. Note that the integrand of the first integral is zero unless $\min\{s, a\} \leq K \leq a$. Similarly, the integrand of the second integral is zero unless $a \leq K \leq \max\{s, a\}$. In particular, the ranges of both integrals are bounded intervals on which g'' is assumed continuous, so both integrals are ordinary Riemann integrals. (One way to prove this identity is to fix s and let $h(a)$ equal the right-hand side. By the standard rules of calculus, we have $h'(a) = 0$ and hence $h(a)$ is a constant. To evaluate that constant, let $a = s$ and note that both integrals vanish since the ranges of integration have zero length.) Consider the case $g(s) = \log s$. We have the identity

$$\log S_T = \log a + \frac{S_T - a}{a} - \int_0^a \frac{(K - S_T)^+}{K^2} dK - \int_a^\infty \frac{(S_T - K)^+}{K^2} dK.$$

This formula is interpreted to mean that a claim with payout $\xi_T = \log S_T$ can be approximately replicated by trading in calls and puts over a large number of strikes.

Although in reality there do not exist contingent claims with log payouts, market practitioners often think of this log contract as being traded since it can be manufactured (approximately) by calls and puts in a straight-forward manner. This line of thinking has lead to the introduction of variance swap contracts which we will consider in the chapter on continuous time models.
The above result says that given enough call prices, it is possible to replicate any claim with payout of the form $g(S_T)$, assuming g is piece-wise linear (or at least to replicate approximately in the case where g is smooth enough to be approximated by a piece-wise linear function). Assuming we know the initial prices of theses call, we can then calculate the initial cost of the approximate hedging portfolio.

We now come to a simple result observed by Breeden and Litzenberger in 1978 . Our setting is a market with prices $\left(P^T, S\right)$ where the first asset is a bond of maturity T, let $\mathbb{Q} \sim \mathbb{P}$ be a T-forward measure, so that $\mathbb{E}^{\mathbb{Q}}\left[S_T\right] = S_0/P_0^T$. Let $C_T^{T,K} = (S_T - K)^+$be the payout of a call with strike $K \geq 0$, and let the initial prices be

$$C_0^{T,K} = P_0^T \mathbb{E}^{\mathbb{Q}}\left[(S_T - K)^+\right].$$

Then for any collection of strikes $\mathcal{K} = \{K_1, \ldots, K_N\}$ the augmented market with prices

$$\left(P^T, S, C^{T,K}\right)_{K \in \mathcal{K}}$$

has no arbitrage by the easy direction of the first fundamental theorem, since $\mathbb{Q}$ is a forward measure for the augmented market.
Note that the function $K \mapsto C_0^{T,K}$ is convex and therefore has right- and left- derivatives at each point. The following gives meaning to these derivatives:

Proposition 3.10.1 (Breeden-Litzenberger formula)**.** *For any $K \geq 0$ we have*

$$\mathbb{Q}\left(S_T > K\right) = \frac{1}{P_0^T} D_K^+ C_0^{T,K}$$

and

$$\mathbb{Q}\left(S_T \geq K\right) = \frac{1}{P_0^T} D_K^- C_0^{T,K}.$$

If $K \mapsto C_0^{T,K}$ is twice-differentiable then the law of the random variable S_T has a density f_{S_T} under $\mathbb{Q}$ given by

$$f_{S_T}(K) = \frac{1}{P_0^T} D_K^2 C_0^{T,K}.$$

Proof. Note that

$$\begin{aligned} D^+ C_0(K) &= \lim_{\varepsilon \downarrow 0} \frac{C_0(K+\varepsilon) - C_0(K)}{\varepsilon} \\ &= -P_0^T \lim_{\varepsilon \downarrow 0} \mathbb{E}^{\mathbb{Q}}\left[g_\varepsilon\left(S_T - K\right)\right] \end{aligned}$$

where

$$g_\varepsilon(x) = \frac{x}{\varepsilon}\mathbb{1}_{[0,\varepsilon)} + \mathbb{1}_{[\varepsilon,\infty)}(x).$$

Note that g_ε is bounded and $g_\varepsilon \to \mathbb{1}_{(0,\infty)}$ pointwise, so the first formula is proven by by the dominated convergence theorem. The formula for the left-derivative is proven similarly. Finally, if C_0 is twice-differentiable the density is recovered by differentiating once more with respect to K. □

3.11 Call prices from moment generating functions

Since a portfolio of calls and puts on a stock can essentially replicate any European contingent claim, it is important to have models where the call prices can be computed easily. Unfortunately, there are few models where there exists nice, elementary formulae for the call prices. However, there are many models (especially when we get to continuous time) where the moment generating functions can be computed explicitly, and we will now see that given the moment generating function we can compute call prices by integration:

Consider a market model with a bond of maturity T, a stock with $S_T \geq 0$ almost surely, and let $\mathbb{Q}$ be a fixed T-forward measure. Let

$$C_0^{T,K} = P_0^T \mathbb{E}^{\mathbb{Q}}\left[(S_T - K)^+\right]$$

for $K > 0$, so that if $C_0^{T,K}$ is the initial price of a call with strike K, then the augmented market has no arbitrage. For complex θ in the vertical strip

$$\Theta = \{\theta = p + \mathrm{i}q : 0 \leq p \leq 1, q \in \mathbb{R}\}$$

define the moment generating function of the log stock price by

$$M(\theta) = \mathbb{E}^{\mathbb{Q}}\left[e^{\theta \log S_T}\mathbb{1}_{\{S_T > 0\}}\right].$$

Note that we have for $\theta = p + \mathrm{i}q \in \Theta$,

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[\left|e^{\theta \log S_T \mathbb{1}_{\{S_T>0\}}}\right|\right] &= \mathbb{E}^{\mathbb{Q}}\left[S_T^p\right] \\
&\leq \mathbb{E}^{\mathbb{Q}}\left[S_T\right]^p \\
&= \left[\frac{S_0}{P_0^T}\right]^p < \infty
\end{aligned}$$

by Jensen's inequality, so the moment generating function is well-defined and finite-valued – indeed it is analytic on Θ but we do not use this fact below. The following result shows how to recover call prices from the moment generating function.

Theorem 3.11.1. *For any* $0 < p < 1$ *the identity*

$$C_0^{T,K} = S_0 - \frac{K^{1-p} P_0^T}{2\pi} \int_{-\infty}^{\infty} \frac{M(p+ix)e^{-\mathrm{i}x\log K}}{(x-\mathrm{i}p)(x+\mathrm{i}(1-p))} dx$$

holds.

Essentially, we are inverting the moment generating function via a complex integral. Variants of this procedure are often called a Bromwich, Fourier or Mellin transform. To prove this formula, we begin with a lemma:

Lemma 3.11.1. *For any* $0 < p < 1$ *the identity*

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{\mathrm{i}ax}}{(x-\mathrm{i}p)(x+\mathrm{i}(1-p))} dx = \begin{cases} e^{-ap} & \text{if } a \geq 0 \\ e^{a(1-p)} & \text{if } a < 0 \end{cases}$$

holds.

Proof. This is a standard application of the Cauchy residue theorem. Consider the case $a \geq 0$. Define the semi-circular contour

$$\Gamma_R = \{x + \mathrm{i}0 : -R \leq x \leq R\} \cup \left\{ Re^{\mathrm{i}\phi} : 0 \leq \phi \leq \pi \right\}$$

in the upper half-plane. Cauchy's theorem gives

$$\begin{aligned} \int_{\Gamma_R} \frac{e^{\mathrm{i}az}}{(z-\mathrm{i}p)(z+\mathrm{i}(1-p))} dz &= \mathrm{i}2\pi \frac{e^{\mathrm{i}az}}{z+\mathrm{i}(1-p)}\bigg|_{z=\mathrm{i}p} \\ &= 2\pi e^{-ap} \end{aligned}$$

since the integrand is meromorphic with a simple pole at $z = \mathrm{i}p$ inside the contour, and the contour integral is evaluated in the anticlockwise sense. On the other hand,

$$\begin{aligned} &\int_{\Gamma_R} \frac{e^{\mathrm{i}az}}{(z-\mathrm{i}p)(z+\mathrm{i}(1-p))} dz \\ &= \int_{-R}^{R} \frac{e^{\mathrm{i}ax}}{(x-\mathrm{i}p)(x+\mathrm{i}(1-p))} dx + \int_0^{\pi} \frac{\mathrm{i}Re^{-aR\sin\phi} e^{\mathrm{i}(aR\cos\phi + \phi)}}{\left(Re^{i\phi} - \mathrm{i}p\right)\left(Re^{\mathrm{i}\phi} + \mathrm{i}(1-p)\right)} d\phi \end{aligned}$$

and the second integral vanishes as $R \to \infty$ since $a \geq 0$. The case $a < 0$ is handled in exactly the same way; just integrate around a semi-circular contour in the lower half-plane enclosing the other pole at $-\mathrm{i}(1-p)$. □

Proof of the theorem. From the lemma, we have the identity

$$
\begin{aligned}
(S_T - K)\,\mathbb{1}_{\{S_T \geq K\}} &+ (S_T - S_T)\mathbb{1}_{\{S_T < K\}} \\
&= S_T - \frac{K^{1-p}}{2\pi} \int_{-\infty}^{\infty} \frac{e^{p \log S_T + \mathrm{i}x \log(S_T/K)} \mathbb{1}_{\{S_T > 0\}}}{(x - \mathrm{i}p)(x + \mathrm{i}(1-p))} dx.
\end{aligned}
$$

Now multiply by P_0^T and compute expectations. The result follows upon interchanging expectation and integration on the right-hand side. This is justified by Fubini's theorem since

$$
\int_{-\infty}^{\infty} \mathbb{E}^{\mathbb{Q}} \left| \frac{e^{p \log S_T + \mathrm{i}x \log(S_T/K)}}{(x - \mathrm{i}p)(x + \mathrm{i}(1-p))} \right| dx = M(p) \int_{-\infty}^{\infty} \frac{dx}{\sqrt{(x^2 + p^2)\,(x^2 + (1-p)^2)}} < \infty.
$$

□

3.12 Super-replication of American claims

We now discuss American claims. Here, things are quite different. The canonical example of an American claim is the American put option – a contract which gives the buyer the right (but not the obligation) to sell the underlying stock at a fixed strike price $K > 0$ at any time between time 0 and a fixed maturity date T. Hence, the payout of the option is $(K - S_\tau)^+$ where $\tau \in \{0, \ldots, T\}$ is a time chosen by the holder of the put to exercise the option. The payout of an American claim is specified by two ingredients:

- a maturity date $T > 0$,
- an adapted process $(\xi_t)_{0 \leq t \leq T}$.

For instance, in the case of an American put, we take $\xi_t = (K - S_t)^+$. Unlike the European claim, the holder of an American claim can choose to exercise the option at any time τ before or at maturity. However we insist that τ is a stopping time.

Now, if an American claim matures at $T > 0$ and is specified by the payout process $(\xi_t)_{0 \leq t \leq T}$, then the actual payout of the claim is modelled by the random variable ξ_τ, where τ is any stopping time for the filtration taking values in $\{0, \ldots, T\}$.
We can think of the American claim then as a family, indexed by the stopping time τ, of European claims with payouts ξ_τ. To simplify matters, we make the following assumption in this subsection:

The market model $P = (P_t)_{0 \leq t \leq T}$ is complete.

Let $Y = (Y_t)_{0 \leq t \leq T}$ be the unique martingale deflator such that $Y_0 = 1$. Intuitively, the seller of such a claim should at time 0 charge at least the amount

$$
\sup_{\tau \leq T} \mathbb{E}\left[Y_\tau \xi_\tau\right]
$$

to be sure that he can hedge the option, where the supremum is taken over the set of stopping times smaller than or equal to T. Indeed, this is the case.

Theorem 3.12.1. *Suppose that the adapted process $(\xi_t)_{0\leq t\leq T}$ specifies the payout of an American claim maturing at $T > 0$. There exists a self-financing pure-investment trading strategy H such that*

- $X_t \geq \xi_t$ *for all* $0 \leq t \leq T$,
- $X_{\tau^*} = \xi_{\tau^*}$ *for some stopping time* τ^*, *and*
- $X_0 = \sup_{\tau\leq T} \mathbb{E}\left[Y_\tau \xi_\tau\right]$,

where $X_t = H_t \cdot P_t = H_{t+1} \cdot P_t$.

Remark 3.12.2. The strategy H dominates the payout of the American claim at all times, but is conservative in the sense that it exactly replicates the optimally exercised claim.

The rest of this subsection is dedicated to proving this theorem. We will need a result of general interest:

Theorem 3.12.3 (Doob decomposition theorem)**.** *Let U be a discrete-time supermartingale. Then there is a unique decomposition*

$$U_t = U_0 + M_t - A_t$$

where M is a martingale and A is a predictable non-decreasing process with $M_0 = A_0 = 0$.

Proof. Let $M_0 = 0 = A_0$ and define

$$\begin{aligned} M_{t+1} &= M_t + U_{t+1} - \mathbb{E}\left[U_{t+1} \mid \mathcal{F}_t\right] \\ A_{t+1} &= A_t + U_t - \mathbb{E}\left[U_{t+1} \mid \mathcal{F}_t\right] \end{aligned}$$

for $t \geq 0$. Since U is assumed to be supermartingale, and hence integrable, the processes M and A are integrable. It is straightforward to check that M is a martingale, and since U is a supermartingale, that A is non-decreasing. Also by induction, we see that A_{t+1} is $\mathcal{F}_t$-measurable. Summing up,

$$\begin{aligned} M_t - A_t &= M_0 - A_0 + \sum_{s=1}^{t} \left(M_s - M_{s-1} - A_s + A_{s-1}\right) \\ &= \sum_{s=1}^{t} \left(U_s - U_{s-1}\right) \\ &= U_t - U_0 \end{aligned}$$

To show uniqueness, assume that $U_t = U_0 + M_t - A_t = U_0 + M'_t - A'_t$. Then $M - M'$ is a predictable discrete-time martingale, that is, a constant. □

Now we introduce the key concept in optimal stopping theory:

Definition 3.12.1. Let $(Z_t)_{0\le t\le T}$ be a given integrable adapted discrete-time process. Define an adapted process $(U_t)_{0\le t\le T}$ by the recursion

$$\begin{aligned} U_T &= Z_T \\ U_t &= \max\{Z_t, \mathbb{E}[U_{t+1} \mid \mathcal{F}_t]\} \qquad \text{for } 0 \le t \le T-1 \end{aligned}$$

The process $(U_t)_{0\le t\le T}$ is called the Snell envelope of $(Z_t)_{0\le t\le T}$.

Remark 3.12.4. The Snell envelope clearly satisfies both

$$U_t \ge Z_t \quad \text{and} \quad U_t \ge \mathbb{E}[U_{t+1} \mid \mathcal{F}_t]$$

almost surely. Thus, another way to describe the Snell envelope of a process is to say it is the smallest supermartingale dominating that process.

In our application Z will be the process $Y\xi$, where Y is the martingale deflator and ξ is the process specifying the payout of the American claim.

Theorem 3.12.5. *Let $(Z_t)_{0\le t\le T}$ be an integrable adapted process, let $(U_t)_{0\le t\le T}$ be its Snell envelope with Doob decomposition $U_t = U_0 + M_t - A_t$. Let $A_{T+1} = +\infty$ and*

$$\tau^* = \min\{t \in \{0, \dots, T\} : A_{t+1} > 0\}.$$

Then τ^* is a stopping time and

$$U_{\tau^*} = U_0 + M_{\tau^*} = Z_{\tau^*}.$$

Proof. That τ^* is a stopping time follows from the fact that the non-decreasing process $(A_t)_{0\le t\le T+1}$ is predictable. Now note that

$$\mathbb{E}[U_{t+1} \mid \mathcal{F}_t] = \mathbb{E}[U_0 + M_{t+1} - A_{t+1} \mid \mathcal{F}_t] = U_0 + M_t - A_{t+1}$$

since M is a martingale and A is predictable so that by the definition of Snell envelope

$$U_0 + M_t - A_t = \max\{Z_t, U_0 + M_t - A_{t+1}\}.$$

Note that $A_{\tau^*} = 0$ and hence

$$U_0 + M_{\tau^*} = \max\{Z_{\tau^*}, U_0 + M_{\tau^*} - A_{\tau^*+1}\}$$

But since $A_{\tau^*+1} > 0$ we must conclude

$$U_{\tau^*} = U_0 + M_{\tau^*} = Z_{\tau^*}.$$

□

Theorem 3.12.6. *Let Z be an adapted integrable process and let U be its Snell envelope. Then*

$$U_0 = \sup \left\{ \mathbb{E}\left[Z_\tau\right] : \textit{ stopping time } 0 \leq \tau \leq T \right\}.$$

Proof. Since U is a supermartingale,

$$U_0 \geq \mathbb{E}\left[U_\tau\right]$$

for any stopping time τ by the optional sampling theorem. But since $U_t \geq Z_t$ by construction,

$$U_0 \geq \mathbb{E}\left[Z_\tau\right]$$

for any stopping time τ. But letting $\tau^* = \min\left\{t \in \{0, \ldots, T\} : A_{t+1} > 0\right\}$ where $U = U_0 + M - A$ is the Doob decomposition of U, we have

$$U_0 = U_0 + \mathbb{E}\left[M_{\tau^*}\right] = \mathbb{E}\left[Z_{\tau^*}\right].$$

again by the optional sampling theorem and the previous result. $\square$

Remark 3.12.7. By a similar argument, one can show that

$$U_t = \operatorname{ess\,sup} \left\{ \mathbb{E}\left[Z_\tau \mid \mathcal{F}_t\right] : \text{stopping time } t \leq \tau \leq T \right\}.$$

for all $0 \leq t \leq T$. This formula allows us to define the Snell envelope for the infinite horizon case $T = \infty$ and also in the continuous time case.

Definition 3.12.2. If Z is an integrable adapted process, a stopping time σ such that $\mathbb{E}\left[Z_\sigma\right] = \sup_{0 \leq \tau \leq T} \mathbb{E}\left[Z_\tau\right]$ is called an optimal stopping time. Obviously the stopping time τ^* defined above is an optimal stopping time.

Returning to finance, let $(\xi_t)_{0 \leq t \leq T}$ be the process specifying the payout of an American option, let Y the unique martingale deflator with $Y_0 = 0$ and let $(U_t)_{0 \leq t \leq T}$ be the Snell envelope of $Y\xi$ with Doob decomposition $U = U_0 + M - A$.
We now will use the assumption that the market is complete: let H be strategy such that $H_T \cdot P_T = \left(U_0 + M_T\right)/Y_T$. Setting $X_t = H_t \cdot P_t$ note that XY is a martingale since it is a local martingale from before, and since the market is complete, it is also bounded. By the martingale property, we have

$$X_t Y_t = U_0 + M_t$$

for all $0 \leq t \leq T$. In particular,

- $X_t = \left(U_0 + M_t\right)/Y_t \geq U_t/Y_t \geq \xi_t$ for all $0 \leq t \leq T$,
- $X_{\tau^*} = \left(U_0 + M_{\tau^*}\right)/Y_{\tau^*} = U_{\tau^*}/Y_{\tau^*} = \xi_{\tau^*}$, and
- $X_0 = U_0 = \sup_{\tau \leq T} \mathbb{E}\left[Y_\tau \xi_\tau\right]$,

completing the proof of the theorem.

3.13 A dual approach to optimal stopping

The final result in discrete time is the following dual approach to optimal stopping:

Theorem 3.13.1. *Let $(Z_t)_{0\leq t\leq T}$ be a discrete-time, integrable adapted process. Then*

$$\sup_{\tau\leq T}\mathbb{E}\left[Z_\tau\right]=\inf_{M}\mathbb{E}\left[\max_{0\leq t\leq T}\left(Z_t-M_t\right)\right]$$

where the supremum on the left-hand side is taken over stopping times τ and the infimum on the right-hand side is taken over martingales M with $M_0=0$.

Proof. Let M be a martingale with $M_0=0$. By the optional stopping theorem we have

$$\begin{aligned}\sup_{\tau\leq T}\mathbb{E}\left[Z_\tau\right]&=\sup_{\tau\leq T}\mathbb{E}\left[Z_\tau-M_\tau\right]\\&\leq\mathbb{E}\left[\max_{0\leq t\leq T}\left(Z_t-M_t\right)\right]\end{aligned}$$

Since the inequality holds for any martingale M, it also holds when we take the infimum of the right-hand side over M. For the reverse inequality, let U be the Snell envelope of Z and let $U=U_0+M^*-A$ be its Doob decomposition. Since $U_t\geq Z_t$ for all t we have

$$\begin{aligned}\inf_{M}\mathbb{E}\left[\max_{0\leq t\leq T}\left(Z_t-M_t\right)\right]&\leq\mathbb{E}\left[\max_{0\leq t\leq T}\left(Z_t-M_t^*\right)\right]\\&\leq\mathbb{E}\left[\max_{0\leq t\leq T}\left(U_t-M_t^*\right)\right]\\&=\mathbb{E}\left[\max_{0\leq t\leq T}\left(U_0-A_t\right)\right]\\&=U_0\end{aligned}$$

The conclusion follows from the last section where we proved that

$$U_0=\sup_{\tau\leq T}\mathbb{E}\left[Z_\tau\right]$$

□