1 Practicalities and background

1.1 Practicalities

The final grading will be based on the final exam. All exercises are valid for a bonus system for the final exam. Upon passing the exam, solved exercises give a bonus to the final grade.
1.2 Background

We assume that the following concepts are familiar:

1. Probability space, random variables, expectation, convergence concepts.
2. Conditional expectations, martingales.
3. The fundamentals of discrete time financial mathematics.

2 Preliminaries

2.1 The market model, contingent claims and numeraire

We work in a discrete time with a finite number $T + 1$ of time points. We fix a filtered complete probability space $(\Omega, (\mathcal{F}_t)_{t=0}^T, \mathcal{F}, P)$, where we assume that $\mathcal{F}_T = \mathcal{F}$. The elements of $\Omega$ are denoted by $\omega$ and the space of $\mathcal{F}_t$-measurable random variables by $L^0(\mathcal{F}_t)$. When $t = T$, we denote simply $L^0 := L^0(\mathcal{F})$.

The market consists of a finite number $d$ of assets which are, e.g., stocks, bonds, commodities, or currencies. The initial prices at $t = 0$ are known, and their future prices at time $t = 1, \ldots, T$ are described as random variables on $(\Omega, (\mathcal{F}_t)_{t=0}^T, \mathcal{F}, P)$. The $i$th asset is available at time $t$ for a price $s^i_t > 0$ that is known to the investor at time $t$. In other words, $s^i_t \in L^0(\mathcal{F}_t)$. The $d$-dimensional random variable $s_t = (s^1_t, \ldots, s^d_t)$ is called the price vector at time $t$ and the $d$-dimensional adapted stochastic process $s := (s_t)_{t=0}^T$ is called the price process.

Throughout, the $d$-dimensional portfolio processes will be denoted by $z$. These are adapted stochastic process $z = (z_t)_{t=0}^T$ with the interpretation that a portfolio $z_t$ is held over the period $(t, t+1]$. We assume that $z_T = 0$, which means that the agent liquidates the position at the terminal time. The set of such processes will be denoted by $\mathcal{N}_0$, i.e.,

$$\mathcal{N}_0 := \{ (z_t)_{t=0}^T | z_t \in L^0(\Omega, \mathcal{F}_t, P; \mathbb{R}^d), z_T = 0 \}.$$

The cost of updating a portfolio from $z_{t-1}$ to $z_t$ at time $t$ is thus $s_t \cdot \Delta z_t$, where

$$\Delta z_t := z_t - z_{t-1}.$$

Summing over $t = 0, \ldots, T$, the total costs are given by the discrete time stochastic integral

$$\sum_{t=0}^T s_t \cdot \Delta z_t.$$
Using summation by parts, we get the total gains
\[
\sum_{t=0}^{T-1} z_t \Delta S_{t+1} = -\sum_{t=0}^{T} s_t \cdot \Delta z_t
\]
over the whole trading period.

**Definition 2.1.** A portfolio \( z \in \mathbb{R}^d \) is called an arbitrage opportunity, if
\[
\sum_{t=0}^{T-1} z_t \Delta s_{t+1} \geq 0 \quad \text{P-a.s. and } \quad P\left( \sum_{t=0}^{T-1} z_t \Delta s_{t+1} > 0 \right) > 0.
\]

If the market model does not admit arbitrage opportunities, we say that the model is arbitrage-free, or that the model has the NA-property.

Clearly, an arbitrage opportunity is an investment strategy without any risk.

Throughout the course, \( c \) denotes a contingent claim that our agent has to deliver at the terminal time \( T \). We assume that \( c \in L^0(\mathcal{F}) \) so that the claim is known at time \( T \) to the agent (recall that \( \mathcal{F}_T = \mathcal{F} \)). A claim \( c \) is called a derivative if it is measurable w.r.t. to the \( \sigma \)-algebra \( \sigma(s_0, \ldots, s_T) \) generated by the price process. In this case, by Doob’s measurability theorem, there exists a measurable \( F : \mathbb{R}^{d \times (T+1)} \rightarrow \mathbb{R} \) such that \( c = F(s) \).

We allow \( c \) to have both positive and negative values so that it may present both gains as well as losses, that is, both assets and liabilities. An agent facing the claim \( c \) and following the trading strategy \( z \) has the terminal wealth
\[
\sum_{t=0}^{T-1} z_t \Delta S_{t+1} - c
\]
that is an \( \mathcal{F}_T \)-measurable random variable.

**Remark 2.2.** Throughout the course, we work with discounted price processes and discounted claims. Assuming that there is an additional asset (bank account) that pays interest with a rate \( r > -1 \) and other prices are given by a price process \( \tilde{s} \), the discounted price process is
\[
s_t = \tilde{s}_t/(1+r)^t.
\]
Similarly, given a claim \( \tilde{c} \), the discounted claim is
\[
c = (1+r)^{-T} \tilde{c}.
\]

### 2.2 Changes of measure and martingales

Let \( Q \) be (another) probability measure on \((\Omega, \mathcal{F})\). The set
\[
L^1(Q) := \{ \eta \in L^0 \mid E^Q|\eta| < \infty \}
\]
is the space of $Q$-integrable random variables. We also denote $L^1 := L^1(P)$.

The measures $P$ and $Q$ on are equivalent if

$$P(A) = 0 \iff Q(A) = 0 \quad \forall A \in F,$$

in which case we denote $P \sim Q$. The measure $Q$ is said to be absolutely continuous w.r.t. $P$ if

$$P(A) = 0 \implies Q(A) = 0 \quad \forall A \in F,$$

in which case we denote $Q \ll P$. When $Q \ll P$, $dQ/dP$ denotes the Radon-Nikodym density of $Q$ w.r.t. $P$, that is, $Q(A) = E^{P}[dQ/dP 1_A]$, or, more generally, $E^Q[\eta] = E^P\left[\frac{dQ}{dP}\eta\right]$ for all $\eta \in L^1(Q)$.

For any $\eta \in L^1(Q)$, we denote the $\mathcal{F}_t$-conditional $Q$-expectation of $\eta$ by $E^Q_t\eta$,

$$E^Q_t\eta := E^Q[\eta | \mathcal{F}_t].$$

We will frequently need the conditional Jensen’s inequality

$$h(E^Q_t\xi) \leq E^Q_t h(\xi) \quad \forall \xi \in L^1(Q), t = 0, \ldots, T,$$

(2.2)

where $h$ is a convex function on the real line. Using conjugates of convex functions, we prove this inequality later on as an exercise.

Our notation for the conditional expectation allows us to write neatly that a stochastic process $(m_t)_{t=0}^T$ is a $Q$-martingale if, for all $t = 0, \ldots, T$, $m_t \in L^1(Q)$ and

$$E^Q_t \Delta m_{t+1} = 0 \quad \forall t = 0, \ldots, T - 1,$$

where, for $d$-dimensional processes, the equation is understood component-wise.

A class of martingales is generated by Radon-Nikodym densities. For $Q \ll P$, the stochastic process $(q_t)_{t=0}^T$ defined by

$$q_t := E^P_t \left[\frac{dQ}{dP}\right]$$

is called the density process of $Q$ (with respect to $P$).

**Lemma 2.3.** Let $Q \ll P$ and $q$ be the density process of $Q$. For any $\eta \in L^1(Q)$ and $t = 1, \ldots, T$, we have

$$E^Q_t \eta = \frac{1}{q_t} E^P_t [q_T \eta] \quad Q\text{-a.s.}$$

**Exercise.** Let $Q \ll P$ and $(q_t)_{t=1}^T$ be the density process of $Q$ and $m$ be a $Q$-martingale. Show that the density process $q$ and the process $(q_t m_t)_{t=0}^T$ are $P$-martingales.
Recall that a random variable \( \tau : \Omega \to \{0, 1, 2, \ldots, T\} \) is called a stopping time if \( \{\tau \leq t\} \in \mathcal{F}_t \) for each \( t \). A stochastic process \((m_t)_{t=0}^{T}\) is a local martingale if there is a nondecreasing sequence \((\nu^\tau)_{\nu=1}^{\infty}\) of stopping times, \(\nu^\tau \to T\), such that each stopped process \(m_{\nu^\tau}^\tau\) defined as

\[
m_{\nu^\tau}^\tau := m_{\min(t, \nu^\tau)}
\]

is a (true) martingale. The sequence of stopping times in the definition is called as a localizing sequence. Here the wording "true" just emphasizes the property that the martingale is not only a local martingale but a martingale.

The following lemma is characteristic of discrete time stochastic analysis and no longer true in the continuous time setting. For a set \( A \), the function

\[
\mathbb{1}_A(c) := \begin{cases} 
1 & \text{if } c \in A, \\
0 & \text{otherwise}
\end{cases}
\]

is the "measure theoretic" indicator function of \( A \).

**Lemma 2.4.** A local martingale \( m \) is a martingale if \( E m_T^+ < \infty \).

**Proof.** Let \((\nu^\tau)_{\nu=1}^{\infty}\) be a localizing sequence for \( m \) and denote \( m_t^\tau := m_{\min(t, \nu^\tau)} \). Since \( m_0^\tau = m_0 \) and

\[
m_t^\tau = \sum_{t'=0}^{t-1} \mathbb{1}_{\{\tau^\nu=t'\}} m_{t'} + \mathbb{1}_{\{\nu^\tau \geq t\}} m_t, \quad t = 1, \ldots, T,
\]

we see, by proceeding inductively forward in time \( t = 1, \ldots, T \), that each \( \mathbb{1}_{\nu^\tau > t} m_t \) is integrable (and thus also \( \mathbb{1}_{\{\tau^\nu=t'\}} m_{t'} \) for \( t' \leq t \)). Thus both summands in the right side \( \Delta m_T^\nu = \mathbb{1}_{\{\tau^\nu=T\}} m_T - \mathbb{1}_{\{\tau^\nu=t\}} m_{T-1} \) are integrable. The martingale property of \( m^\nu \), \( E_{T-1} \Delta m_T^\nu = 0 \), gives us

\[
E_{T-1}[\mathbb{1}_{\{\tau^\nu=T\}} m_T] = E_{T-1}[\mathbb{1}_{\{\tau^\nu=T\}} m_{T-1}] = \mathbb{1}_{\{\tau^\nu=T\}} \mathbb{1}_{\{\tau^\nu=t\}} m_{T-1},
\]

where last equality follows from the fact that \( \mathbb{1}_{\{\tau^\nu=T\}} m_{T-1} \) is integrable and \( \mathcal{F}_{T-1} \)-measurable. Since \( \tau^\nu(\omega) = T \) for \( \nu \) large enough, we get that \( E_{T-1} m_T^+ = m_{T-1}^+ \), so \( m_{T-1}^+ \) is integrable. Proceeding recursively backwards in time, we see that each \( m_t^+ \) is integrable. In particular, the right side in

\[
(m^\nu)_T^+ = \left( \sum_{t'=0}^{T-1} \mathbb{1}_{\{\tau^\nu=t'\}} m_{t'} + \mathbb{1}_{\{\tau^\nu=T\}} m_T \right)^+ \leq \sum_t m_t^+
\]

is integrable so that, by Fatou’s lemma,

\[
E m_T^- \leq \liminf E(m_t^+)^- = \liminf[-E m_T^- + E(m_t^+) < \infty.
\]

Again, we may proceed inductively backwards to deduce that each \( m_t^- \) is integrable and that \( (m^\nu)_T^- \leq \sum_t m_t^- \). Thus, by the dominated convergence theorem for conditional expectation,

\[
E_{T-1} m_T^- = \lim E_{T-1} m_T^- = \lim m_T^- = m_{T-1}.
\]

Now the proof is finished by backward induction \( t = T - 1, \ldots, 0 \). \( \square \)
We finish the introductory part on classical martingale inequalities. We will use these only at the end of the course when we study continuous time limits of discrete time models.

Given a stochastic process \( y \), we define its maximal process \( y^* = (y^*_t)_{t=0}^T \) as
\[
y^*_t := \max_{t' \leq t} y_{t'}.
\]

Recall that an adapted stochastic process \( y \) with \( y_t \in L^1(Q) \) for all \( t \), is a \( Q \)-submartingale, or \( Q \)-supermartingale, respectively, if
\[
E_t \Delta y_{t+1} \geq 0 \quad \forall t = 0, 1, \ldots, T - 1,
\]
\[
E_t \Delta y_{t+1} \leq 0 \quad \forall t = 0, 1, \ldots, T - 1.
\]

**Theorem 2.5** (Doob’s inequality). Assume that \( y \) is a nonnegative submartingale. Then, for every \( \lambda > 0 \) and \( p \geq 1 \),
\[
P(y^*_T \geq \lambda) \leq \frac{E\left[ y^p_T \mathbb{1}_{\{y^*_T \geq \lambda\}} \right]}{\lambda^p}.
\]

**Proof.** Notice first that the process \( (y^p_t)_{t=0}^T \) is also a submartingale: for any \( 0 \leq t \leq T - 1 \), we have
\[
y^p_t \leq |E_t y_{t+1}|^p \leq E_t y^p_{t+1},
\]
where the first inequality comes from the submartingale property of \( y \) and from the fact that \( h(y) = y^p \) is nondecreasing on the positive axis, and the second inequality is the conditional Jensen’s inequality (2.2).

Denoting \( y_{-1} := 0 \), we have
\[
E\left[ y^p_T \mathbb{1}_{y^*_T \geq \lambda} \right] = E\left[ \sum_{t=0}^T y^p_t \mathbb{1}_{y^*_t \geq \lambda} \right] \\
\quad \geq E\left[ \sum_{t=0}^T y^p_t \mathbb{1}_{y^*_t < \lambda \cap y^*_t \geq \lambda} \right] \\
\quad \geq E\left[ \sum_{t=0}^T \lambda^p \mathbb{1}_{y^*_t < \lambda \cap y^*_t \geq \lambda} \right] \\
\quad = \lambda^p P(y^*_T \geq \lambda),
\]
where the first inequality follows from the submartingale property of \( y^p \) and the fact that each \( \{y^*_t < \lambda, y^*_t \geq \lambda\} \in \mathcal{F}_t \).
2.2.1 Exercises

Exercise 2.2.1. Assume that $(\Omega, \mathcal{F}, P) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$ (for the the usual Borel-$\sigma$-algebra $\mathcal{B}(\mathbb{R})$) and that the distribution of the random variable $\xi$ defined as $\xi(\omega) = \omega$ ($\xi$ is the coordinate mapping) with respect to the measure $P$ is the normal distribution $N(\mu, \sigma^2)$ with mean $\mu$ and standard deviation $\sigma$. For $a \in \mathbb{R}$, compute $E e^{a \xi}$.

Let $Q$ be a measure such that the Radon–Nikodym density is

$$
\frac{dQ}{dP}(\omega) = e^{\frac{-\alpha \omega - \frac{1}{2} \alpha^2}{\sigma^2}}.
$$

Show that the distribution of $\xi$ with respect to $Q$ is $N(0, \sigma^2)$.
Hint: Show that the cumulative distribution function is that of a normal distribution with standard deviation $\sigma$.

Exercise 2.2.2. Let $Q \ll P$ and $(q_t)_{t=1}^T$ be the density process of $Q$ and $m$ be a $Q$-martingale. Show that the density process $q$ and the process $(q_t m_t)_{t=0}^T$ are $P$-martingales.

Exercise 2.2.3. Assume that $m$ is a martingale. Show that the stochastic process $x = (x_t)_{t=0}^T$, given as a stochastic integral of $z \in \mathcal{N}_0$ w.r.t. $m$, $x_0 := 0,$

$$
x_t := \sum_{t'=1}^{t} z_{t-1} \Delta m_{t'} \quad t = 1, \ldots, T,\n$$

is a local martingale.
Hint: Show first that if there is $M \in \mathbb{R}$ such that $|z_t| \leq M$ for all $t$, then $x$ is a martingale,
2.3 Geometric price processes

In this section we assume that the price process is of a "geometric" form

\[ s_t := s_0 \prod_{t'=1}^{t} \frac{1 + R_{t'}^s}{1 + r} = s_{t-1} \frac{1 + R_{t'}^s}{1 + r} \quad \forall \ t = 1, \ldots, T \tag{2.3} \]

where \( r > -1 \) is the interest rate (discount factor) of the underlying riskless bank account, and

\[ R_t^s = \mu + \sigma \xi_t \tag{2.4} \]

for a sequence \((\xi_t)_{t=1}^T \) of independent identically distributed random variables with \( E\xi_t = 0 \) and \( E(\xi_t)^2 = 1 \). We also assume that the filtration is generated by \( \xi_t \), i.e. \( F_0 \) is trivial and \( F_t = \sigma((\xi_{t'})_{t'-1}^{t-1}) \) for \( t = 1, \ldots, T \). In addition, we assume that \( P(\xi_t = 0) = 0 \) so that \( F_t = \sigma((s_{t'})_{t'=1}^{t-1}) \)

Remark 2.6. For the nondiscounted price process \( \tilde{s}_t = s_t (1 + r)^t \), we have the identity

\[ R_t^s = \frac{\tilde{s}_t - \tilde{s}_{t-1}}{\tilde{s}_{t-1}}, \]

so \( R_t^s \) is the return of \( \tilde{s} \) at time \( t \). We also see that \( \tilde{s} \) is a solution of the stochastic difference (differential) equation

\[ \Delta \tilde{s}_t = \mu \tilde{s}_{t-1} \Delta t + \sigma \tilde{s}_{t-1} \Delta w_t, \]

where \( w_t := \sum_{t'} \xi_{t'} \) is a martingale with independent increments.

Definition 2.7 (Martingale measures). A probability measure \( Q \) is called a martingale measure of \( s \), if \( s \) is \( Q \)-martingale.

Synonyms for a martingale measures of a price process are a risk-neutral measure and a pricing measure. These terms, however, are misleading when one generalizes to more general illiquid "nonlinear market models".

Given a density process \( q = (q_t)_{t=0}^T \) of some \( Q \ll P \), it also has the product representation

\[ q_t = \prod_{t'=1}^{t} (1 + R_{t'}^q) \quad \forall \ t = 1, \ldots, T \tag{2.5} \]

for some adapted stochastic process \( R^q \).

Theorem 2.8. Assume that the price process satisfies (2.3) with \( \xi_t^2 = 1 \) and that \( Q \) is a martingale measure of \( s \). Then \((\xi_t)_{t=0}^T \) are independent w.r.t. \( Q \) with

\[ E^Q \xi_t = (r - \mu)/\sigma \quad Q\text{-a.s.} \quad \forall \ t = 1, \ldots, T, \]

and the density process with the representation (2.5) satisfies

\[ E[\xi_t R_t^q] = (r - \mu)/\sigma \quad P\text{-a.s.} \quad \forall \ t = 1, \ldots, T, \]

Proof. Exercise.
2.3.1 Exercises

In all the three exercises below, we work under the assumptions of Section 2.3.

Exercise 2.3.1. Consider a one step pricing model with a bank account \( s_0 = 1 \), \( s_1 = (1 + r) \), and with one stock \( \tilde{s}^1 \), which has dynamics \( \tilde{s}_0 = \tilde{s}_0 = s_0 \) and \( \tilde{s}_1 = s_0(1 + \xi) \), where \( \xi \) is a random variable on \((\Omega, \mathcal{F}, P)\) with two possible values \( u, d \in \mathbb{R}, -1 < d < u \):

\[
P(\xi = u) = p = 1 - P(\xi = d).
\]

Show that there is no arbitrage in this model, if \( d < r < u \). Defining a probability measure \( Q \) by

\[
Q(\xi = u) = \frac{r - d}{u - d},
\]

compute \( E_Q s_1 \). Is \( Q \) equivalent to \( P \)?

Exercise 2.3.2. Prove Theorem 2.8.

Exercise 2.3.3 (Binomial model). Assume that \( \xi_t \) in (2.4) can take only two values, \( P(\xi_t = 1) = 1 - P(\xi_t = -1) \), and that \( \mu \) and \( \sigma \) are such that \( -1 < R^t_s \) almost surely. Use Theorem 2.8 to show that there is only one martingale measure \( Q \). Find also a formula for each \( R^Q_t \) in (2.5).

Hint: Make an Ansatz that \( R^Q_t \) is a linear function of \( \xi_t \).
3 Asset liability management and indifference pricing

Our agent liable to a claim $c$ and following the trading strategy $z$ needs to come up with

$$c - \sum_{t=1}^{T} z_{t-1} \Delta s_t$$

amount of cash after settling the claim at the terminal time. This is a random variable belonging to $L^0$. We assume that the agent’s risk preferences are given by an extended real-valued function $V : L^0 \rightarrow \mathbb{R}$ so that she faces an optimization problem

$$\text{minimize} \quad V \left( c - \sum_{t=1}^{T} z_{t-1} \Delta s_t \right) \text{ over } z \in \mathcal{N}_0. \quad \text{(ALM)}$$

We call this optimization problem as the Asset liability management problem. Indeed, since we allow $c$ have both positive and negative values, it can describe both assets as well as liabilities. We assume throughout that our risk preference satisfies the following assumption. The last item means that $V$ is a convex function, so we may say that $V$ is a convex risk preference.

**Assumption 3.1.** The risk preference $V : L^0 \rightarrow \mathbb{R}$ is nonconstant and satisfies

1. (Normalization) $V(0) = 0$,
2. (Monotonicity) $V(c^1) \leq V(c^2)$ if $c^1 \leq c^2$ $P$-almost surely,
3. (Diversification principle) $V$ is convex: for any $c^1, c^2 \in L^0$ and $\alpha \in [0, 1]$,

$$V(\alpha c^1 + (1 - \alpha) c^2) \leq \alpha V(c^1) + (1 - \alpha) V(c^2).$$

Diversification principle means that "diversified position" $\alpha c^1 + (1 - \alpha) c^2$ has at most the same "risk" as the (correspondingly weighted) average risk of the individual positions.

**Example 3.1 (Expected loss).** Assume that $V$ is a nondecreasing convex function on the real line with $V(0) = 0$. Such $V$ is called a loss function. Then

$$V(c) = EV(c)$$

defines a convex risk preference. Here and what follows, we define the expectation of a random variable as $+\infty$ unless the positive part is integrable. This means, in particular, that if neither positive nor negative part is integrable, then the expectation is $+\infty$.

Notice that, by Jensen’s inequality, $EV$ is risk averse in the sense that

$$V(Ec) \leq EV(c)$$
for every integrable \( c \). Notice also that the mapping \( c \mapsto -V(-c) \) is a one-to-one mapping between loss functions and utility functions (concave nondecreasing functions) that are zero at the origin. The expected loss and the expected utility \( c \mapsto EU(c) \) thus amount to the same thing, up to these changes of signs.

We analyze the dependence of the optimal value of \((\text{ALM})\) on the claim. To this end, we introduce the value function \( \phi : L^0 \to \mathbb{R} \) of \((\text{ALM})\),

\[
\phi(c) := \inf_{z \in \mathbb{N}_0} V \left( c - \sum_{t=1}^T z_{t-1} \Delta s_t \right). \tag{3.1}
\]

**Exercise.** The value function \( \phi \) of \((\text{ALM})\) is a convex monotone function.

The set

\[
C := \{ c \in L^0 \mid \exists z \in \mathbb{N}_0 : \sum_{t=0}^{T-1} z_t \Delta s_t \geq c \text{ P-a.s.} \} \tag{3.2}
\]

of freely super-hedgeable claims plays an important role throughout the pricing theory. Given a set \( A \), the function

\[
\delta_A(c) := \begin{cases} 
0 & \text{if } c \in A, \\
+\infty & \text{otherwise}
\end{cases}
\]

is the indicator function of \( A \) (in the terminology of optimization theory).

**Example 3.2** (Super-hedging). Assume that \( V(c) = \delta_{L^0}(c) \). Then the optimal value of \((\text{ALM})\) is finite (and zero) if and only if there is a portfolio \( z \) such that

\[
\sum_{t=0}^{T-1} z_t \Delta s_t \geq c \text{ P-a.s.},
\]

that is, there exists a portfolio that super-hedges \( c \). Thus, we have that

\[
\phi = \delta_C.
\]

You may convince yourself that this \( V \) is the "most risk-averse" risk preference in the sense that it is the largest function satisfying Assumption 3.1.

A claim \( c \) belonging to the set

\[
C \cap (-C)
\]

is called redundant whereas \( c \) is called replicable if there exists \( z \in \mathbb{N}_0 \) such that \( c = \sum_{t=0}^{T-1} z_t \cdot \Delta s_{t+1} \). Recall that a set \( A \) is a cone if \( \lambda c \in A \) for every \( c \in A \) and \( \lambda \geq 0 \).

**Exercise.** Show that \( C \) is a convex cone and that \( C \cap (-C) \) is a linear space. Show also that every replicable claim is redundant, and, if the model is arbitrage-free, the converse holds as well.
Next we turn our attention to indifference pricing. Assume now that our agent has liabilities in the market whose terminal cash-value is given by $\bar{c}$. For instance, the initial wealth $w_0$ of an agent is described by a constant $\bar{c} = w_0$ (up to discounting). In general, we can think of $\phi(\bar{c})$ as the risk-level of the agent (when the agent trades optimally in the market), so

$$\{c \in L^0 \mid \phi(c) \leq \phi(\bar{c})\}$$

is the set of financial positions which are at least as acceptable to her as her current position $\bar{c}$.

For another claim $c$,

$$\pi_s(c; \bar{c}) := \inf_{\alpha \in \mathbb{R}} \{\alpha \mid \phi(\bar{c} + c - \alpha) \leq \phi(\bar{c})\}$$

describes the least amount of cash the agent needs to cover the claim without worsening her risk profile. The value of $\pi_s(c; \bar{c})$ is known as the indifferece price of $c$; the agent is willing to have the liability $c$ for $\pi_s(c; \bar{c})$ amount of cash.

Symmetrically,

$$\pi_b(c; \bar{c}) := \sup_{\alpha \in \mathbb{R}} \{\alpha \mid \phi(\bar{c} - c + \alpha) \leq \phi(\bar{c})\}$$

describes the largest amount of cash the agent is willing to pay to reduce her liabilities by $c$ without worsening her risk profile. We have that the buying price and selling price are related by

$$\pi_b(c; \bar{c}) = -\pi_s(-c, \bar{c}). \tag{3.3}$$

We denote

$$\pi_{\sup}(c) := \inf_{\alpha \in \mathbb{R}} \{\alpha \mid c - \alpha \in C\},$$

$$\pi_{\inf}(c) := \sup_{\alpha \in \mathbb{R}} \{\alpha \mid \alpha - c \in C\}.$$

These correspond to $\pi_s$ and $\pi_b$ in the case when $\bar{c} = 0$ and the risk preference is the one of superhedging, $V = \delta_{L^0}$. Here $\pi_{\sup}$ gives the superhedging costs and $\pi_{\inf}$ gives the subhedging costs, and we have

$$\pi_{\inf}(c) = -\pi_{\sup}(-c).$$

To elaborate, $\pi_{\sup}(c)$ is the least amount of cash the agent needs to cover the claim $c$ without any risk, and $\pi_{\inf}(c)$ the highest amount of cash the agent can pay without having any risk. Having this in mind, the inequalities in the following theorem make a lot of sense. The properties of $\pi_s(\cdot, \bar{c})$ in the theorem mean that it is a convex risk measure on $L^0$ in the sense of Exercise 3.1.5.

**Theorem 3.3.** Assume that the price process is arbitrage-free and that $\bar{c} \in \text{dom } \phi$. Then $\pi_s(\cdot, \bar{c})$ is a monotone convex function with $\pi_s(0, \bar{c}) \leq 0$ and

$$\pi_s(c + \alpha, \bar{c}) = \pi_s(c, \bar{c}) + \alpha. \tag{3.4}$$
We have $\pi_s(c; \bar{c}) \leq \pi_{\text{sup}}(c)$. If $\pi_s(0; \bar{c}) = 0$, then

$$\pi_{\text{inf}}(c) \leq \pi_b(c; \bar{c}) \leq \pi_s(c; \bar{c}) \leq \pi_{\text{sup}}(c),$$

where we have equalities throughout if $c$ is replicable.

**Proof.** That $\pi_s(\cdot, \bar{c})$ is a monotone convex function, $\pi_s(0, \bar{c}) \leq 0$, and (3.4) holds, are left as an exercise.

To prove $\pi_s(c; \bar{c}) \leq \pi_{\text{sup}}(c)$, let $\alpha > \pi_{\text{sup}}(c)$, which we may assume to be finite. There exists $z' \in N_0$ such that $c - \alpha + \sum_{t=1}^{T} z'_{t-1} \Delta s_t \leq 0$. Thus, for $z \in N_0$ that is an $\varepsilon$-solution for $\varphi(\bar{c})$, we get, from monotonicity of $V$, that

$$V \left( \bar{c} - c - \alpha - \sum_{t=1}^{T} (z + z')_{t-1} \Delta s_t \right) \leq V \left( \bar{c} - \sum_{t=1}^{T} z_{t-1} \Delta s_t \right) \leq \varphi(\bar{c}) + \varepsilon.$$

We deduce that $\varphi(\bar{c} - c - \alpha) \leq \varphi(\bar{c})$, which means that $\pi_s(c; \bar{c}) \leq \alpha$. Since $\alpha > \pi_{\text{sup}}(c)$ was arbitrary, we must have $\pi_s(c; \bar{c}) \leq \pi_{\text{sup}}(c)$.

Assume now that $\pi_s(0; \bar{c}) = 0$. By convexity,

$$0 = \pi_s((1/2)(-c) + (1/2)c) \leq (1/2)\pi_s(-c) + (1/2)\pi_s(c),$$

which together with (3.3) gives the claimed chain of inequalities. If $c$ is replicable, it is redundant (see the exercises), so $c \in C$ and $-c \in C$, which implies $\pi_b(c) = \pi_s(c) = 0$.

In the theorem above, the translation invariance (3.4) is characteristic of linear market models with a cash account. If, e.g., one adds portfolio constraints or illiquidity effects to the model, this property of indifference prices often disappears.
3.1 Exercises

Exercise 3.1.1. Consider (ALM) in the one step model $T = 1$ with $d = 1$, trivial initial $\sigma$-algebra $\mathcal{F}_0 = \{\Omega, \emptyset\}$ and with an expected loss $\mathcal{V} = EV$. Then the problem becomes

$$\text{minimize} \quad EV(c - z_0 \Delta s_1) \quad \text{over} \quad z_0 \in \mathbb{R}.$$ 

- Is there an optimal solution if $s$ is not arbitrage-free?
- For an "exponential disutility" function $V(c) = e^c - 1$, find a price process $s$ such that $EV(z_0 \Delta s_1) = +\infty$
  whenever $z_0 \notin [-1, 1]$. This shows that there might be "implicit constraints" in (ALM).

Exercise 3.1.2. Consider again the setting in the previous exercise. Find a price process $s$ and a loss function $V$ so that the optimal solution for this problem is $z = 0$ for every essentially bounded claim $c$.

Hint: We never assumed that the price process has to be positive.

Exercise 3.1.3. Let $X$ and $U$ be linear spaces. Show that

1. a set $K \subset X$ is convex if and only if $\delta_K$ is a convex function,
2. if $f^1, f^2$ are extended-real-valued convex functions on $X$, then $f^1 + f^2$ is convex (with the convention that $-\infty - \infty = -\infty + \infty = \infty$),
3. if $h$ is a convex function on $U$ and $A : X \to U$ is linear, then $f(x) = h(Ax)$ is convex,
4. if $\{f_i\}_{i \in J}$ are convex functions on $X$ ($J$ is an arbitrary index set), then $f(x) := \sup_{i \in J} f_i(x)$ is a convex function,
5. if $f$ is a jointly convex function on $X \times U$, then $\phi(u) := \inf_{x \in X} f(x, u)$ defines a convex function on $U$.

Exercise 3.1.4. Show that the value function of (ALM) is a convex monotone function.
Exercise 3.1.5 (Risk measures). An extended real valued function $\mathcal{R}$ on $L^\infty$ (the space of essentially bounded random variables), is called a convex risk measure if it is convex, nondecreasing, and $\mathcal{R}(c + \alpha) = \mathcal{R}(c) + \alpha$ for every constant $\alpha$. Show that

$$\mathcal{V}(c) = \inf_{d \in L^\infty} \{ \mathcal{R}(d) \mid d \geq c \}$$

is a convex risk preference.

Exercise 3.1.6. Show that the following statements are equivalent.

(a) The market model is arbitrage-free.

(b) The set

$$\mathcal{L} = \left\{ z \in \mathcal{N}_0 \left| \sum_{t=1}^T z_{t-1} \Delta s_t \leq 0 \ P\text{-a.s.} \right. \right\}$$

is a linear space.

(c)

$$\mathcal{C} \cap L_+^0 = \{0\}. \quad \text{(NA)}$$

Exercise 3.1.7. Show that $\mathcal{C}$ is a convex cone and that $\mathcal{C} \cap (-\mathcal{C})$ is a linear space. Show also that every replicable claim is redundant, and, if the model is arbitrage-free, the converse holds as well.

Exercise 3.1.8. Prove that $\pi_s(\cdot, \bar{c})$ is a monotone convex function, $\pi_s(0, \bar{c}) \leq 0$, and (3.4) holds.
4 Duality in indifference pricing

Both the asset liability problem (ALM) and the indifference prices are defined via convex optimization problems. This allows us to analyze them using the conjugate duality theory from convex optimization. In particular, we will see how the familiar formulas for superhedging prices via martingale measures are obtained as a special case of this theory.

4.1 Convex conjugates and duality in optimization

Recall that, for \( p \in [1, \infty) \), the Lebesque space \( L^p := L^p(\Omega, \mathcal{F}, P) \) is a Banach space under the norm \( \| q \| := (E|q|^p)^{1/p} \).

For \( p > 1 \), its Banach dual is \( L^{p'} \) for \( p' \) satisfying \( 1/p + 1/p' = 1 \). For \( p = 1 \), we set \( p' = \infty \), and the Banach dual of \( L^1 \) is the space \( L^\infty = L^\infty(\Omega, \mathcal{F}, P) \) of essentially bounded random variables. For all \( p = [1, \infty) \), the pairing between \( L^p \) and \( L^{p'} \) is given by

\[
\langle q, c \rangle := E[qc].
\]

Assume now that \( g \) is an extended real-valued convex function on \( L^{p'} \). The function \( g \) is said to be proper if it is not identically +\( \infty \) and if it never takes the value \( -\infty \). The function \( g \) is lower semicontinuous if the level-sets

\[
\{ c \in L^{p'} \mid g(c) \leq \alpha \}
\]

are closed for each \( \alpha \in \mathbb{R} \). Equivalently, \( g \) is lower semicontinuous if the epigraph

\[
\text{epi } g := \{ (c, \alpha) \in L^{p'} \times \mathbb{R} \mid g(c) \leq \alpha \}
\]

is closed, or if, for every \( c \in L^{p'} \),

\[
\liminf_{c' \to c} g(c') \geq g(c)
\]

for every sequence \( c' \to c \). We remark that, since \( L^{p'} \) are duals of Banach spaces \( L^p \), here it suffices to consider ordinary sequences. In general, one has to pass from sequences to “generalized sequences” or to “nets”.

Given an extended real-valued function \( g \) on \( L^{p'} \), its conjugate \( g^* : L^p \) is defined by

\[
g^*(q) = \sup_{c \in L^{p'}} \{ \langle q, c \rangle - g(c) \}.
\]

The function \( g^* \) is also known as Legendre-Fenchel transform, polar function, or convex conjugate of \( g \). Since \( g^* \) is a supremum of lower semicontinuous functions (actually, continuous functions), \( g^* \) is always a lower semicontinuous function on \( L^p \). The Fenchel inequality

\[
g(c) + g^*(q) \geq \langle q, c \rangle
\]
follows directly from the definition of the convex conjugate. In the exercises, we will familiarize ourselves with this transformation by calculating conjugates of convex functions defined on $\mathbb{R}^d$.

The \textit{biconjugate} of $g$ is the function

$$g^{**}(c) = \sup_q \{ \langle q, c \rangle - g^*(q) \}.$$ 

By the Fenchel inequality, we always have

$$g \geq g^{**}.$$ 

The following \textit{biconjugate theorem} is the fundamental theorem on convex conjugates. The formula in the theorem is known as the \textit{dual representation}.

\textbf{Theorem 4.1} (Biconjugate theorem). Assume that $g$ is a proper lowersemi-continuous convex extended real-valued function on $L^p$. Then $g = g^{**}$, i.e., $g$ has the dual representation

$$g(c) = \sup_{q \in L^p} \{ \langle c, q \rangle - g^*(q) \}.$$ 

\textit{Proof.} We start by stating (without a proof) the geometric fact (the Hahn-Banach theorem) that a closed convex set in a locally convex space (especially in $L^p \times \mathbb{R}$), is the intersection of all closed half spaces containing the set. Applying this to the epigraph of $g$, we get that $g$ is the supremum of all affine functions less than $g$, i.e.,

$$g(c) = \sup_{q \in L^p, \alpha \in \mathbb{R}} \{ \langle c, q \rangle - \alpha \mid \langle c, q \rangle - \alpha \leq g(c) \}.$$ 

Let $q \in L^p$ be fixed. Then $c \mapsto \langle c, q \rangle - \alpha$ is smaller than $g$ if $c \mapsto \langle c, q \rangle - \alpha \leq g(c)$ for all $c$, i.e.

$$\alpha \geq \sup_c \{ \langle c, q \rangle - g(c) \} = g^*(q).$$ 

We see from this that, if $g^*(q) < \infty$, then largest affine function with the slope $q$ less than $g$ is $c \mapsto \langle c, q \rangle - g^*(q)$, whereas if $g^*(q) = \infty$, there are no affine functions with slope $q$ less than $g$. Since this is valid for any $q$, we get that $c \mapsto \sup \{ \langle c, q \rangle - g^*(q) \}$ is the supremum of all affine functions less than $g$. Combining this with the first paragraph of the proof, we get the result.

In this course, we will need the following basic properties of convex functions. We give a proof in the case when the probability space is discrete. The general case is proved exactly in the same manner, but this requires some knowledge of functional analysis\footnote{Students familiar with functional analysis may notice that the Banach-Alaoglu theorem (the unit ball in the dual of a Banach space is weak*-compact) follows as a special case.}. We denote

$$B_\epsilon := \{ c \in \mathbb{R}^d \mid |c| < \epsilon \}.$$
Theorem 4.2. Let $g$ be a convex function on $L^{p'}$ that is finite at $\bar{c} \in L^{p'}$. Then $g$ is bounded from above on a neighborhood of $\bar{c}$ if and only if $g$ is continuous at $\bar{c}$. When $g$ is proper and lower semicontinuous, the following are equivalent:

1. $g$ is bounded from above on a neighborhood of $\bar{c}$,
2. for every $\alpha \in \mathbb{R}$, \( \{ q \in L^p \mid g^*(q) - \langle \bar{c}, q \rangle \leq \alpha \} \) are weakly compact.

Proof. We prove the claim when $g$ is a convex function on $\mathbb{R}^d$.

Replacing $g$ by $g(c + \bar{c}) - g(\bar{c})$, we may assume that $\bar{c} = 0$ and that $g(0) = 0$. Assume that $g$ is bounded from above at the origin. This means that there exists $B \epsilon$ and $\gamma > 0$ such that $g(c) \leq \gamma$ for all $c \in B \epsilon$. By convexity, $0 = g((1/2)((-c) + (1/2)c) \leq (1/2)g(-c) + (1/2)g(c)$, which implies that, $-g(-c) \leq g(c)$. Thus, for $c \in B \epsilon$, $|g(c)| \leq \gamma$. Now, using convexity again, it is an exercise to show that this implies continuity.

To prove that 1. implies 2., note that we have (see the exercises)
\[ g(c) \leq g(0) + \delta_{B \epsilon}(c) \quad \forall c \iff g^*(q) \geq \epsilon |q| - \gamma \quad \forall q, \]
so, for any $\alpha \in \mathbb{R}$,
\[ \{ q \in \mathbb{R}^d \mid g^*(q) \leq \alpha \} \subset \{ q \in \mathbb{R}^d \mid |q| \leq \alpha + \gamma \}, \]
where the set on the right side is bounded.

That 2. implies 1., we may again do translations so that $g^*(0) = \inf_q g^*(q) = 0$; the details are left as an exercise. Let $\gamma > 0$ and denote
\[ K := \{ q \in \mathbb{R}^d \mid g^*(q) \leq \gamma \}. \]

If $q \notin K$, we have
\[ j_K(q) := \inf_{\lambda > 0} \{ \lambda \mid q \in \lambda K \} \]
\[ = \inf_{\lambda > 0} \{ \lambda \mid q \in \lambda K \} \]
\[ = \inf_{\lambda > 1} \{ \lambda \mid g^*(q/\lambda) \leq \gamma \} \]
\[ \leq \inf_{\lambda > 1} \{ \lambda \mid g^*(q)/\lambda \leq \gamma \} \]
\[ = g^*(q)/\gamma. \]

If $q \in K$, we have $j_K(q) \leq 1$, so putting these together we get that $g^*(q) \geq \gamma j_K(q) - \gamma$. Using the change of variables $\tilde{q} = q/\lambda$, we see that
\[ j_K^*(c) = \sup_{q \in \mathbb{R}^d, \lambda > 0} \{ c \cdot q - \lambda \mid q/\lambda \in K \} \]
\[ = \sup_{q \in \mathbb{R}^d} \{ \sup_{\lambda > 0} \{ \lambda(c \cdot \tilde{q} - 1) \} \mid \tilde{q} \in K \} \]
\[ = \delta_K(c), \]

$^2$In $\mathbb{R}^d$, a closed set is weakly compact if and only if it is bounded.
where \( K^\circ := \{ c | c \cdot q \leq 1 \ \forall q \in K \} \) (polar set of \( K \)). It is an exercise to show that \( K^\circ \) contains some open ball centered at the origin (since \( K \) is assumed to be bounded in part 2.). Combining this with \( g^*(q) \geq \gamma_j K(q) - \gamma \), we see, by conjugating as in the proof of part 1. above, that
\[
g(c) \leq \delta_{K^\circ}(c/\gamma) + \gamma.
\]
Therefore, \( g \) is bounded above in a neighborhood of the origin.

Now we turn our attention to convex optimization problems. Assume that \( F \) is a jointly convex function on \( \mathcal{N}_0 \times L^p \) and consider the value function
\[
\phi(c) := \inf_{z \in \mathcal{N}_0} F(z, c).
\]
This function gives the optimal value of the convex minimization problem
\[
\text{minimize } F(z, c) \quad \text{over } z \in \mathcal{N}_0 \tag{P}
\]
as a function of \( c \). The value function is always convex (see the exercises), so, when it is proper and lower semicontinuous, the biconjugate theorem gives
\[
\inf_{z \in \mathcal{N}_0} F(z, c) = \phi(c) = \sup\{ (q, c) - \phi^*(q) \}.
\]
In this context, the optimization problem (P) is referred to as the primal problem and the optimization problem
\[
\text{maximize } \langle q, c \rangle - \phi^*(q) \quad \text{over } q \in L^p \tag{D}
\]
is called the dual problem.

When \( \phi \) is closed, their optimal values coincide, and one says that there is no duality gap, or there is strong duality between (P) and (D). Even when \( \phi \) is not closed, the inequality \( \phi \geq \phi^{**} \) means that, at least, the dual problem gives lower bounds to optimal value of the primal problem. This is often called weak duality between (P) and (D).

Conjugate duality also leads to optimality conditions that are given in terms of subgradients of convex functions. However, we do not go into this direction during this course.
4.1.1 Exercises

Exercise 4.1.1. Let $f$ be a convex lower semicontinuous function on the real line. Convince yourself that, given a "slope" $v$, $f^*(v)$ is the smallest constant $\alpha$ such that the affine function $x \rightarrow vx - \alpha$ is majorized by $f$. What does this mean geometrically?

Exercise 4.1.2. Calculate the conjugates of the following functions on the real line:

1. $f(x) = |x|
2. f(x) = \delta_{\mathbb{B}}(x)$, where $\mathbb{B} = \{x \mid |x| \leq 1\}$.
3. $f(x) = \frac{1}{p}|x|^p$, for $p > 1$.
4. $V(x) = (e^{ax} - 1)/a$.

Exercise 4.1.3. Let $V$ be a loss function and analyze $V^*$ using the geometric idea from the first exercise.

1. Is $V^*$ positive?
2. Is $V^*$ zero somewhere?
3. Is $V^*$ monotone?
4. Where is $V^*$ finite?
5. Is $V^*$ necessarily finite at the origin?

Hint: For the last three question, the answer depends on your choice of $V$. 
4.2 The dual problem of the asset liability problem

In this section we will derive the dual representation for the optimal value function $\varphi$ of (ALM). We assume from now on the risk preference $V$ is an expected loss, that is,

$$V(c) := EV(c)$$

for some nonconstant nondecreasing lower semicontinuous convex function $V : \mathbb{R} \to \mathbb{R}$ with $V(0) = 0$. Recall that such $V$ is called a loss function.

The following is a special case of the theorem on "interchange of integration and minimization" whose proof relies on a measurable selection argument.

**Theorem 4.3.** Let $h$ be an extended real-valued convex function on the real line and let

$$g(c) = Eh(c)$$

be defined on $L^p'$. Then $g$ is a proper lower semicontinuous convex function with

$$g^*(q) = Eh^*(q).$$

**Proof.** We prove the conjugate formula. That $g$ is proper and lower semicontinuous are left as an exercise. By Fenchel inequality,

$$h^*(q) \geq qc - h(c),$$

so, by taking expectations and supremum over $c \in L^p'$, we get

$$Eh^*(q) \geq g^*(q).$$

Thus, we only need to show the opposite inequality

$$g^*(q) \geq Eh^*(q).$$

Let $\bar{c}$ and $\bar{q}$ be such that $h(\bar{c})$ and $h^*(\bar{q})$ are finite. Let $\epsilon > 0$ and consider the set \{$(c, \omega) \in \mathbb{R} \times \Omega | h(c) - cq(\omega) \leq -h^*(q(\omega)) - \epsilon$\}. By a measurable selection theorem\(^3\), there is a random variable $c \in L^0$ such that, for all $\omega$,

$$h(c(\omega)) - c(\omega)q(\omega) \leq -h^*(q(\omega)) - \epsilon.$$ 

For $\nu = 1, 2, \ldots$, defining $c^\nu = \mathbb{1}_{c \leq q^\nu} + \mathbb{1}_{c > q^\nu} \bar{c}$, we have that $c^\nu \in L^p'$, $c^\nu \to c$ and $h(c^\nu) - c^\nu q \to h(c) - cq$ almost surely. Since

$$h(c^\nu) - c^\nu q \leq \max\{h(c) - cq, h(\bar{c}) - \bar{c}q\},$$

(reverse) Fatou’s lemma gives

$$\lim \sup E[h(c^\nu) - c^\nu q] \leq E[h(c) - c(q)] \leq E[-h^*(q(\omega)) - \epsilon].$$

This means that, for $\nu$ large enough, $(c^\nu, q) - Eh(c^\nu) - \epsilon \leq Eh^*(q)$. Since $\epsilon > 0$ was arbitrary, we have $g^*(q) \leq Eh^*(q)$.

\(^3\)We omit the details.
The set of positive multiples of $P$-absolutely continuous martingale measures of $s$ will be denoted by $Q$. By Lemma ??,
$$Q := \{ q \in L^1 \mid E_t[q \Delta s_{t+1}] = 0 \}. \quad (4.1)$$

**Theorem 4.4.** Assume that $s^*_T \in L^{p'}$. Then the conjugate of the value function $\varphi : L^{p'} \mapsto \mathbb{R}$ associated with (ALM) has the expression
$$\varphi^*(q) = \begin{cases} EV^*(q) & \text{if } q \in Q, \\ +\infty & \text{otherwise}. \end{cases}$$

**Proof.** Assume that $q \in Q$, $c$, and $z$ are such that $EV^*(q) < \infty$ and that $EV(c - \sum_{t=1}^T z_{t-1} \Delta s_t) < \infty$. We get from the Fenchel inequality
$$V(u) + V^*(y) \geq yu \quad \forall u, y \in \mathbb{R}$$
that
$$V \left( c - \sum_{t=1}^T z_{t-1} \Delta s_t \right) + V^*(q) \geq qc - q \sum_{t=1}^T z_{t-1} \Delta s_t \quad P\text{-a.s.}$$
Since the positive part of the left side is integrable, we get from Exercise ?? and Lemma 2.4 that
$$E \left[ q \sum_{t=1}^T z_{t-1} \Delta s_t \right] = 0.$$ 
Thus
$$\varphi(c) + EV^*(q) \geq \langle q, c \rangle,$$
which implies $\varphi^*(q) \leq EV^*(q)$.

To prove the other direction, let $q \in L^p$ be arbitrary. By the assumption that $s^*_T \in L^{p'}$, we have that $\sum_{t=1}^T z_{t-1} \Delta s_t$ belongs to $L^p$ for any $z \in \mathcal{N}_0^\infty$. Thus
$$\varphi^*(q) = \sup_{c \in L^{p'}} \left\{ \langle c, q \rangle - \inf_{z \in \mathcal{N}_0} EV \left( c - \sum_{t=1}^T z_{t-1} \Delta s_t \right) \right\}$$
$$= \sup_{c \in L^{p'}, z \in \mathcal{N}_0} \left\{ \langle c, q \rangle - EV \left( c - \sum_{t=1}^T z_{t-1} \Delta s_t \right) \right\}$$
$$\geq \sup_{c \in L^{p'}, z \in \mathcal{N}_0^\infty} \left\{ \langle c, q \rangle - EV \left( c - \sum_{t=1}^T z_{t-1} \Delta s_t \right) \right\}$$
$$= \sup_{z \in \mathcal{N}_0^\infty} \left\{ E[|q \sum_{t=1}^T z_{t-1} \Delta s_t|] + \langle \tilde{c}, q \rangle - EV(\tilde{c}) \right\}$$
$$= \sup_{z \in \mathcal{N}_0^\infty} \{ E[q \sum_{t=1}^T z_{t-1} \Delta s_t] + EV^*(q) \}$$
$$= \delta_Q(q) + +EV^*(q)$$

22
where the second last inequality follows Theorem 4.3 and the last equality by applying (4.1).

\[ \square \]

**Corollary 4.5.** Assume that \( s^*_T \in L^{p'} \). Then the polar cone of the set of superhedgeable claims \( \mathcal{C} \) is \( \mathcal{Q} \), i.e.

\[ \sigma_{\mathcal{C}}(q) = \delta_{\mathcal{Q}}. \]

**Proof.** This is a special case of the theorem above, when \( V = \delta_{\mathbb{R}^-} \). Indeed, then \( \varphi = \delta_{\mathcal{C}} \) and \( V^* = \delta_{\mathbb{R}^+} \).

\[ \square \]

**Theorem 4.6.** Assume that \( s^*_T \in L^{p'} \). If \( \varphi \) is proper and lower semicontinuous at \( c \), then there is no duality gap between the asset liability problem

\[
\begin{align*}
\text{minimize} & \quad \mathbb{E} V \left( c - \sum_{t=1}^{T} z_{t-1} \Delta s_t \right) \quad \text{over} \quad z \in \mathcal{N}_0 \\
\text{and the dual problem} & \\
\text{maximize} & \quad \mathbb{E} [qc - V^*(q)] \quad \text{over} \quad q \in \mathcal{Q}.
\end{align*}
\]

If \( V \) is a finite loss function, then under any of the following conditions:

1. \( \mathbb{E} V \) is continuous on \( L^{p'} \)
2. \( L^{p'} = L^\infty \),
3. \( (\Omega, \mathcal{F}, P) \) is a discrete probability space,

the value function \( \varphi \) is proper and lower semicontinuous everywhere and the dual problem has a solution.

**Proof.** The absence of the duality gap follows from the biconjugate theorem and Theorem 4.4. When \( V \) is a finite loss function, the rest of the claims follow from Theorem 4.2. We omit the details, but it is a good exercise to verify the claims under the first and the third condition. Under the second condition this is considerably harder.

\[ \square \]
4.2.1 Exercises

Exercise 4.2.1. Let $h$ be an extended real-valued proper lower semicontinuous convex function on the real line and

$$g(c) = Eh(c)$$

be defined of $L^p'$. Show that $g$ is proper and lower semicontinuous. Prove the inequality

$$g^*(q) \geq Eh^*(q)$$

when $q$ is a simple random variable in $L^p$. Hint: Use Fatou’s lemma and the sequential characterization of lower semicontinuity. For the inequality, consider simple $c \in L^p'$ that are constants on the same sets as $q$.

Exercise 4.2.2. Prove the conditional Jensen’s inequality (2.2) using Theorem 7.9. Hint: Use the dual representation of $Eh$.

Exercise 4.2.3. Consider the asset liability problem in the case $s_T^* \in L^1$ and $c \in L^\infty$ and for an expected loss with an exponential loss function from Exercise 4.1.2. Find an expression for the optimal value in terms of martingale measures $Q_m := \{ q \in Q \mid Eq = 1 \}$ of $s$. Hint: $Q = \{ \alpha q \mid \alpha \in \mathbb{R}_+, q \in Q_m \}$.
4.3 Pricing formulas for indifference prices

Recall that, given a current financial position $\bar{c}$, the indifference price of a claim $c$ is given by

$$\pi_s(c; \bar{c}) := \inf_{\alpha \in \mathbb{R}} \{ \alpha \mid \varphi(\bar{c} + c - \alpha) \leq \varphi(\bar{c}) \}.$$ 

In this section we concentrate on claims that belong to $L^p'$. Then the indifference price can be expressed

$$\pi_s(c, \bar{c}) = \inf_{\alpha \in \mathcal{A}(\bar{c})} \{ \alpha \mid c - \alpha \in \mathcal{A}(\bar{c}) \},$$

where the acceptance set

$$\mathcal{A}(\bar{c}) := \{ c \in L^p' \mid \varphi(c + \bar{c}) \leq \varphi(\bar{c}) \},$$

describes the claims that the agent finds "acceptable" given her risk preferences and her ability to trade in the market. In the next theorem, the function

$$\sigma_{\mathcal{A}(\bar{c})}(c) = \sup_{c \in \mathcal{A}(\bar{c})} E[cq]$$

is known as the support function of $\mathcal{A}(\bar{c})$. We denote $Q^p := Q \cap L^p$.

**Theorem 4.7.** Assume that $\mathcal{F}_0$ is the trivial $\sigma$-field, $\varphi$ is proper and lower semicontinuous, and that the price process is arbitrage-free. Then

$$\pi(c; \bar{c}) = \sup_{q \in Q^p} \{ Ecq - \sigma_{\mathcal{A}(\bar{c})}(q) \mid Eq = 1 \}.$$ 

If $\inf \varphi < \varphi(\bar{c})$, then

$$\sigma_{\mathcal{A}(\bar{c})}(q) = \inf_{\lambda > 0} \{ \lambda [EV^*(q/\lambda) - \varphi(\bar{c})] - E[\bar{c}q] + \delta_Q(q) \}.$$ 

**Proof.** We have

$$\pi^*_s(q; \bar{c}) = \sup_{c \in L^p'} \{ E[cq] - \pi_s(c, \bar{c}) \}$$

$$= \sup_{\alpha \in \mathbb{R}, c \in L^p'} \{ E[cq] - \alpha \mid c - \alpha \in \mathcal{A}(\bar{c}) \}$$

$$= \sup_{\alpha \in \mathbb{R}, \tilde{c} \in L^p'} \{ E[(\tilde{c} + \alpha)q] - \alpha \mid \tilde{c} \in \mathcal{A}(\bar{c}) \}$$

$$= \sup_{\alpha \in \mathbb{R}, \tilde{c} \in L^p'} \{ E[(\tilde{c}q + \alpha(q - 1)) \mid \tilde{c} \in \mathcal{A}(\bar{c}) \}$$

$$= \begin{cases} \sigma_{\mathcal{A}(\bar{c})}(q) & \text{if } Eq = 1, \\ +\infty & \text{otherwise}. \end{cases}$$

Thus the first formula follows from the biconjugate theorem as soon as $\pi_s(\cdot, \bar{c})$ is proper and lower semicontinuous. This will be a topic in the exercises.
The support function $\sigma_{A(\bar{c})}$ satisfies

\[
\sigma_{A(\bar{c})}(q) = \sup_{c \in L^{p'}} \{ (c, q) \mid \varphi(c + \bar{c}) \leq \varphi(\bar{c}) \}
\]

\[
= \inf_{\lambda \geq 0} \sup_{c \in L^{p'}} \{ (c, q) - \lambda(\varphi(\bar{c} + c) - \varphi(\bar{c})) \}
\]

\[
= \inf_{\lambda \geq 0} \sup_{c \in L^{p'}} \{ (\bar{c} - c, q) - \lambda(\varphi(\bar{c}) - \varphi(\bar{c})) \}
\]

\[
= \inf_{\lambda \geq 0} \{ \lambda \varphi^\ast(q/\lambda) - E[\bar{c}q] + \lambda \varphi(\bar{c}) \}.
\]

Here the second line is based on Lagrangian duality and the assumption that \(\inf \varphi < \varphi(\bar{c})\). Thus the result follows from Theorem 4.4.

\[ \square \]

**Example 4.8** (Superhedging prices). For the superhedging prices, we have \(V = \delta_{R_-}\) and \(\bar{c} = 0\), so the above formula reduces to

\[
\pi_{\sup}(c) = \sup_{q \in \mathcal{Q}} \{ E[cq] \mid E[q] = 1 \}.
\]

We have recovered the pricing formula for superhedging prices in terms of absolutely continuous martingale measures of \(s\).
5 Continuous time limits of discrete time models

In this section we are interested in characterizing limits of derivative prices given by discrete time market models when the agent trades with higher and higher frequency. In the $n$-th step model, $n = 1, 2, 3, \ldots$, we allow the agent to trade at equidistant (dyadic) time points $\{ k/2^n \}_{k=0}^{2^n}$. At each $n$-th stage, the prices are given by an arbitrage-free price process $(s_t^{(n)})_{t=0}^{2^n}$, whose absolutely continuous martingale measures we denote by $Q^{(n)}_m$.

Denoting by $C := C([0,1])$ the space of continuous functions, we define a linear interpolation operator $W_n : \mathbb{R}^{2^n+1} \rightarrow C$ by

$$[W_n((y_k)_{k=0}^{2^n})]_t := ([2^n t] + 1 - 2^n t) y_{[2^n t]} + (2^n t - [2^n t]) y_{[2^n t]+1}, \quad 0 \leq t \leq 1,$$

where $[2^n t]$ denotes the integer part of $2^n t$. Then $W_n(s^{(n)}(\omega))$ defines a random continuous path. Assume now that $F : C \rightarrow \mathbb{R}$ is a continuous function. Then, for $n = 1, 2, \ldots$,

$$c^{(n)} := F\left(W_n(s^{(n)})\right)$$

defines a claim (a path-dependent derivative) in the $n$-th step model. If $F$ depends only on the terminal point $t = 1$, i.e. $F(y) = F_1(y_1)$ for some continuous function $F_1$, then $F$ is a "European type claim".

We know from our discrete time results, how to price such claims for each discrete time model $(n)$. We want to understand how super-hedging prices converge when the length of the time steps converges to zero. This leads us to study weak convergence of stochastic processes. Indeed, assuming $F$ is a bounded continuous function on $C$, and that the continuous stochastic processes $W(s^{(n)})$ converge weakly along $Q^{(n)}$ to some $s$ under $Q$, then, as we will learn, by the definition of weak convergence,

$$\lim E^{Q^{(n)}} F\left(W(s^{(n)})\right) = E^Q F(s).$$

This gives a formula for the superhedging prices at the limit. For example, we will see how the famous Black & Scholes formula arises in this way. However, some addition complications arise since many derivatives, like European put and call options in the Black & Scholes formula, are not given in terms of bounded $F$. 

27
5.1 Convergence in distribution on metric spaces

We give the definition and fundamental results on convergence in distribution for random variables that take values in a complete separable metric space \( S \). Later on, \( S \) will be either \( C, \mathbb{R}^d \) or \( \mathbb{R} \). We equip all metric spaces with their Borel-\( \sigma \)-algebras.

We assume throughout this section that \( \xi, \xi^{(n)}, n = 1, 2, \ldots, \) are \( S \)-valued random variables defined on \((\Omega, \mathcal{F}, \mathbb{P}), (\Omega^{(n)}, \mathcal{F}^{(n)}, Q^{(n)}), n = 1, 2, \ldots, \) respectively. The expectation w.r.t. \( Q^{(n)} \) is denoted by \( E^{(n)} \).

**Definition 5.1.** The sequence \((\xi^{(n)})_{n=1}^{\infty}\) converges in distribution to \( \xi \) if

\[
E^{(n)} F(\xi^{(n)}) \to E F(y)
\]

for every continuous bounded function \( F \) on \( S \). This convergence is denoted by

\[
\xi^{(n)} \xrightarrow{d(Q^{(n)})} \xi.
\]

We begin by collecting some famous theorems about the convergence in distribution. For a set \( A \) in a metric space \( S \), we denote its interior by \( \text{int} A \), closure by \( \text{cl} A \), and boundary by \( \partial A := \text{cl} A \cap \text{cl}(A^C) \). The distance in \( S \) is denoted by \( d \), and \( \epsilon \)-fattening by \( ^\epsilon \! A := \{ y \in S \mid d(A, y) \leq \epsilon \} \).

**Theorem 5.2** (Portmanteau theorem). The following conditions are equivalent:

1. \( \xi^n \xrightarrow{d(Q^{(n)})} \xi \),
2. \( \liminf_{n \to \infty} Q^{(n)}(\xi^{(n)} \in G) \geq Q(\xi \in G) \) for any open \( G \subseteq S \),
3. \( \limsup_{n \to \infty} Q^{(n)}(\xi^{(n)} \in F) \leq Q(\xi \in K) \) for any closed \( K \subseteq S \),
4. \( \lim_{n \to \infty} Q^{(n)}(\xi^{(n)} \in B) \to Q(\xi \in B) \) for any \( B \in \mathcal{B}(S) \) with \( Q(\xi \in \partial B) = 0 \).

**Proof.** (i) \( \Rightarrow \) (ii): given an open set \( G \subseteq S \), there is a sequence \((F^n)_{n=1}^{\infty}\) of nonnegative continuous functions on \( S \) such that \( F^n \xrightarrow{\mathcal{L}} 1_G \). Assume for a contradiction that there is \( \epsilon > 0 \) and a subsequence \( n \) (we denote this subsequence with the same \( n \)) and \( N \) such that, for all \( n > N \), \( Q^{(n)}(\xi^{(n)} \in G) + \epsilon < Q(\xi \in G) \). By the monotone convergence theorem, there is \( M \) such that, for all \( \nu \geq M \), \( E 1_G(\xi) \leq EF^{(\nu)}(\xi) + \epsilon \). For \( \nu \geq M \), choose \( n \geq N \) such that \( EF^{(\nu)}(\xi) \leq E^{(n)} F^{(\nu)}(\xi^{(n)}) + \epsilon \). Since \( E^{(n)} F^{(\nu)}(\xi^{(n)}) \leq Q^{(n)}(\xi^{(n)} \in G) \), this is a contradiction.

(ii) \( \Rightarrow \) (i). Given a nonnegative continuous bounded function \( F \) on \( S \), each \( \{ y \in C \mid F(y) > t \} \) is open, so, by Exercise, Fatou’s lemma, and Fubini’s

---

4We omit the proof of this property of metric spaces.
theorem,

\[ EF(\xi) = \int_0^\infty Q(F(\xi) > t) dt \]

\[ \leq \int_0^\infty \liminf_n Q^{(n)}(F(\xi^{(n)}) > t) dt \]

\[ \leq \liminf_n \int_0^\infty Q^{(n)}(F(\xi^{(n)}) > t) dt \]

\[ \leq \liminf_n E^{(n)} \int_0^\infty 1_{\{F(\xi^{(n)}) > t\}} dt \]

\[ = \liminf_n E^{(n)} F(\xi^{(n)}). \]

Applying the inequality to \( L \pm F \), where \( |F| \leq L \), gives (i).

The conditions (ii) and (iii) are equivalent by taking complements.

(ii) and (iii) \( \Rightarrow \) (iv): for any \( B \in \mathcal{B}(S) \), we have

\[ Q(\xi \in \text{int } A) \leq \liminf Q^{(n)}(\xi^{(n)} \in \text{int } A) \]

\[ \leq \liminf Q^{(n)}(\xi^{(n)} \in \text{cl } A) \]

\[ \leq \limsup Q^{(n)}(\xi^{(n)} \in \text{cl } A) \]

\[ \leq Q(\xi \in \text{cl } A). \]

When \( Q(\xi \in \partial B) = 0 \), the left side and the last line are equal, so we have (iv).

(iv) \( \Rightarrow \) (ii): given a closed set \( K \subseteq S \), the sets \( \partial K \) are disjoint for different \( \epsilon > 0 \), so \( \{ \xi \in \partial K \} \) is a \( Q \)-null-set for almost every \( \epsilon > 0 \). Indeed, \( \omega \mapsto d(K, \xi(\omega)) \) is a real-valued random variable, so \( Q(d(K, \xi) = \epsilon) \) is nonzero only for at most countably many \( \epsilon \). We have \( Q^{(n)}(\xi^{(n)} \in F) \leq Q^{(n)}(\xi^{(n)} \in 'F) \), so (iii) follows by letting \( n \to \infty \) and then \( \epsilon \searrow 0 \); the rigorous proof is via a contradiction argument just as in the first paragraph of the proof. \( \square \)

**Theorem 5.3** (Continuous mapping theorem). Let \( f : S \to V \) be a continuous mapping between metric spaces. If \( x^n \xrightarrow{d(Q^{(n)})} x \), then \( f(x^n) \xrightarrow{d(Q^{(n)})} f(x) \).

**Proof.** For any continuous bounded function \( F \) on \( V \), the composition \( y \mapsto F(f(y)) \) is continuous and bounded on \( V \). \( \square \)

The following theorem shows that on product spaces, the convergence in distribution is characterized by “cylindrical” bounded continuous functions.

**Theorem 5.4.** Let \( \tilde{S} \) be another metric space and let \( \tilde{\xi}, \tilde{\xi}^{(n)} \), \( n = 1, 2, \ldots \) be \( S \)-valued random variables on \( (\Omega, F, Q), (\Omega^{(n)}, F^{(n)}, Q^{(n)}) \), respectively. Then \( (\xi^{(n)}, \tilde{\xi}^{(n)}) \xrightarrow{d(Q^{(n)})} (\xi, \tilde{\xi}) \) on \( S \times \tilde{S} \) if and only if

\[ E^{(n)}[F(\xi^{(n)}) \tilde{F}(\tilde{\xi}^{(n)})] \to E[Q(F(\xi) \tilde{F}(\tilde{\xi}))] \]

for every \( F \in C_b(S) \) and \( \tilde{F} \in C_b(\tilde{S}) \).
Theorem 5.5. Let $\eta_k(n)$ be $S$-valued random variables defined, for every $n$ and $k$, on $(\Omega^{(n)}, \mathcal{F}^{(n)}, Q^{(n)})$. Assume that $\eta_k(n) \xrightarrow{d(Q^{(n)})} \eta_k$ and that $\eta_k \xrightarrow{d(Q)} \xi$. Then $\xi^{(n)} \xrightarrow{d(Q^{(n)})} \xi$ as soon as, for every $\epsilon > 0$,

$$\lim_{k \to \infty} \limsup_n Q^{(n)}(d(\eta_k^{(n)}, \xi^{(n)}) > \epsilon) < \epsilon.$$  

Proof. We use the third condition in the Portmanteau theorem. For any closed $K \subseteq S$ and $\epsilon > 0$, we have

$$Q^{(n)}(\xi^{(n)} \in F) \leq Q^{(n)}(\eta_k^{(n)} \in F) + Q^{(n)}(d(\eta_m^{(n)}, \xi^{(n)}) > \epsilon)).$$

Thus

$$\limsup_n Q^{(n)}(\xi^{(n)} \in F) \leq Q(\eta_k \in F) + \limsup Q^{(n)}(d(\eta_m^{(n)}, \xi^{(n)}) > \epsilon),$$

and letting first $\epsilon$ tend to zero and then $k$ to infinity proves the claim, since $1_{\eta_k \in F} \to 1_{\eta \in F}$ almost surely.

Theorem 5.6. Assume that $a^{(n)}, \xi^{(n)}, \xi$ are $S$-valued random variables such that $\xi^{(n)} \xrightarrow{d(Q^{(n)})} \xi$ and that $a^{(n)} \xrightarrow{Q^{(n)}} a$ for some constant $a$. Then

1. $(\xi^{(n)}, a^{(n)}) \xrightarrow{d(Q^{(n)})} (\xi, a)$.

2. (Slutsky’s theorem) If $Y = C$ or $Y = \mathbb{R}$, then $a^{(n)} + b^{(n)} \xi^{(n)} \xrightarrow{d(Q^{(n)})} a + b \xi$ for any sequence of $Y$-valued random variables such that $b^{(n)} \xrightarrow{Q^{(n)}} b$ for some constant $b$.

Guide: To prove 1., apply Theorem 5.4 and Theorem 5.5. To prove 2., apply the first part and the continuous mapping theorem, and use the fact the multiplication and addition are continuous mappings on $Y$. The details are left as an exercise.

Theorem 5.7 (The Skorokhod representation theorem). Assume in addition that $S$ is separable. Then there exists a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, with $S$-valued random variables $\tilde{\xi}$, $\xi^{(n)}$, $\nu = 1, 2, \ldots$ with the same distributions as $\xi, \xi^{(n)}$, respectively, and such that $\tilde{\xi}^{(n)}(\omega) \rightarrow \tilde{\xi}(\omega)$ $\tilde{P}$-almost surely.

We end this this section that relates tightness of distributions to relative compactness. The sequence $(\xi^{(n)})_{n=1}^{\infty}$ is relatively compact in distribution is if its every subsequence has a further subsequence that convergences in distribution. A sequence $(\xi^{(n)})_{n=1}^{\infty}$ is tight if

$$\sup_K \liminf_{2^n} Q^{(n)}(\xi^{(n)} \in K) | K \subset S \text{ is compact} = 1.$$
**Theorem 5.8.** The sequence \((\xi^\nu)_{\nu=1}^\infty\) is tight if and only if it is relatively compact in distribution.

**Example 5.9.** A single \(S\)-valued random variable \(\xi\) is tight in the sense that
\[
\sup_K \{P(\xi \in K) \mid K \subset C \text{ is compact}\} = 1.
\]

This follows from the property that Borel-measures on complete separable metric spaces are inner regular in the sense that measures of Borel sets can be approximated from inside by compact sets precisely in the above sense.
Exercises

Exercise 5.1.1. Consider the Portmanteau theorem in the case of $S = \mathbb{R}^1$. Denoting the cumulative distribution functions by $\Psi^{(n)}(x) := Q^{(n)}(\xi^{(n)} \leq x)$ and $\Psi^{(n)}(x) := Q(\xi \leq x)$, $x \in \mathbb{R}$, show that the conditions in the Portmanteau theorem are equivalent to

$$
\Psi^{(n)}(x) \to \Psi(x) \quad \text{for all } x \text{ with } \Delta \Psi(x) = 0,
$$

where $\Delta \Psi(x) = \Psi(x) - \lim_{x' \nearrow x} \Psi(x')$ is the jump of $\Psi$ and $x$.

Assume that the $S$-valued random variables $\eta, \eta_k, k = 1, 2, \ldots$ are defined on the same probability space with a probability measure $Q$. Recall that the sequence $(\eta_k)_{k=1}^{\infty}$ convergences in probability to $\eta$, denoted by $\eta_k \overset{Q}{\to} \xi$, if

$$
\lim_k Q(d(\eta_k, \eta) > \epsilon) = 0,
$$

which can be written equivalently as

$$
\lim_k E_Q[\max\{d(\eta_k, \eta), 1\}] = 0.
$$

Exercise 5.1.2. Show that $\eta_k \overset{Q}{\to} \eta$ implies $\eta_k \overset{d(Q)}{\to} \eta$, and the two conditions are equivalent when $\eta$ is almost surely a constant. Hint: A sequence converging in probability has a subsequence that converges almost surely.

Exercise 5.1.3. Prove Theorem 5.6.
5.2 Convergence of continuous stochastic processes

We consider continuous time processes only on the "time interval" [0, 1]. A family \((y_t)_{t \in [0, 1]}\) of \(\mathbb{R}\)-valued random variables is called an \(\mathbb{R}\)-valued continuous time stochastic process. Given \(\omega \in \Omega\), the function \(t \mapsto y_t(\omega)\) is called as a path, or a trajectory or a realization, of the process \(y\). Instead of considering a stochastic process as an indexed family of \((\mathbb{R}^d)\)-valued random variables, one may thus think of stochastic process as a function-valued random variable, or as a random path, trajectory, etc. If the paths of a continuous time process are \(P\)-almost surely continuous, then the process is called continuous.

We denote the space of continuous functions with \(C := C([0,1])\) and equip it with the supremum norm

\[
\|y\| := \sup_{t \in [0,1]} |y_t|.
\]

The space \(C\) is a Banach space with respect to this norm, so in particular it is a complete separable metric space. The modulus of continuity is the function

\[
w_h(y) := \sup \{ |y_{t'} - y_t| : |t - t'| \leq h \} \quad y \in C, h \in \mathbb{R}_+.
\]

The modulus of continuity \(w_h(y)\) is decreasing w.r.t. \(h\) and continuous w.r.t. \(y\), and\(^6\) \(\lim_{h \to 0} w_h(y) = 0\).

The mapping \(\pi_t y = y_t\) from \(C\) to \(\mathbb{R}\) is the valuation mapping at \(t\). We recall the following famous theorem from real analysis.

**Theorem 5.10.** [Arzelà Ascoli]. Let \(D\) be dense in \([0, 1]\). Then a set \(K \subset C\) is relatively compact iff \(\pi_t K\) is relatively compact in \(\mathbb{R}\) for every \(t \in D\) and

\[
\lim_{h \to 0} \sup_{y \in K} w_h(y) = 0.
\]

The following lemma shows that continuous random processes are exactly the measurable random variables from \((\Omega, \mathcal{F})\) to \((C, \mathcal{B}(C))\).

**Lemma 5.11.** The Borel-\(\sigma\)-algebra \(\mathcal{B}(C)\) of \(C\) satisfies

\[
\mathcal{B}(C) = \sigma(\{\pi_t; t \in [0,1]\}).
\]

In particular, each continuous stochastic process \(y\) and probability measure \(Q\) on \((\Omega, \mathcal{F})\) defines a measure \(\text{Law}(y; Q)\) on \(C\) via

\[
(\text{Law}(y; Q))(A) = Q(\{\omega : y(\omega) \in A\}).
\]

Like the notation suggests, this is called the law of \(y\) under \(Q\) on \(C\). A law on \(C\) is determined by finite dimensional distributions in the following sense.

\(^6\) Extending the modulus of continuity to noncontinuous functions, \(\lim_{h \to 0} w_h(y)\) gives the size of the largest jump of \(y\).
Theorem 5.12. Let $y$ and $\tilde{y}$ be continuous stochastic processes defined from $(\Omega, F, Q)$ and $(\tilde{\Omega}, \tilde{F}, \tilde{Q})$. Then

$$\text{Law}(y; Q) = \text{Law}(\tilde{y}, \tilde{Q})$$

if and only if, for all $k \in \mathbb{N}$ and for all $0 \leq t_1, \ldots, t_k \leq 1$,

$$\text{Law}((y_{t_1}, \ldots, y_{t_k}); Q) = \text{Law}((\tilde{y}_{t_1}, \ldots, \tilde{y}_{t_k}); \tilde{Q}).$$

Proof. Necessity is obvious from Lemma 5.11. To prove the sufficiency, we use the monotone class theorem, some details are left in the exercises. Let $D = \{B \in \mathcal{B}(C) \mid Q(y \in B) = \tilde{Q}(\tilde{y} \in B)\}$ and $A$ consist of all sets

$$A = \{y \in C \mid (y_{t_1}, \ldots, y_{t_k}) \in E\} \quad 0 \leq t_1, \ldots, t_k \leq 1, E \in \mathcal{B}(\mathbb{R}^k).$$

Then $A$ is a $\pi$-system (i.e., it is closed under finite intersections) and $D$ is a $\pi$-system (i.e., $C \in A$, $(A \setminus \tilde{A}) \in A$ whenever $A, \tilde{A} \in A$ with $\tilde{A} \subset A$, and $A_1, A_2, \ldots \in A$ implies $(\bigcup_i A_i) \in A$). By assumption $A \subset D$, so the monotone class theorem gives $\sigma(A) \subset D$. Thus the result follows from Lemma 5.11. \qed

We will need Prohorov’s theorem that characterizes the convergence in distribution in terms of finite dimensional distributions, tightness, and modulus of continuity.

Denoting $C_b(\mathbb{R}^k)$ the set of continuous bounded function on $\mathbb{R}^k$, the sequence $(y^\nu)^\infty_{\nu=1}$ of stochastic processes is said to converge in finite dimensional distributions to $y$, denoted by $y^\nu \overset{fd}{\rightarrow} y$ if,

$$EF(y^\nu_{t_1}, \ldots, y^\nu_{t_k}) \rightarrow EF(y_{t_1}, \ldots, y_{t_k}), \forall k \in \mathbb{N}, 0 \leq t_1 \leq \cdots \leq t_k \leq 1, F \in C_b(\mathbb{R}^k).$$

We will analyze convergence of finite dimensional distributions in more detail in the next section. Here we note that finite dimensional random variables may be viewed as continuous stochastic processes by, instead of $[0, 1]$, considering a finite set $S$ and equipping it with the discrete metric. When we write results concerning both continuous processes and finite dimensional random variables, we write that $y$ takes values in $C(S)$, where $S = [0, 1]$ or $S$ is a finite set, respectively.

We remark that $\omega \mapsto w_h(y(\omega))$ is measurable for every continuous stochastic process $y$.

Theorem 5.13 (Prohorov’s theorem). Let $y, y^\nu, \nu = 1, 2, 3, \ldots$ be $C(S)$-valued random variables.

1. $y^\nu \overset{d(Q^{(\infty)})}{\rightarrow} y$ if and only if $y^\nu \overset{fd(Q^{(\infty)})}{\rightarrow} y$ and $(y^\nu)^\infty_{\nu=1}$ is relatively compact in distribution.

2. $y^\nu \overset{d(Q^{(\infty)})}{\rightarrow} y$ if and only if $y^\nu \overset{fd(Q^{(\infty)})}{\rightarrow} y$ and $(y^\nu)^\infty_{\nu=1}$ is tight.
3. If $C(S) = C$ and $y^{n} \overset{fd}{\longrightarrow} y$, then $(y^{n})_{n=1}^{\infty}$ is tight if and only if
\[
\lim_{h \to 0} \limsup_{\nu \to \infty} E^{Q^{\nu}} [w_{h}(y^{\nu}) \wedge 1] = 0. \tag{5.1}
\]

Proof. To prove 1., assume first that $y^{\nu} \overset{d(Q^{(n)})}{\longrightarrow} y$. For any $(t_{1}, \ldots, t_{k})$, the mapping $y \mapsto (y_{1}, \ldots, y_{k})$ is continuous. Thus by the continuous mapping theorem, Theorem 5.3, $y^{\nu} \overset{fd(Q^{(n)})}{\longrightarrow} y$. The relative compactness in distribution is trivial.

Assume now that $y^{\nu} \overset{fd(Q^{(n)})}{\longrightarrow} y$ and that $(y^{n})_{n=1}^{\infty}$ is relatively compact in distribution, but assume to the contrary that we do not have $y^{n} \overset{d(Q^{(n)})}{\longrightarrow} y$. Then there exists $f \in C_{b}(C)$, $\epsilon > 0$, and some subsequence (still indexed by the same $n$) such that $|E^{Q^{(n)}} f(y^{(n)}) - E^{Q} f(y)| > \epsilon$. By the relative compactness in distribution, there exists a continuous stochastic process $\tilde{y}$ such that $y^{(n)} \overset{d(Q^{(n)})}{\longrightarrow} \tilde{y}$ along a further subsequence. Since finite dimensional distributions converge to $y$ by assumption, we have $Law(y) = Law(\tilde{y})$ by Theorem 5.12. Thus $y^{(n)} \overset{d(Q^{(n)})}{\longrightarrow} y$ along the previously chosen further subsequence, which is a contradiction with $|E^{Q^{(n)}} f(y^{(n)}) - E^{Q} f(y)| > \epsilon$.

The second claim follows from the first one and Theorem 5.8.

To prove 3., assume first that $(y^{\nu})_{\nu=1}^{\infty}$ is tight. Given an $\epsilon > 0$, there is a compact $K \subset C$ such that $\limsup Q^{(n)}(y^{(n)} \in K^{C}) < \epsilon$. By the Arzelà Ascoli theorem, there is $h > 0$ such that $w_{h} \leq \epsilon$ on $K$ so that $Q^{(n)}(w_{h}(y^{(n)}) \geq \epsilon) < \epsilon$. Since $\epsilon > 0$ was arbitrary, we get (5.1).

Assume now that (5.1) holds. We have that $w_{h}(y^{(n)}) \downarrow_{\nu} 0$ $Q^{(n)}$-almost surely for each $h$, so $\limsup$ can be replaced by $\sup$ in (5.1). Thus, given an $\epsilon > 0$, there exist numbers $h^{k} > 0$ such that
\[
Q^{(n)}(w_{h^{k}}(y^{(n)}) > 2^{-k}) \leq 2^{-k-1}\epsilon \quad k = 1, 2, \ldots.
\]
Let $D$ be a dense set in $[0, 1]$. For every $t_{k} \in D$, the sequence $(y^{(n)}_{t_{k}})_{n=1}^{\infty}$ is tight by 2. so, there are compact sets $A_{k} \in \mathbb{R}$ such that
\[
Q^{(n)}(y_{t_{k}}^{(n)} \notin A_{k}^{C}) \leq 2^{-k-1}\epsilon \quad k = 1, 2, \ldots.
\]
Defining
\[
K = \bigcap_{k} \{w \in C \mid w_{t_{k}} \in A_{k}, w_{h^{k}}(w) \leq 2^{-k}\},
\]
we get from the Arzelà Ascoli theorem that $K$ is compact. Since $\sup_{n} Q^{(n)}(y^{(n)} \notin K) \leq \epsilon$, we have tightness of $(y^{(n)})_{n=1}^{\infty}$.

The following is Kolmogorov’s tightness criterion. Usually it is stated for all $t, t' \in [0, 1]$, not just for dyadic points
\[
D := \bigcup_{n} D_{n},
\]
where \( D_n \) is the \( n \)-th dyadic partition of \([0,1]\),

\[
D_n := \{i/2^n \mid i = 0, \ldots, 2^n\}.
\]

Moreover, Kolmogorov's tightness criterion would also give Hölder continuity of the processes as well as of the limiting process. However, we do not prove this fact here.

**Theorem 5.14** (Kolmogorov’s tightness criterion). Let \( y^\nu, \nu = 1, 2, 3, \ldots \) be continuous stochastic processes with \( y^\nu_0 = y_0 \) for some constant \( y_0 \) and

\[
E |y^\nu_t - y^\nu_{t'}|^a \leq L |t - t'|^{1+b} \quad \forall t, t' \in D, \ \nu = 1, 2, \ldots
\]

for some constants \( a \geq 1, b, L > 0 \). Then the sequence \((y^\nu)_{\nu=1}^\infty\) is tight.

**Proof.** Let \( h = 2^{-n} \). Let \( y \) be a continuous stochastic process. Denoting

\[
\xi_k = \max_{t, t' \in D_k} \{|y_t - y_{t'}| \mid |t - t'| \leq 2^{-k}\},
\]

we have, for some constant \( L \),

\[
w_{2^{-n}}(y) = \sup_{t, t'} \{|y_t - y_{t'}| \mid |t - t'| \leq 2^{-n}\}
= \sup_k \max_{t, t' \in D_k} \{|y_t - y_{t'}| \mid |t - t'| \leq 2^{-n}\}
\leq L \sum_{k \geq n} \xi_k,
\]

where the inequality is left as an exercise. Note that

\[
E \xi_k^a \leq \sum_{t, t' \in D_k, |t - t'| \leq 2^{-k}} E |y_t - y_{t'}|^a;
\]

so, by (5.2), \( \xi_k \) corresponding to \( y^\nu \) satisfies \( E \xi_k^a \leq L 2^k 2^{-k(1+b)} = L 2^{-kb} \) (for some other constant \( L \)). In conjunction with the monotone convergence theorem and Hölder’s inequality, this implies

\[
E \sum_{k \geq n} \xi_k = \sum_{k \geq n} E \xi_k \leq \sum_{k \geq n} (E \xi_k^a)^{1/a} \leq L^{1/a} \sum_{k \geq n} 2^{-kb/a} \leq M 2^{-nb/a}
\]

for some constant \( M > 0 \). Combining the inequalities we get

\[
\lim_{h \to 0} \limsup_{\nu \to \infty} E [w_h(y^\nu) \wedge 1] \leq \lim_{n \to \infty} \limsup_{\nu \to \infty} E [w_{2^{-n}}(y^\nu)] \leq \lim_{n \to \infty} M 2^{-nb/a} = 0.
\]

\[\square\]
5.2.1 Exercises

Recall that a sequence \((\xi^{(n)})_{n=1}^{\infty}\) of \(-\)-valued random variables in \(L^1\) is \textit{uniformly integrable} if

\[
\lim_{M \to \infty} \limsup_{n \to \infty} E[|\xi^{(n)}| 1_{|\xi^{(n)}| \geq M}] = 0.
\]

**Exercise 5.2.1.** Assume that \((\xi^{(n)})_{n=1}^{\infty}\) is a sequence of nonnegative \(-\)-valued random variables in \(L^1\) such that \(\xi^{(n)} \xrightarrow{d(P)} \xi\). Show that \(\liminf E\xi^{(n)} \geq E\xi\). Show also that \(E\xi^{(n)} \to E\xi\) if and only if \((\xi^{(n)})_{n=1}^{\infty}\) are uniformly integrable.

Hint: "cut-off" functions of the form \(F(x) = x^+ \wedge r\) and \(F(x) = x^+ (r-x)^+\), \(r > 0\), are bounded and continuous.

**Exercise 5.2.2.** Prove that, for \(k\)-dimensional random vectors \(\xi^{(n)}\), we have that \((\xi^{(n)}) \xrightarrow{d(P)} \xi\) if and only if \(\Delta\xi^{(n)} \xrightarrow{d(P)} \Delta\xi\), where \(\Delta\xi = (\xi_1, \xi_2 - \xi_1, \ldots, \xi_k - \xi_{k-1})\).

**Exercise 5.2.3.** Prove the inequality (5.3) when \(y\) is a deterministic continuous function on \([0, 1]\).

**Exercise 5.2.4.** The continuous time process

\[
Y_t^{(n)} := S_{\lfloor 2^n t \rfloor}^{(n)} \quad t \in [0, 1]
\]

defines a piecewise constant "continuous time extension" of the \(n\)-th market discrete time market model. Show that, for every \(t\),

\[
W_t^{(n)}(\xi^{(n)}) - Y_t^{(n)}
\]

converge to zero in probability as \(n\) tends to infinity.
5.3 Convergence of the finite dimensional distributions

In this section we give conditions in order to get that the finite dimensional distributions of the sequence \((s(n))_{n=1}^{\infty}\) converge to a geometric Brownian motion. For this purpose, we will

- recall some facts on weak convergence of finite dimensional distributions, and central limit theorem,
- approximate Brownian motion by a scaled symmetric random walk,
- approximate geometric Brownian motion by a binomial tree.

In the finite dimensional case, the weak convergence of random variables is classically given in terms of the cumulative distribution functions. A cumulative distribution function \(\Psi\) is continuous at point \(x\), if \(\Delta \Psi(x) = \Psi(x) - \Psi(x-) = 0\); here \(\Psi(x-) = \lim_{y \downarrow x, y \to x} \Psi(y)\).

**Lemma 5.15.** Let, for each \(n = 1, 2, \ldots\), the random variable \(\eta^n\) be defined on \((\Omega^n, F^n, Q^n)\). Then \(\text{Law}(\eta^n, Q^n) \xrightarrow{d} \Psi\), if at every continuity point \(x\) of \(\Psi\) we have

\[ Q^n(\eta^n \leq x) \to \Psi(x), \text{ as } n \to \infty. \tag{5.4} \]

We consider the case, where the limit is the cumulative distribution function of a normal random variable:

\[ \Phi_{\mu, \sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{-\frac{1}{2\sigma^2}(y-\mu)^2} dy, \]

where \(\mu \in \mathbb{R}, \sigma > 0\). If \(\xi \sim N(\mu, \sigma^2)\), then \(E\xi = \mu\) and \(\text{Var}(\xi) = \sigma^2\). If \(\mu = 0\) and \(\sigma = 1\), then the random variable \(\xi\) has the standard normal distribution. If \(\xi \sim N(0, 1)\), then \(\mu + \sigma \xi \sim N(\mu, \sigma^2)\). We use the notation

\[ \eta^n \xrightarrow{d(Q^n)} \Phi_{\mu, \sigma^2}, \]

if \((\eta^n, Q^n)\) converge in distribution to a random variable with the distribution function \(\Phi_{\mu, \sigma^2}\).

We will need the following variant of the central limit theorem. In the classical one \((\xi_k)_{k=1}^{\infty}\) is the same sequence for all \(n\). The proof, that uses characteristic functions (Fourier transform), of our variant is exactly the same.

**Proposition 5.16 (Central limit theorem).** Let

\[ Y_1^n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \xi_k^{(n)}, \]

for an \(Q^n\)-i.i.d. \((\xi_k^{(n)})_{k=1}^{\infty}\) sequence of random variables with \(E^{Q^n} \xi_k^{(n)} = 0\) and \(E^{Q^n} (\xi_k^{(n)})^2 = 1\). We have

\[ Y_1^n \xrightarrow{d(Q^n)} \Phi_{0,1} \tag{5.5} \]
5.3.1 Convergence of random walks to a Brownian motion

Definition 5.17. A continuous stochastic process \( W \) is a Brownian motion, if it has independent increments: for all \( 0 \leq t_0 < t_1 < \cdots < t_n \) the random variables \( W_{t_i} - W_{t_{i-1}} \) are independent, and \( W_t - W_s \sim \mathcal{N}(0, t-s) \) for all \( 0 \leq s < t \leq 1 \).

Theorem 5.18. Let

\[
y^{(n)}_t = \frac{1}{\sqrt{2^n}} \sum_{k=1}^{2^n} \xi^{(n)}_k \quad t = 0, \ldots, 2^n
\]

for an \( Q^{(n)} \)-i.i.d. \( (\xi^{(n)}_k)_{k=1}^{\infty} \) sequence of random variables with \( \mathbb{E}_{Q^{(n)}} \xi^{(n)}_k = 0 \) and \( \mathbb{E}_{Q^{(n)}} (\xi^{(n)}_k)^2 = 1 \). Then

\[
\mathcal{W}(y^{(n)}) \xrightarrow{fd} W
\]

for a Brownian motion \( W \).

Proof. The continuous time process given by

\[
Y^{(n)}_t := \frac{1}{\sqrt{2^n}} \sum_{k=1}^{\lfloor 2^n t \rfloor} \xi^{(n)}_k \quad t = 0, \ldots, 2^n
\]

defines a piecewise linear extension (in time) of \( (k/2^n, \xi^{(n)}_k) \in [0,1] \times \mathbb{R}, k = 1, \ldots, 2^n \). Since \( \mathcal{W}(y^{(n)}) \) is the linearly interpolated extension, we have that the differences

\[
\mathcal{W}(\xi^{(n)})_t - Y^{(n)}_t = (2^n t - \lfloor 2^n t \rfloor) \frac{\xi^{(n)}_{\lfloor 2^n t \rfloor + 1}}{\sqrt{2^n}}
\]

converge to zero in probability as \( n \) tends to infinity. By Slutsky’s theorem, it thus suffices to show that \( Y^{(n)} \xrightarrow{fd} W \).

Using Slutky’s theorem, the central limit theorem and the fact that \( \frac{\lfloor nt \rfloor}{n} \to t \) when \( n \to \infty \), we get

\[
Y^{(n)}_t = \frac{\sqrt{2^n}}{\sqrt{2^n}} \sum_{k=1}^{\lfloor 2^n t \rfloor} \xi^{(n)}_k \xrightarrow{d} \Phi_{0,t},
\]

as \( n \to \infty \). Let now \( t < u \). The random variables \( Y^{(n)}_u - Y^{(n)}_t \) are independent from the variables \( Y^{(n)}_t \), since

\[
Y^{(n)}_u - Y^{(n)}_t = \frac{1}{\sqrt{2^n}} \sum_{k=\lfloor 2^n t \rfloor + 1}^{\lfloor 2^n u \rfloor} \xi^{(n)}_k
\]
and the random variables $\xi_k^{(n)}$ are independent. Repeating the previous arguments we get

$$Y_u^{(n)} - Y_t^{(n)} \xrightarrow{d(P^n)} \Phi_{0, u-t}.$$ 

We observe that the variables $\Delta Y_{t_i}^{(n)} := Y_{t_i}^{(n)} - Y_{t_{i-1}}^{(n)}$ are mutually independent for all $0 \leq t_0 < t_1 < \cdots < t_n \leq T$. Thus the process $Y^{(n)}$ has independent increments, and we have that

$$Q^n \left\{ Y_{t_i}^{(n)} - Y_{t_{i-1}}^{(n)} \leq x_i, i = 1, \ldots, n \right\} \rightarrow \prod_{i=1}^n \Phi_{0, t_i-t_{i-1}}(x_i) = P\{ W_{t_i} - W_{t_{i-1}} \leq x_i, i = 1, \ldots, n\}.$$

This finishes the proof, since all the finite dimensional distributions of the processes $Y^{(n)}$ converge to the finite dimensional distributions of the process $W$; see exercise ??.

Using Slutsky’s theorem, the following theorem is proved like Theorem 5.18 above.

**Theorem 5.19.** Let

$$y_t^{(n)} = \sum_{k=1}^t \left( \frac{\mu^{(n)}}{2^n} + \frac{\sigma^{(n)}}{\sqrt{2^n}} \xi_k^{(n)} \right), \quad t = 0, \ldots, 2^n$$

for an $Q^n$-i.i.d. $(\xi_k^{(n)})_{k=1}^\infty$ sequence of random variables with $E^{Q^n} \xi_k^{(n)} = 0$ and $E^{Q^n} (\xi_k^{(n)})^2 = 1$, and deterministic sequences $(\mu^{(n)})_{n=1}^\infty$, $(\sigma^{(n)})_{n=1}^\infty$ with $\mu^{(n)} \rightarrow \mu$ and $\sigma^{(n)} \rightarrow \sigma$. Then

$$W(y^{(n)}) \xrightarrow{f_d(Q^n)} W^{\mu, \sigma}$$

for a Brownian motion $W^{\mu, \sigma}_t := \mu t + \sigma W_t$.

### 5.3.2 Convergence of a geometric binomial tree to a geometric Brownian motion

**Theorem 5.20.** Let

$$S_t^{(n)} = S_0 \prod_{k=1}^{2^n t} (1 + R_k^{(n)}),$$

for a $Q^n$-i.i.d. sequence $(\xi_k^{(n)})_{k=1}^{2^n}$ with $E^{Q^n} \xi_k^{(n)} = 0$ and $(\xi_k^{(n)})^2 \equiv 1$ and

$$R_k^{(n)} = \frac{\mu}{2^n} + \frac{\sigma}{\sqrt{2^n}} \xi_k^{(n)}.$$
Then
\[ S^{(n)} \xrightarrow{fd(Q^{(n)})} S^{\mu,\sigma} \]
for a geometric Brownian motion \( S^{\mu,\sigma}_t = S_0 \exp(\mu t + \sigma W_t - \frac{\sigma^2}{2} t) \).

Proof. We write \( S^{(n)}_t \) as
\[ S^{(n)}_t = S_0 \prod_{k=1}^{\lfloor 2^n t \rfloor} (1 + R_k^{(n)}) = S_0 \exp\{ \sum_{k=1}^{\lfloor 2^n t \rfloor} \log(1 + R_k^{(n)}) \}. \]

Next we study the convergence of the finite dimensional distributions of the sum \( \sum_{k=1}^{\lfloor n t \rfloor} \log(1 + R_k^{(n)}) \). Using the Taylor expansion for the logarithm: \( \log(1 + x) = x - \frac{x^2}{2} + O(x^3) \), we get
\[ \sum_{k=1}^{\lfloor 2^n t \rfloor} \log(1 + R_k^{(n)}) = \sum_{k=1}^{\lfloor 2^n t \rfloor} \left( \frac{\mu}{2^n} + \frac{\sigma}{\sqrt{2^n}} \xi_k^{(n)} \right) - \frac{1}{2} \sum_{k=1}^{\lfloor 2^n t \rfloor} \left( \frac{\mu}{2^n} + \frac{\sigma}{\sqrt{2^n}} \xi_k^{(n)} \right)^2 + \sum_{k=1}^{\lfloor 2^n t \rfloor} O((\frac{\mu}{2^n} + \frac{\sigma}{\sqrt{2^n}} \xi_k^{(n)})^3) \]
\[ =: I + II + III. \]

Sum I

We observe that
\[ \sum_{k=1}^{\lfloor 2^n t \rfloor} \left( \frac{\mu}{2^n} + \frac{\sigma}{\sqrt{2^n}} \xi_k^{(n)} \right) \xrightarrow{df(Q^{(n)})} \Phi_{\mu,\sigma^2}, \]
by the results of the previous section. We can write this using the Brownian motion \( W_t^{\mu,\sigma} := \mu t + \sigma W_t \) as
\[ \sum_{k=1}^{\lfloor 2^n t \rfloor} \left( \frac{\mu}{2^n} + \frac{\sigma}{\sqrt{2^n}} \xi_k^{(n)} \right) \xrightarrow{df(Q^{(n)})} W^{\mu,\sigma}. \]

Sum II

Next we will consider the limit of the sum \( \sum_{k=1}^{\lfloor 2^n t \rfloor} \left( \frac{\mu}{2^n} + \frac{\sigma}{\sqrt{2^n}} \xi_k^{(n)} \right)^2 \). Note first that \( \lfloor 2^n t \rfloor (\frac{\mu}{2^n})^2 \to 0 \). Since, for every \( k \), \( (\xi_k^{(n)})^2 = 1 \), we obtain that
\[ \sum_{k=1}^{\lfloor 2^n t \rfloor} \left( \frac{\sigma}{\sqrt{2^n}} \xi_k^{(n)} \right)^2 = \frac{\lfloor 2^n t \rfloor}{2^n} \sigma^2 Q^{(n)} \to \sigma^2 t. \]
For the product \( \frac{\sigma}{\sqrt{2n}} \xi_k^{(n)} \), we have

\[-\frac{\sigma^2}{2n^{3/2}} \leq \frac{\sigma}{\sqrt{2n}} \xi_k^{(n)} \leq \frac{\sigma}{2n^{3/2}},\]

so

\[\sum_{k=1}^{[nt]} \left( \frac{\sigma}{\sqrt{n}} n \xi_k^{(n)} \right) \xrightarrow{Q(n)} 0.\]

Combining these three observations, we get

\[\sum_{k=1}^{[nt]} \left( \frac{\mu}{n} + \frac{\sigma}{\sqrt{n}} \xi_k^{(n)} \right)^2 \xrightarrow{Q(n)} \sigma^2 t.\] (5.6)

**Remark 5.21.** Note that the convergence in (5.6) does not depend on the measure \( P \) at all.

**Sum III**

The expression \(|\frac{\mu}{2n} + \xi_k \frac{\sigma}{\sqrt{2n}}|^3\) can be dominated by

\[\left| \frac{\mu}{2n} + \xi_k \frac{\sigma}{\sqrt{2n}} \right|^3 \leq \frac{K(\mu, \sigma)}{2n^{3/2}},\]

for example with the constant \( K(\mu, \sigma) = |\mu|^3 + 3\mu^2\sigma + 3|\mu|\sigma^2 + \sigma^3 \). Using this estimate, we get that

\[\sum_{k=1}^{[nt]} O((\frac{\mu}{n} + \frac{\sigma}{\sqrt{n}})^3) \xrightarrow{Q(n)} 0\] (5.7)

**Remark 5.22.** Note that the convergence in (5.7) does not depend on the measure \( P \) at all.
5.3.3 Exercises

For a stochastic process \( y \), we define quadratic variation and the \( Q \)-predictable quadratic variation, respectively as

\[
[y, y]_t := \sum_{t' = 0}^{t} (y_{t'} - y_{t' - 1})^2, \quad t = 0, 1, \ldots, T
\]

\[
\langle y, y \rangle_t := \sum_{t' = 0}^{t} E_{t'}^Q (y_{t'} - y_{t' - 1})^2, \quad t = 0, 1, \ldots, T,
\]

where \( y_{-1} := 0 \) and \( F_{-1} := F_0 \).

Exercise 5.3.1. Let \( (m_t)_{t=0}^{T} \) be a square integrable \( Q \)-martingale. Show that

1. The process defined by \( y_t := m_t^2 - \langle m, m \rangle_t \) is a \( Q \)-martingale.

Exercise 5.3.2. Let \( (\xi_k)_{k=1}^{\infty} \) be an i.i.d. sequence of nonnegative random variables that have the same law as the random variable \( \eta \), whose cumulative distribution function is \( P(\eta \leq x) = e^{-x^2}, \quad x \geq 0 \). Let

\[
\xi_n^* := n^{-1/2} \max \{\xi_1, \ldots, \xi_n\},
\]

\[
s_k(n) := \frac{1}{\sqrt{2^n}} \xi_k, \quad k = 1, \ldots, 2^n,
\]

and let \( f : C \to \mathbb{R} \) be the continuous mapping \( f(y) := \|y\| = \sup_{t \in [0,1]} |y_t| \).

1. Show that \( \xi_n^* \) has the same law as \( \eta \).

2. What is the random variable \( f(\mathcal{W}_{(n)}(s^{(n)})) \)?

Exercise 5.3.3. Continuing 5.3.2,

1. show that \( \mathcal{W}_{(n)}(s^{(n)}) \xrightarrow{d} 0 \).
   Hint: Show that all the finite dimensional vectors of the form

\[
(\mathcal{W}_{(n)}(s^{(n)}))_{t_0}, \ldots, \mathcal{W}_{(n)}(s^{(n)})_{t_k})
\]

converge to zero in probability.

2. combine the observations of these two exercises to conclude that \( \mathcal{W}(s^{(n)}) \) cannot converge in distribution (though the finite dimensional distributions converge).
   Hint: Apply the continuous mapping theorem to arrive at contradiction.
5.4 The weak convergence of the price processes to the Black & Scholes Model

We work with the binomial tree structure introduced in Section 2.3: the discounted price process at the $n$-th stage is of the form

$$s^{(n)}_t = s_0 \prod_{t' = 1}^t \frac{1 + R^{(n)}_{t'}}{1 + r/2^n} \quad \forall \ t = 1, \ldots, 2^n \quad (5.8)$$

where $r > -1$ is the interest rate, and

$$R^{(n)}_t = \frac{\mu}{2^n} + \left( \frac{\sigma}{\sqrt{2^n}} \right) \xi_t$$

for a sequence $(\xi_t)_{t=1}^T$ of independent identically distributed essentially bounded random variables with $E\xi_t = 0$ and $\xi_t^2 = 1$. By Theorem 2.8 under a martingale measure $Q^{(n)}$ of $s^{(n)}$, $\xi_t$ are also i.i.d. and the martingale measure has the density process

$$q^{(n)}_t = s_0 \prod_{t' = 1}^t \left( 1 + \frac{1}{\sqrt{2^n}} \frac{r - \mu}{\sigma} \xi_{t'} \right) \quad \forall \ t = 1, \ldots, 2^n. \quad (5.10)$$

It is an exercise to check that, under the martingale measures, we have

$$E^{Q^{(n)}}(\xi_t) = \frac{1}{\sqrt{2^n}} \frac{r - \mu}{\sigma},$$

$$\text{Var}^{Q^{(n)}}(\xi_t) = 1 - E\xi_t^3 \frac{1}{\sqrt{2^n}} \frac{r - \mu}{\sigma} - \frac{1}{2^n} \left( \frac{r - \mu}{\sigma} \right)^2.$$

Thus, we may write

$$s^{(n)}_t = s_0 \prod_{t' = 1}^t \left( 1 + \frac{\sigma^{(n)}}{\sqrt{2^n}} \xi^{(n)}_{t'} \right) \quad \forall \ t = 1, \ldots, 2^n, \quad (5.11)$$

for an $Q^{(n)}$-i.i.d. sequence $(\xi^{(n)})_{t=1}^{2^n}$ with $E^{Q^{(n)}}(\xi^{(n)}_t) = 0$, $E^{Q^{(n)}}(\xi^{(n)}_t)^2 = 1$, and $\sigma^{(n)} \to \sigma$. Applying the results of the previous section, we get

$$\mathcal{W}(s^{(n)}) \xrightarrow{d(Q^{(n)})} \exp \left( \sigma W_t - \frac{\sigma^2 t}{2} \right)$$

for a Brownian motion $W$. It is an exercise to verify that the limit process is indeed a martingale. It is called a geometric Brownian motion.

To pass from the discounted price process to the original one, note that

$$(1 + r/2^n)|2^nt| \to e^{rt},$$

so Slutsky’s lemma gives that the nondiscounted price process converges under the martingale measure to

$$\exp \left( \sigma W_t - \frac{\sigma^2 t}{2} + rt \right)$$

44
Let \( S_t = S_0 \exp(\mu t + \sigma W_t - \frac{\sigma^2}{2} t) \). One can show that the price process \( S \) is a unique solution to the stochastic differential equation
\[
dS_t = S_t (\mu dt + \sigma dW_t), \quad S_0 = S_0. \tag{5.12}
\]

Let the value \( S^0 \) of the bank account at time \( t \) be \( S^0_t = e^{rt} \).

The market model \((S^0, S, T)\), where \( S^0_t = e^{rt} \) and the stock price \( S \) is given by
\[
S_t = e^{\sigma W_t + \mu t - \frac{\sigma^2}{2} t/2}
\]
is called *Black & Scholes market model*. This is the famous continuous time market model whose analysis is based on the theory of stochastic integration. This is a topic in the course "Finanzmathematik II".

### 5.4.1 Tightness

For a stochastic process \( y \), we define **quadratic variation** and the **\( Q \)-predictable quadratic variation**, respectively as
\[
[y, y]_t := \sum_{\nu=0}^{t} (y_{\nu} - y_{\nu-1})^2, \quad t = 0, 1, \ldots, T
\]
\[
\langle y, y \rangle_t := \sum_{\nu=0}^{t} \mathbb{E}^Q_{t-1} (y_{\nu} - y_{\nu-1})^2, \quad t = 0, 1, \ldots, T,
\]
where \( y_{-1} := 0 \) and \( \mathcal{F}_{-1} := \mathcal{F}_0 \).

The following theorem is a special case (the general case holds, e.g., for every \( p \geq 1 \)) of Burkholder-Davis-Gundy inequalities.

**Theorem 5.23** (Burkholder-Davis-Gundy). *Let \( p \geq 2 \). There is a constant \( L \) such that, for any square integrable martingale \( m \),
\[
E[m^*_T]^p \leq L E[m, m^{p/2}]^p.
\]

The following is the key lemma in our analysis. It does not only give tightness for the weak convergence of our price processes but also uniform integrability in an appropriate sense so that we get convergence results also for unbounded derivatives, like the classical European bid and call options.

We say that a stochastic process \( m \) has **finite moments** if \( E(m^*_T)^p < \infty \) for all \( p \geq 1 \).

**Lemma 5.24.** *For any \( n = 1, 2, \ldots \) let \( \mathbb{Q}_n \) be a probability measures on \( \Omega \) and let \( m^{(n)} = \{m^{(n)}_k\}_{k=0, \ldots, 2^n} \) be a positive \( \mathbb{Q}^n \)-martingale with \( m^{(n)}_0 = 1 \) and with finite moments. Assume that there exists a constant \( L \) such that, for every \( n \) and \( k = 0, \ldots, 2^n \),
\[
|m^{(n)}_{k+1} - m^{(n)}_k| \leq \frac{L}{\sqrt{2^n}} m^{(n)}_k, \quad k = 0, \ldots, 2^n - 1. \tag{5.13}
\]
Then the sequence \( L \) aw\((W(n) | Q^n), n = 1, 2, \ldots, \) is tight on the space \( C \). Furthermore, for any \( m = 1, 2, \ldots, \)
\[
\sup_{n=1,2,\ldots} E^{Q^n} \left[ \max_{k=0,\ldots,2^n} (m_k^{(n)})^{2m} \right] < \infty, \tag{5.14}
\]

Proof. We establish first (5.14).
\[ E^{Q^n}(m_k^{(n)})^{2m} = E^{Q^n}(m_{k-1}^{(n)} + \Delta m_k^{(n)})^{2m} \]
\[ = \sum_{j=0}^{2m} E^{Q^n} \left( \binom{2m}{j} (m_{k-1}^{(n)})^{2m-j} (\Delta m_k^{(n)})^j \right) \]
\[ = \sum_{j=0}^{2m} E^{Q^n} \left( \binom{2m}{j} (m_{k-1}^{(n)})^{2m} E^{Q^n}_{k-1} \left( \frac{\Delta m_k^{(n)}}{m_{k-1}^{(n)}} \right)^j \right) \]
\[ = E^{Q^n} \left( (m_{k-1}^{(n)})^{2m} + \sum_{j=2}^{2m} \binom{2m}{j} (m_{k-1}^{(n)})^{2m-j} E^{Q^n}_{k-1} \left( \frac{\Delta m_k^{(n)}}{m_{k-1}^{(n)}} \right)^j \right) \]
\[ \leq E^{Q^n} \left( (m_{k-1}^{(n)})^{2m} \left( 1 + \sum_{j=2}^{2m} \binom{2m}{j} \left( \frac{L}{\sqrt{2^n}} \right)^j \right) \right) \]
\[ \leq E^{Q^n} \left( (m_{k-1}^{(n)})^{2m} \left( 1 + \sum_{j=2}^{2m} \left( \frac{2m}{j} \left( \frac{L}{\sqrt{2^n}} \right)^j \right) \right) \right) . \]

Repeating the argument for \( k = 2^n, 2^n - 1, \ldots, 0 \), we get
\[ E^{Q^n}(m_{2^n}^{(n)})^{2m} \leq \left( 1 + \sum_{j=2}^{2m} \left( \frac{2m}{j} \left( \frac{L}{\sqrt{2^n}} \right)^j \right) \right) 2^n . \]

Here the right side converges to a finite number as \( n \) tends to infinity, so
\[ \sup_n E^{Q^n}(m_{2^n}^{(n)})^{2m} < \infty. \]
Thus
\[ \sup_n E^{Q^n}((m_{2^n}^{(n)})^{\ast})^{2m} < \infty \]
as well, by Doob’s inequality, Theorem 2.5.

To have tightness, we use Kolmogorov’s tightness criterion. This follows from Burkholder-Davis-Gundy inequality (Theorem 5.23) and (5.14); for \([2^n u], [2^n t] \in \]

46
$D, u \leq t$, we get

\[
E^Q_n \left| W^{(n)}(m^{(n)}_{[2^n u]}) - W^{(n)}(m^{(n)}_{[2^n t]}) \right|^4
\]

\[
= E^Q_n \left| m^{(n)}_{[2^n u]} - m^{(n)}_{[2^n t]} \right|^4
\]

\[
\leq LE^Q_n \left( \sum_{k=[2^n t]}^{[2^n u] - 1} (m_{k+1}^{(n)} - m_k^{(n)})^2 \right)^2
\]

\[
\leq LE^Q_n \left( \sum_{k=[2^n t]}^{[2^n u] - 1} \left( \frac{L m_k^{(n)}}{\sqrt{2^n}} \right)^2 \right)^2
\]

\[
\leq LE^Q_n \left( \sum_{k=[2^n t]}^{[2^n u] - 1} \left( \frac{(L m^{(n)}_{[2^n u]})^2}{2^n} \right) \right)^2
\]

\[
\leq LE^Q_n \left( |t - u| 2^n \left( \frac{(L m^{(n)}_{[2^n u]})^2}{2^n} \right) \right)^2
\]

\[
\leq L |t - u|^2 \sup_n E^Q_n (m^{(n)}_{[2^n u]})^4
\]

\[
\leq L |t - u|^2
\]
5.4.2 Convergence of the superhedging prices

We have already established weak convergence of the price processes. The following lemma will allow us to cover the case of polynomially growing continuous payoffs $F$ rather than merely bounded ones.

**Theorem 5.25.** Assume that $F$ is polynomially growing continuous function on $C$, i.e., there exists $p \geq 1$ such that

$$|F(y)| \leq L(1 + \|y\|^p) \quad \forall \ y \in C$$

We have that the superhedging prices of the discrete time models converge in the following sense:

$$\lim_{n \to \infty} E^Q F(W(s^{(n)})) = E^Q F(S),$$

where $S_t = e^{\sigma W_t - \frac{1}{2} \sigma^2 t + rt}$ is the price process of the Black & Scholes market model.

**Proof.** By Lemma 5.24, the sequence $(\text{Law}(W(s^{(n)}), Q^{(n)}))_{n=1}^{\infty}$ is a tight sequence of laws on $C$. By Theorem 5.20, $W(s^{(n)}) \stackrel{d}{\to} S$, so $W(s^{(n)}) \stackrel{d}{\to} S$, by Theorem 5.13. The Skorokhod’s representation theorem, Theorem 5.7, gives a probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{Q})$ and $\hat{S}^{(n)}$, $\hat{S}$ on $\hat{\Omega}$ such that $\hat{S}^{(n)} \to S$ almost surely and such that $\text{Law}(\hat{S}^{(n)}; Q) = \text{Law}(W(s^{(n)}); Q^{(n)})$ and $\text{Law}(\hat{S}; \hat{\mathcal{F}}) = \text{Law}(S, Q)$.

Consider the function $G(x) = |x|^2$. Since $|F(y)| \leq L(1 + \|y\|^p)$ on $C$, we have that

$$G(|F(y)|) \leq L^2 + 2L^2 \|y\|^p + L^2 \|y\|^{2p},$$

so, by (5.14),

$$\sup_n E^Q G(F(\hat{S}^{(n)})) = \sup_n E^Q G(F(W(s^{(n)})))$$

$$\leq \sup_n L^2 E^Q [1 + 2 \max_{k=1, \ldots, 2^n} |s_k^{(n)}|^p + \max_{k=1, \ldots, 2^n} |s_k^{(n)}|^{2p}]$$

$$< \infty.$$  

By the Dunford-Pettis criterion, $(F(\hat{S}^{(n)}))_{n=1}^{\infty}$ is uniformly integrable on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{Q})$. Since $F$ is continuous and since $\hat{S}^{(n)} \to S$ almost surely, we get

$$\lim_{n} E^Q F(W(s^{(n)})) = \lim_{n} E^Q F(\hat{S}^{(n)}) = E^Q F(S).$$

\[\square\]
5.4.3 Black & Scholes for European put and call options

We know that the approximating price \( \pi_s(n)(c^{(n)}) \) for the European contingent claim \( f(s_T^{(n)}) \) is

\[
\pi_s(n)(c) = (1 + \frac{r}{n})^{-\lfloor nT \rfloor} E^Q f(S_T^{(n)}).
\]

We also know that \( s_T^{(n)} \) converges weakly to the limit \( S_T = e^{\sigma W_T - \frac{1}{2} \sigma^2 T + r T} \), and also that

\[
E^n Q^n f(s_T^{(n)}) \to E^Q f(S_T),
\]
as \( n \to \infty \). In the limit we have the following

\[
\pi_s(f(S_T)) = e^{-r T} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(S_0 e^{(r - \frac{\sigma^2}{2}) T + \sigma \sqrt{T} y}) e^{-\frac{y^2}{2}} dy.
\]

You will learn in the course "Finanzmathematik II" why this is also a hedging price for European options in the Black & Scholes market model.

**Example 5.26** (Black & Scholes- pricing formula for European call). Let

\[
f(S_T) = (S_T - K)^+
\]

be the European call with strike \( K \).

From the equation (5.15) we get

\[
\pi_s(f(S_T)) = e^{-r T} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (S_0 e^{(r - \frac{\sigma^2}{2}) T + \sigma \sqrt{T} y} - K)^+ e^{-\frac{y^2}{2}} dy
\]

\[
= e^{-r T} \frac{1}{\sqrt{2\pi}} \int_{y_0}^{\infty} (S_0 e^{(r - \frac{\sigma^2}{2}) T + \sigma \sqrt{T} y} - K) e^{-\frac{y^2}{2}} dy
\]

\[
= S_0 \frac{1}{\sqrt{2\pi}} \int_{y_0}^{\infty} e^{-\frac{1}{2}(y - \sigma \sqrt{T})^2} dy - e^{-r T} K (1 - \Phi(y_0))
\]

\[
= S_0 (1 - \Phi(y_0 - \sigma \sqrt{T}) - e^{-r T} K (1 - \Phi(y_0))).
\]

Using some properties of normal distributions like symmetry etc. one can write

\[
\pi_s(f(S_T)) = S_0 \Phi\left(\frac{\log(S_0/K) + (r + \frac{\sigma^2}{2}) T}{\sigma \sqrt{T}}\right) - e^{-r T} K \Phi\left(\frac{\log(S_0/K) + (r - \frac{\sigma^2}{2}) T}{\sigma \sqrt{T}}\right).
\]

For more details we refer to exercises.
5.4.4 Exercises

Exercise 5.4.1. Compute $ES_t$ and $\text{Var}(S_t)$, when $S_t$ is the geometric Brownian motion

$$S_t = S_0 e^{\mu t + \sigma W_t - \frac{1}{2} \sigma^2 t}.$$ 

Let $\xi \sim N(\mu, \sigma^2)$. Show that if $K \in \mathbb{R}$ then

$$E(e^{\xi 1_{\{\xi \leq K\}}}) = \exp(\mu + \frac{1}{2} \sigma^2) \Phi(d),$$

where $d = \frac{K - \mu - \sigma^2}{\sigma}$. [The B&S pricing formula is based on such computations.]

Exercise 5.4.2. Consider a Black & Scholes market model $S_t = S_0 e^{\mu t - \frac{1}{2} \sigma^2 t + \sigma W_t}$, where $W$ is a Brownian motion with respect to the measure $P$.

1. What is the risk neutral price of an European option $f_T := S_T 1_{\{K \leq S_T \leq L\}}$, when $r = 0$ and $L > K > 0$.

2. Compute the Black – Scholes hedging price for a call option with the parameters $\mu = 1$, $\sigma = 0.4$, $T = 3$, $r = 0.05$, $K = 100$ and $S_0 = 90$.

Exercise 5.4.3. Put

$$u(t, x) = \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f \left( x e^{y \sigma \sqrt{T-t} + r(T-t) - \frac{\sigma^2 (T-t)}{2} - \frac{\log(x)}{\sigma \sqrt{T-t}}} \right) e^{-\frac{y^2}{2}} dy.$$ 

Show that one can write $u(t, x)$ as

$$u(t, x) = \int_{0}^{\infty} e^{-r(T-t)} \frac{\phi \left( \frac{\log(y - r(T-t) + \sigma^2 (T-t)/2 - \log(x)}{\sigma \sqrt{T-t}} \right)}{y \sigma \sqrt{T-t}} f(y) dy,$$

where $\phi(y) = \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}$. 

50
5.5 Convergence of the asset liability management problem

Recall from Theorem 4.6 that, for the binomial model

\[
s^{(n)}_t = \prod_{t'=1}^t \frac{1 + \mu/2^n + (\sigma/\sqrt{2^n})\xi_t}{1 + r/2^n} \quad \forall \ t = 1, \ldots, 2^n \tag{5.16}
\]

where \( r > -1 \) is the interest rate and \((\xi_t)_{t=1}^T\) are independent identically distributed with \( E\xi_t = 0 \) and \( \xi_t^2 = 1 \), there is no duality gap between the asset liability problem

\[
\text{minimize} \quad EV \left( c^{(n)} - \sum_{t=1}^T z_{t-1} \Delta s^{(n)}_t \right) \quad \text{over} \quad z \in \mathcal{N}^{(n)}_0 \quad (\text{ALM}^{(n)})
\]

with a finite loss function and the dual problem

\[
\text{maximize} \quad E\left[ qc - V^*(q) \right] \quad \text{over} \quad q \in \mathcal{Q}^{(n)}.
\]

Moreover, the dual problem has a solution, and since there is only one martingale measure (with the density process \( q^{(n)} \)) for each \( s^{(n)} \), this can be written as

\[
\text{maximize} \quad E[\alpha q^{(n)} c - V^*(\alpha q^{(n)})] \quad \text{over} \quad \alpha \in \mathbb{R}_+ \quad (\text{D}^{(n)})
\]

By (5.10), the density processes satisfy

\[
q^{(n)}_t = s_0 \prod_{t'=1}^t \left( 1 + \frac{1}{\sqrt{2^n}} \frac{r - \mu}{\sigma} \xi_{t'} \right) \quad \forall \ t = 1, \ldots, 2^n. \tag{5.17}
\]

Thus the density processes are of a similar form as the price processes (5.11), and we may conclude that

\[
W^{(n)}(q^{(n)}) \xrightarrow{d(P)} Z,
\]

where \( Z_t = e^{r - \mu B_t - \frac{1}{2}(r - \mu)^2 t} \) for a Brownian motion \( B \). Recall also that

\[
W^{(n)}(s^{(n)}) \xrightarrow{d(Q^{(n)})} S
\]

where \( S_t = e^{\sigma W_t - \frac{1}{2} \sigma^2 t + rt} \) for (another) Brownian motion \( W \). When \( B \) and \( W \) are the same brownian motion, it is an exercise (using Exercise 5.4.1 and Lemma 2.3), that the measure \( Q \) given by \( dQ/dP := Z_T \) is a martingale measure for \( S \). Thus we see that the dual expression in the following theorem is of the same form as those of the corresponding discrete time problems.

Theorem 5.27. Assume that the claim \( c \) is a derivative of the form \( c = \mathcal{F}(W^{(n)}(s^{(n)})) \) for polynomially growing \( \mathcal{F} \) as in Theorem 5.25 and that \( V^* \) is polynomially growing on \( \mathbb{R}_+ \), i.e., there exists \( p \geq 1 \) such that

\[
|V^*(y)| \leq L(1 + \|y\|^p) \quad \forall \ y \in \mathbb{R}_+.
\]
If $V$ is bounded from below, the optimal values of $(\text{ALM}^{(n)})$ converge to

$$\alpha E^Q F(S) - EV^*(\alpha Z_T),$$

where $\alpha = \lim \alpha^{(n)}$ and each $\alpha^{(n)}$ is optimal for $(D^{(n)})$.

Proof. Denoting $g^{(n)}(\alpha) := -\alpha E^Q F(W^{(n)}(s^{(n)})) + EV^*(\alpha q^{(n)}_T)$, we have that

$$\limsup_n \sup_\alpha g^{(n)}(\alpha) \geq \sup_\alpha \lim_n g^{(n)}(\alpha).$$

Fix $\alpha \in \mathbb{R}_+$. By the discussion before the theorem, we know that $W^{(n)}(s^{(n)}) \Rightarrow S$ and that $W^{(n)}(q^{(n)}) \Rightarrow Z$. We also know from Theorem 5.25 that

$$\lim_n E^Q F(W^{(n)}(s^{(n)})) = E^Q F(S).$$

Since $q^{(n)}$ are of the form (5.17), we have that

$$|q^{(n)}_{k+1} - q^{(n)}_k| \leq \frac{\mu - r}{\sqrt{2^n}} q^{(n)}_k \quad \forall k = 0, \ldots, 2^2 - 1.$$

By Lemma 5.24, we thus have, for any $l = 1, 2, \ldots$

$$\max_n E^P(q^{(n)}_T) < \infty.$$

As in the proof of Theorem 5.25, using the Skorokhod representation theorem, there is a probability space $(\hat{\Omega}, \hat{F}, \hat{P})$ with $\hat{q}^{(n)}$ and $\hat{Z}$ that have the same laws as $q^{(n)}$ and $\hat{Z}$ such that $q^{(n)}_T \Rightarrow \hat{Z}_T$ almost surely. Using the polynomial growth of $V^*$ like in the proof of Theorem 5.25 for $F$, we get that $(V^*(\alpha q^{(n)}_T)_{n=1}^\infty$ is uniformly integrable. Thus

$$\lim_n E^P V^*(\alpha q^{(n)}_T) = E^P V^*(\alpha \hat{Z}_T) = E(\alpha \hat{Z}_T).$$

We conclude that

$$\lim_n g^{(n)}(\alpha) = g(\alpha) := -\alpha E^Q F(S) + EV^*(\alpha Z_T) + \delta_{\mathbb{R}_+}(\alpha).$$

The function $g$ is an extended real-valued convex function on the real line, and its recession function satisfies, by the monotone convergence theorem,

$$g^\infty(\alpha) := \lim_{\lambda \to \infty} \frac{g(\lambda \alpha) - g(0)}{\lambda}$$

$$= \alpha E^Q F(S) E(V^*)^\infty(\alpha Z_T) + \delta_{\mathbb{R}_+}(\alpha)$$

$$= \delta_{\{0\}}(\alpha).$$

From this, it is possible to show that the level sets of $g$ are bounded. Assume now that $\alpha^{(n)}$ are optimizers for $(D^{(n)})$. Since there is no duality gap for each $n$,
\(-g(n)(\alpha(n))\) equals the optimal value of \((\text{ALM}(n))\). By the assumption that \(V\) is bounded from below, these optimal values are bounded from below, so there is \(L \in \mathbb{R}\) such that \(L \geq g(n)(\alpha(n)) \geq g(\alpha(n))\). Since the level sets of \(g\) were shown to be bounded, there is a subsequence (still denoted by \(\alpha(n)\)) that converge to some \(\alpha\).

By Slutsky's lemma, \(\alpha(n)W(n)(q(n)) \to \alpha Z\), and we may repeat the previous part of the proof (still having the appropriate uniform integrability, since \(\alpha(n)\) form a bounded sequence) to get that

\[
\lim EV^*(\alpha(n)q_2(n)) = EV^*(\alpha Z_T).
\]

Of course, we also have \(\lim \alpha(n)E_Q^{(n)} F(W(n)(s(n))) = \alpha E_Q^F(S)\), which proves the claim.

**Remark 5.28.** Just like the limits of the superhedging prices in binomial prices can be shown to be superhedging prices in the Black & Scholes model, the dual expression in Theorem 5.27 corresponds to a dual problem of the asset liability problem in the Black & Scholes model. This analysis is possible using the duality theory introduced in this course and the theory of stochastic integration taught in "Finanzmathematik II".