II Stochastic Financial Models

Ishan Nath, Michaelmas 2023

Based on Lectures by Dr. Michael Tehranchi

August 31, 2024

Contents

1   Introduction

Probability and Measure is desirable; not required. We will talk about about σ-algebras later. Moreover sometimes Fubini's or the dominated convergence theorem is used, without proof.

In the following models, we will use the following assumptions. Note that these are clearly false in the real world:

- No dividends.

- Shares are infinitely divisible. This is a reasonable assumption for a large investor.

- No bid-ask spread, which is reasonable for liquid markets.

- No price impact of buying/selling, so our investing size is not too large.

- No transaction fees.

- No short-selling constraints.

In the following, we use the conventions:

- We will have d stocks (risky assets).

- S_t^i is the price of asset i at time t. For this course, $t \in \{0, 1\}$ or $\{0, 1, 2, \ldots\}$ or $[0, \infty]$. The first few lectures will be focused on the one-period model.

- $S_t = (S_t^1, \ldots, S_t^d)^T \in \mathbb{R}^d$ is the vector of asset prices.

- We will also assume there is a constant risk-free interest rate r.

1.1   One-period Setup

For one-period, we will assume that S_0 is not random, but S_1 is a random vector. Moreover will assume that there is no uncertainty in the distribution.

Our investor holds θ^i shares of asset i. Denote by $\theta = (\theta^1, \ldots, \theta^d)^T \in \mathbb{R}^d$ the vector of the total position of the investor.

Note that $\theta^i < 0$ means that the investor is shorting $|\theta^i|$ shares.

The investor will also put θ^0 in the bank, gaining the risk-free interest rate on this investment.

Let X_t denote the wealth, or liquidation value, of the portfolio at time t. At time

0, this is simply

$$X_0 = \theta^0 + \sum_{i=1}^{d} \theta^i S_0^i = \theta^0 + \theta \cdot S_0.$$

As time progresses, θ^0 accrues interest, so

$$X_1 = \theta^0(1+r) + \theta \cdot S_1.$$

Hence

$$X_1 = (1+r)X_0 + \theta \cdot (S_1 - (1+r)S_0). \tag{1.1}$$

In general,

$$X_t = (1+r)X_{t-1} + \theta_t \cdot (S_t - (1+r)S_{t-1}). \tag{1.2}$$

2   The Mean-Variance Portfolio Problem

This is a problem first proposed by Markowitz in 1952. There are several variations; sometimes we want to find solutions where $\theta^i \geq 0$ for all i, i.e. we are not allowed to short.

Given X_0, our goal is to find the portfolio $\theta \in \mathbb{R}^d$ to minimise $\text{Var}(X_1)$, subject to $\mathbb{E}[X_1] \geq m$.

We will say that $\mathbb{E}[S_1] = \mu \in \mathbb{R}^d$, and $\text{Cov}(S_1) = \mathbb{E}[(S_1 - \mu)(S_1 - \mu)^T] = V$. This covariance matrix is symmetric and non-negative definite.

For this problem assume that $\mu \neq (1+r)S_0$, and that V is positive definite, hence V^{-1} exists. Then, from (1.1) we get

$$\mathbb{E}[X_1] = (1+r)X_0 + \theta \cdot (\mu - (1+r)S_0),$$

$$\text{Var}(X_1) = \theta^T V \theta.$$

Our problem is hence to minimize $\theta^T V \theta$ subject to $\theta^T(\mu - (1+r)S_0) \geq m - (1+r)X_0$.

We will not prove this, but note the following:

Theorem 2.1. *The optimal portfolio is*

$$\theta = \lambda V^{-1}(\mu - (1+r)S_0),$$

where

$$\lambda = \frac{(m - (1+r)X_0))^+}{(\mu - (1+r)S_0)^T V^{-1}(\mu - (1+r)S_0)}.$$

Here we use the convention $a^+ = \max\{a, 0\}$.

Lemma 2.1. *The minima of $\theta^T V \theta$ subject to $\theta^T a = b$ is*

$$\frac{b^2}{a^T V^{-1} a},$$

where V is positive definite, $a \in \mathbb{R}^d$, $a \neq 0$ and $b \in \mathbb{R}$ is given. The unique minimiser is $\theta = \lambda V^{-1} a$, where

$$\lambda = \frac{b}{a^T V^{-1} a}.$$

Proof: The first method will be taking the Lagrangian

$$L(\theta, \lambda) = \frac{1}{2}\theta^T V \theta + \lambda(b - a^T\theta).$$

We can minimize by setting the derivatives 0, i.e.

$$D_\theta L = V\theta - \lambda a = 0 \implies \theta = \lambda V^{-1}a.$$

Pick λ to be feasible: so

$$\theta^T a = b \implies \lambda a^T V^{-1}a = b \implies \lambda = \frac{b}{a^T V^{-1}a}.$$

This isn't very formal; let's work backwards instead. Define $\lambda = b/(a^T V^{-1}a)$, and suppose $\theta^T a = b$. Then we get

$$\begin{aligned}
\theta^T V\theta &= \theta^T V\theta + 2\lambda(b - a^T\theta) \\
&= (\theta - \lambda V^{-1}a)^T V(\theta - \lambda V^{-1}a) + 2\lambda b - \lambda^2 a^T V^{-1}a \\
&\geq 2\lambda b - \lambda^2 a^T V^{-1}a = \frac{b^2}{a^T V^{-1}a}.
\end{aligned}$$

Since V is positive definite, we have equality if and only if $\theta = \lambda V^{-1}a$.

Let's look at another proof. Note that

$$\begin{aligned}
(\theta^T a)^2 &= ((V^{1/2}\theta)^T (V^{-1/2}a))^2 \\
&\leq (\theta^T V\theta)(a^T V^{-1}a) \\
\implies \theta^T V\theta &\geq \frac{(a^T\theta)^2}{a^T V^{-1}a} = \frac{b^2}{a^T V^{-1}a},
\end{aligned}$$

when $a^T\theta = b$. We have equality in Cauchy-Schwarz if and only if $V^{1/2}\theta = \lambda V^{-1/2}a$ for some $\lambda \in \mathbb{R}$, or $\theta = \lambda V^{-1}a$.

As an application, if $a = \mu - (1 + r)S_0$, and $b = \mathbb{E}[X_1] - (1 + r)X_0$, then

$$\text{Var}(X_1) \geq \frac{(\mathbb{E}[X_1] - (1 + r)X_0)^2}{(\mu - (1 + r)S_0)^T V^{-1}(\mu - (1 + r)S_0)},$$

with equality if and only if $\theta = \lambda V^{-1}(\mu - (1 + r)S_0)$, where

$$\lambda = \frac{\mathbb{E}[X_1] - (1 + r)X_0}{(\mu - (1 + r)S_0)^T V^{-1}(\mu - (1 + r)S_0)}.$$

Definition 2.1. The *mean-variance efficient frontier* is the lower boundary of the set of possible pairs $(\mathbb{E}[X_1], \mathrm{Var}(X_1))$, i.e. the points on the parabola.

A portfolio is *mean-variance efficient* if it is the optimal solution of some mean-variance portfolio problem.

Recall that we wanted to minimize $\mathrm{Var}(X_1)$ subject to $\mathbb{E}[X_1] \geq m$. However we can figure out what do to by looking at the above. If $m \leq (1+r)X_0$ then we choose $\theta = 0$ (i.e. we don't invest in risky assets). If $m \geq (1+r)X_0$, we choose $\theta = \lambda V^{-1}(\mu - (1+r)S_0)$, where λ is as before.

Theorem 2.2 (Mutual fund theorem). *θ is mean-variance efficient if and only if $\theta = \lambda V^{-1}(\mu - (1+r)S_0)$ for some $\lambda \geq 0$.*

> **Proof:** If it is mean-variance efficient, we know it must have this form. Conversely, given $\lambda \geq 0$, let $m = (1+r)X_0 + \lambda(\mu-(1+r)S_0)^T V^{-1}(\mu-(1+r)S_0)$. Then θ is the optimal solution to the mean-variance problem with target m.

2.1   Capital Asset Pricing Model

This subsection will be an application of the mutual fund theorem.

Theorem 2.3 (Linear regression coefficients). *Let X and Y be square-integrable, with $\mathrm{Var}(X) > 0$. There are unique constants a and b such that*

$$Y = a + bX + Z,$$

where $\mathbb{E}[Z] = 0$, and $\mathrm{Cov}(X, Z) = 0$.

> **Proof:** Suppose this formula holds, so $Z = Y - a - bX$. Then
>
> $$\mathbb{E}[Z] = \mathbb{E}[Y] - a - b\mathbb{E}[X],$$
>
> $$\mathrm{Cov}(X, Z) = \mathrm{Cov}(X, Y) - b\,\mathrm{Var}(X).$$
>
> These are two unknowns, so their unique solution to $\mathbb{E}[Z] = 0$ and $\mathrm{Cov}(X, Z) = 0$ is given by
> $$b = \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(X)}, \qquad a = \mathbb{E}[Y] - b\mathbb{E}[X].$$

Definition 2.2. $\theta_{\mathrm{mar}} = V^{-1}(\mu - (1+r)S_0)$ is called the *market portfolio*.

Definition 2.3. Fix $X_0 > 0$ and $\theta \in \mathbb{R}^d$. The *excess return of the portfolio* R^{ex} is

$$\frac{X_1}{X_0} - (1 + r) = \frac{\theta^T}{X_0}(S_1 - (1 + r)S_0).$$

Then the excess return of the market is

$$R_{\text{mar}}^{\text{ex}} = \frac{\theta_{\text{mar}}^T}{X_0}(S_1 - (1 + r)S_0).$$

Theorem 2.4 (α is zero). *Fix $X_0 > 0$ and $\theta \in \mathbb{R}^d$. Let α, β be such that*

$$R^{\text{ex}} = \alpha + \beta R_{\text{mar}}^{\text{ex}} + \varepsilon,$$

where $\mathbb{E}[\varepsilon] = 0$ and $\text{Cov}(\varepsilon, R_{\text{mar}}^{\text{ex}}) = 0$. Then $\alpha = 0$.

Proof:　This is just a calculation. We need to calculate

$$\text{Cov}(R^{\text{ex}}, R_{\text{mar}}^{\text{ex}}) = \text{Cov}\left(\frac{\theta^T}{X_0}(S_1 - (1 + r)S_0), \frac{\theta_{\text{mar}}^T}{X_0}(S_1 - (1 + r)S_0)\right)$$

$$= \frac{\theta^T}{X_0}\frac{V\theta_{\text{mar}}}{X_0} = \frac{\theta^T}{X_0^2}(\mu - (1 + r)S_0)$$

$$= \frac{\mathbb{E}[R^{\text{ex}}]}{X_0}.$$

Hence we get

$$\text{Var}(R_{\text{max}}^{\text{ex}}) = \frac{\mathbb{E}[R_{\text{mar}}^{\text{ex}}]}{X_0} = \text{Cov}(R_{\text{mar}}^{\text{ex}}, R_{\text{mar}}^{\text{ex}}).$$

Therefore

$$\beta = \frac{\text{Cov}(R^{\text{ex}}, R_{\text{mar}}^{\text{ex}})}{\text{Var}(R_{\text{mar}}^{\text{ex}})} = \frac{\mathbb{E}[R^{\text{ex}}]}{\mathbb{E}[R_{\text{mar}}^{\text{ex}}]},$$

and so

$$\alpha = \mathbb{E}[R^{\text{ex}}] - \beta\mathbb{E}[R_{\text{mar}}^{\text{ex}}] = 0.$$

The assumptions in our model are as follows:

- There are K agents, where agent k holds portfolio θ_k.

- There are n_i shares of asset i in total. Let $n = (n_1, \ldots, n_d)^T > 0$.

- All investors are mean-variance efficient, and they agree on the mean and covariance of S_1.

- By mutual fund theorem $\theta_k = \lambda_k \theta_{\mathrm{mar}}$ for $\lambda_k \geq 0$ scalar.

- As supply equals demand,

$$\sum_{k=1}^{K} \theta_k = n = \left(\sum_{k=1}^{K} \lambda_k\right) \theta_{\mathrm{mar}} = \Lambda \theta_{\mathrm{mar}}.$$

The implication is that $R_{\mathrm{mar}}^{\mathrm{ex}}$ is the excess return of a broad market index (such as S&P 500 or FTSE), and that

$$R^{\mathrm{ex}} = \beta R_{\mathrm{mar}}^{\mathrm{ex}} + \varepsilon.$$

This is a testable theory, and we can test it in the market and see that it often fails; we can find (negative) α in places. The reason is that some of the assumptions we have made are very strong: we assume a frictionless market without spreads and market impact, but we also assume that all other investors behave nicely.

3   Expected Utility Hypothesis

So far, implicitly, an investor would prefer X over Y if either:

- $\mathbb{E}[X] > \mathbb{E}[Y]$, $\text{Var}(X) \leq \text{Var}(Y)$, or

- $\mathbb{E}[X] \geq \mathbb{E}[Y]$, $\text{Var}(X) < \text{Var}(Y)$.

This is quite strong. Let's assume a different condition.

Proposition 3.1 (Expected Utility Hypothesis)**.** *An agent prefers X to Y if and only if*

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

for a given function U, called the agent's utility function.

Note that if $\hat{U}(x) = a + bU(x)$ for all x, where $a, b \in \mathbb{R}$ and $b > 0$, then the preferences induced by $\hat{U}$ and U are the same. The von-Neumann-Morgenstern theorem characterizes the preferences equivalent to the expected utility hypothesis.

Definition 3.1. An expected utility agent is *indifferent* to X and Y if and only if $\mathbb{E}[U(X)] = \mathbb{E}[U(Y)]$.

3.1   Risk-aversion and Concavity

For now, we will assume:

- U is strictly increasing,

$$U(x) > U(y) \iff x > y.$$

  The implication is, if $X \geq Y$ a.s. and $\mathbb{P}(X > Y) > 0$, then $\mathbb{E}[U(X)] > \mathbb{E}[U(Y)]$, so X is preferred to Y.

- U is concave,

$$U(px + (1 - p)y) > pU(x) + (1 - p)U(y),$$

  for all $x \neq y$ and $0 < p < 1$.

  The implication is, if X is not constant, then by Jensen's inequality

$$\mathbb{E}[U(X)] < U(\mathbb{E}[X]).$$

  Hence $\mathbb{E}[X]$ is preferred to X.

3.2   Concave functions

First we assume that U is always increasing, strictly concave and twice differentiable.

The derivative U' is called the *marginal utility*. This is positive as U is increasing, but decreasing as U is concave. This means that U'' is negative.

Definition 3.2. The *coefficient of absolute risk aversion* is

$$\frac{-U''(x)}{U'(x)}.$$

The *coefficient of relative risk aversion* is

$$\frac{-xU''(x)}{U'(x)},$$

when $x > 0$.

Example 3.1.

Here are some important utility functions.

1. $U(x) = -e^{-\gamma x}$ is the CARA (constant absolute risk aversion) utility function. Indeed,

$$U'(x) = \gamma e^{-\gamma x}, \qquad U''(x) = -\gamma^2 e^{-\gamma x},$$

$$\implies \frac{U''(x)}{U'(x)} = \gamma$$

   is constant.

2. $U(x) = x^{1-R}/(1-R)$, for $R > 0$, $R \neq 1$ is the CRRA (constant relative risk aversion) utility function. We can calculate that

$$\frac{-xU''(x)}{U'(x)} = R.$$

   Note this function is not defined everywhere. To be fully define the functions, we can let

$$U(x) = \begin{cases} \frac{x^{1-R}}{1-R} & x > 0, \\ -\infty & x \leq 0. \end{cases}$$

3. $U(x) = \log x$ is the CRRA utility function with $R = 1$.

4. $U(x) = x$ is a utility function that is not strictly concave. This is the *risk-neutral utility*.

Let's prove some basic facts about concave functions.

Theorem 3.1. *Concave functions are continuous, and differential concave functions have graphs that lie beneath their tangents.*

To prove this we first prove a lemma:

Lemma 3.1. *Let U be concave and $a < b < c$. Then,*
$$\frac{u(b) - u(a)}{b - a} \geq \frac{u(c) - u(b)}{c - b}.$$

Proof: Note that
$$b = \left(\frac{b - a}{c - a}\right)c + \left(\frac{c - b}{c - a}\right)a,$$
and these coefficients are in $(0, 1)$. Hence applying the definition of concavity,
$$u(b) \geq \left(\frac{b - a}{c - a}\right)u(c) + \left(\frac{c - b}{c - a}\right)u(a).$$
Rearranging gives what we want.

Now we can apply this lemma to prove continuity.

Proof: Fix x, $0 < \varepsilon < t$. Apply the lemma to
$$x - t < x - \varepsilon < x < x + \varepsilon < x + t.$$
Then we get that
$$\frac{\varepsilon}{t}(u(x) - u(x - t)) \geq u(x) - u(x - \varepsilon) \geq u(x + \varepsilon) - u(x) \geq \frac{\varepsilon}{t}(u(x + t) - u(x)),$$
which goes to 0 as $\varepsilon \to 0$.

Moreover if u is differentiable at x, then we have
$$u(x) - u(x - t) \geq tu'(x) \geq u(x + t) - u(x),$$
by taking $\varepsilon \to 0$. This gives
$$u(y) \leq u(x) + u'(x)(y - x).$$
Note if u is twice differentiable, we can use second-order Taylor's:
$$u(y) = u(x) + u'(x)(y - x) + u''(z)\frac{(y - x)^2}{2},$$
where $u''(z) \leq 0$.

Theorem 3.2. *Increasing concave functions are unbounded to the left, i.e.* $u(x) \to -\infty$ *as* $x \to -\infty$.

Proof: Pick $x < a < b$. Then,

$$u(x) \leq u(a) + \left(\frac{x-a}{b-a}\right)(u(b) - u(a)),$$

which goes to $-\infty$ as $x \to -\infty$

3.3   Optimal Investment and Marginal Utility

Theorem 3.3. *Consider the problem of maximizing* $\mathbb{E}[U(X_1)]$ *given* X_0, *where*

$$X_1 = (1+r)X_0 + \theta^T(S_1 - (1+r)S_0),$$

where U *is increasing, concave and suitable. If* θ^* *is optimal, then*

$$S_0 = \frac{1}{1+r}\frac{\mathbb{E}[U'(X_1^*)S_1]}{\mathbb{E}[U'(X_1^*)]},$$

where

$$X_1^* = (1+r)X_0 + (\theta^*)^T(S_1 - (1+r)S_0).$$

Proof: Let

$$f(\theta) = \mathbb{E}[U((1+r)X_0) + \theta^T(S_1 - (1+r)S_0))].$$

At the maximum,

$$Df(\theta^*) = 0 = \mathbb{E}[U'(X_1^*)(S_1 - (1+r)S_0)].$$

This rearranges to what we want.

4   Pricing Contingent Claims

For us, a contingent claim is a random variable (the payout is random). The payout for us is at time 1. We will introduce the following framework:

- There are d assets given (fundamental).

- We introduce a new contingent claim with payout Y.

- Often $Y = g(S_1)$, a function of the price of the stocks at time 1, but not necessarily.

We are interested in finding a good time 0 price of the claim.

The most important example is a *call option*: the right but not the obligation to buy some given asset at a fixed price K, often called the *strike price*.

We will assume that the actors are rational. That is, if at time 1 the asset has price greater than K, then the buyer will exercise the option and sell the asset on the market, making $S_1 - K$. If the asset has price less than K, then the buyer will not exercise the right. This gives a payout

$$
\text{payout} = \begin{cases} S_1 - K & S_1 > K, \\ 0 & S_1 \le K, \end{cases}
$$
$$
= (S_1 - K)^+ = g(S_1).
$$

4.1   Utility Indifference Pricing

Fix X_0 to be an initial wealth, and U an increasing concave utility function. Let

$$
\mathcal{X} = \{ X_1 = (1 + r)X_0 + \theta^T(S_1 - (1 + r)S_0) \mid \theta \in \mathbb{R}^d \}.
$$

Note that the investor would prefer to buy one claim with payout Y and price π if there exists $X_1^* = X^*$ such that

$$
\mathbb{E}[U(X^* + Y - (1 + r)\pi)] > \mathbb{E}[U(X)],
$$

for all $X \in \mathcal{X}$.

We will assume U, the distribution of S_1, X_0, Y, the values r, S_0 and π are such that every utility maximisation problem has an optimal solution.

Definition 4.1. The (utility) *indifference (reservation) price* of the claim with payout Y is any π such that

$$
\max_{X \in \mathcal{X}} \mathbb{E}[U(X + Y - (1 + r)\pi)] = \max_{X \in X} \mathbb{E}[U(X)].
$$

Here are some properties of the indifference price. First, our assumptions lead to the following.

Theorem 4.1. *The indifference price exists and is unique.*

From this, we let $\pi(Y)$ denote the indifference price of the claim with payout Y. Now we will relate indifference prices of different claims.

Theorem 4.2. *If $Y_0 \leq Y_1$ almost surely and $\mathbb{P}(Y_0 < Y_1) > 0$, then $\pi(Y_0) < \pi(Y_1)$.*

Moreover, we can show that indifference prices are concave:

Theorem 4.3 (Indifference prices are concave)**.**

$$\pi(pY_1 + (1-p)Y_0) \geq p\pi(Y_1) + (1-p)\pi(Y_0),$$

for all payouts Y_0, Y_1 and all $0 < p < 1$.

Definition 4.2. The *marginal utility price* of Y is

$$\pi_0(Y) = \frac{\mathbb{E}[U'(X^*)Y]}{(1+r)\mathbb{E}[U'(X^*)]},$$

where X^* solves

$$\mathbb{E}[U(X^*)] = \max_{X \in \mathcal{X}} \mathbb{E}[U(X)].$$

Now we have a nice little theorem.

Theorem 4.4. $\pi(Y) \leq \pi_0(Y)$ *for any claim.*

In fact, we can find the following.

Theorem 4.5.
$$\lim_{\varepsilon \to 0} \frac{\pi(\varepsilon Y)}{\varepsilon} = \pi_0(Y).$$

To prove this, we introduce the following definition:

Definition 4.3. We define

$$V(Z) = \max_{X \in \mathcal{X}} \mathbb{E}[U(X + Z)].$$

This is the *indirect utility* of the random payout Z.

In this notation, π is the indifference price of payout Y if and only if

$$V(Y - (1+r)\pi) = V(0).$$

Lemma 4.1. *If $Z_0 \leq Z_1$ almost surely, and $\mathbb{P}(Z_0 < Z_1) > 0$, then $V(Z_0) < V(Z_1)$.*

Proof: Let X^i be such that

$$V(Z_i) = \mathbb{E}[U(X^i + Z_i)],$$

for $i = 0, 1$. Then

$$\begin{aligned}
V(Z_1) = \mathbb{E}[U(X^1 + Z_1)] &\geq \mathbb{E}[U(X^0 + Z_1)] \text{ as } X^1 \text{ is a maximizer} \\
&> \mathbb{E}[U(X^0 + Z_0)] \text{ as } Z_1 \geq Z_0 \\
&= V(Z_0).
\end{aligned}$$

Lemma 4.2. *The indirect utility satisfies*

$$V(pZ_1 + (1-p)Z_0) \geq pV(Z_1) + (1-p)Z_0,$$

for all $0 < p < 1$.

Proof: Let X^i be as before. Note that

$$\begin{aligned}
pX^1 + (1-p)X^0 &= p((1+r)X_0 + \theta_1^T(S_1 - (1+r)S_0)) \\
&\quad + (1-p)((1+r)X_0 + \theta_0^T(S_1 - (1+r)S_0)) \\
&= (1+r)X_0 + (p\theta_1 + (1-p)\theta_0)^T(S_1 - (1+r)S_0) \in \mathcal{X},
\end{aligned}$$

so we have

$$\begin{aligned}
pV(Z_1) + (1-p)V(Z_0) &= \mathbb{E}[pU(X^1 + Z_1) + (1-p)U(X^0 + Z_0)] \\
&\leq \mathbb{E}[U(pX^1 + (1-p)X^0 + pZ_1 + (1-p)Z_0)] \\
&\leq \max_{X \in \mathcal{X}} \mathbb{E}[U(X + pZ_1 + (1-p)Z_0)] \\
&= V(pZ_1 + (1-p)Z_0).
\end{aligned}$$

Now we can prove the existence and uniqueness of indifference prices.

Proof: We show there is a unique solution to $V(Y - (1+r)\pi) = V(0)$. Note that

$$V(0) = \mathbb{E}[U(X^*)] \in (U(-\infty), U(\infty)) = (-\infty, U(\infty)).$$

Now consider $x \mapsto V(Y + x)$. This is strictly increasing and concave, by our lemma. Hence it is continuous (by concavity), with range $(\infty, \lim V(Y + x))$.

But by MCT (actually not entirely sure, Fatou and MCT require non-

negativity),
$$V(Y + x) \geq \mathbb{E}[U(X^* + Y + x)] \uparrow U(\infty).$$

Hence there exists a unique x such that $V(Y + x) = V(0)$, and the unique π is $-x/(1 + r)$.

Now we can prove that indifference prices increase.

Proof: Let $Y_0 \leq Y_1$ almost surely, and $\mathbb{P}(Y_0 < Y_1) > 0$. Then

$$V(Y_0 - (1 + r)\pi(Y_0)) = V(0) = V(Y_1 - (1 + r)\pi(Y_1))$$
$$\geq V(Y_0 - (1 + r)\pi(Y_1)),$$

and so $-(1 + r)\pi(Y_0) > -(1 + r)\pi(Y_1)$ since $x \mapsto V(Y_0 + x)$ is strictly increasing. Hence $\pi(Y_1) > \pi(Y_0)$.

Here is a proof of concavity.

Proof: We are given Y_0, Y_1 and $0 < p < 1$. Then

$$V(pY_1 + (1 - p)Y_0 - (1 + r)\pi(pY_1 + (1 - p)Y_0)) = V(0)$$
$$= V(Y_1 - (1 + r)\pi(Y_1)) = V(Y_0 - (1 + r)\pi(Y_0))$$
$$= pV(Y_1 - (1 + r)\pi(Y_1)) + (1 - p)V(Y_0 - (1 + r)\pi(Y_0))$$
$$\leq V(pY_1 + (1 - p)Y_0 - (1 + r)(p\pi(Y_1) + (1 - p)\pi(Y_0))),$$

hence $\pi(pY_1 + (1 - p)Y_0) \geq p\pi(Y_1) + (1 - p)\pi(Y_0)$.

Now we prove the marginal utility price is bigger than the indifference price.

Proof: Let X^* be such that $V(0) = \mathbb{E}[U(X^*)]$. Let X^1 be such that $V(Y - (1 + r)\pi(Y)) = \mathbb{E}[U(X^1 + Y - (1 + r)\pi(Y_0))]$. Then,

$$V(0) = V(Y - (1 + r)\pi(Y)) = \mathbb{E}[U(X^1 + Y - (1 + r)\pi(Y))]$$
$$\leq \mathbb{E}[U(X^*) + U'(X^*)(X^1 - X^* + Y - (1 + r)\pi(Y))]$$
$$= V(0) + \mathbb{E}[U'(X^*)(X^1 - X^*)] + \mathbb{E}[U'(X^*)(Y - (1 + r)\pi(Y))].$$

Now we claim the bit in the middle is 0. Assuming this, we have

$$\pi(Y) \leq \frac{1}{1 + r} \frac{\mathbb{E}[U'(X^*)Y]}{\mathbb{E}[U'(X^*)]} = \pi_0(Y).$$

To prove this claim note that

$$\mathbb{E}[U'(X^*)(X^* - X^1)] = (\theta^* - \theta^1)^T\mathbb{E}[U'(X^*)(S_1 - (1+r)S_0)] = 0,$$

by the first marginal utility pricing result.

We now prove the convergence of the indifference price to the marginal utility price.

Proof: Fix Y, and let $\pi_t = \pi(tY)/t$, the average indifference price of t shares. Let

$$p = \sup_{t>0} \pi_t = \lim_{t\to 0} \pi_t,$$

from the example sheet, where we show $t \mapsto \pi_t$ is decreasing. This is done by using concavity of π, and the fact $\pi(0) = 0$. Let $X^* \in \mathcal{X}$ maximize $\mathbb{E}[U(X)]$, as before. Then

$$0 = \frac{V(tY - (1+r)t\pi_t) - V(0)}{t}$$

$$\geq \frac{1}{t}\mathbb{E}[U(X^* + t(Y - (1+r)\pi_t)) - U(X^*)]$$

$$\geq \mathbb{E}\left[\frac{U(X^* + t(Y - (1+r)p)) - U(X^*)}{t}\right]$$

$$\to \mathbb{E}[U'(X^*)(Y - (1+r)p)],$$

hence we get

$$p \geq \pi_0 = \frac{\mathbb{E}[U'(X^*)Y]}{(1+r)\mathbb{E}[U'(X^*)]}.$$

But on the other hand,

$$\pi_t \leq \pi_0 \implies p \leq \pi_0,$$

hence $p = \pi_0$.

5   Risk Neutral Measures

Definition 5.1. Consider two probability measures $\mathbb{P}$ and $\mathbb{Q}$, defined on a σ-algebra $\mathcal{F}$.

Then $\mathbb{P}$ and $\mathbb{Q}$ are equivalent if and only if $\mathbb{P}(A) = 0 \iff \mathbb{Q}(A) = 0$.

Remark. Analogously, $\mathbb{P} \sim \mathbb{Q}$ if and only if $\mathbb{P}(A) = 1 \iff \mathbb{Q}(A) = 1$, or $\mathbb{P}(A) > 0 \iff \mathbb{Q}(A) > 0$ or $\mathbb{P}(A) < 1 \iff \mathbb{Q}(A) < 1$.

Example 5.1.

Take $\Omega = \{1, 2, 3\}$ and $\mathcal{F} = \mathcal{P}(\Omega)$. Let

$$\mathbb{P}(1) = \frac{1}{2}, \qquad \mathbb{P}(2) = \frac{1}{2}, \qquad \mathbb{P}(3) = 0,$$

$$\mathbb{Q}(1) = \frac{1}{3}, \qquad \mathbb{Q}(2) = \frac{2}{3}, \qquad \mathbb{Q}(3) = 0.$$

Then $\mathbb{P} \sim \mathbb{Q}$.

Moreover, let Z be a random variable such that $\mathbb{P}(Z > 0) = 1$, and such that $\mathbb{E}[Z] = 1$. Define $\mathbb{Q}(A) = \mathbb{E}[Z \mathbb{1}_A]$.

Then $\mathbb{Q}$ is a probability measure on $(\Omega, \mathcal{F})$, and $\mathbb{P} \sim \mathbb{Q}$.

Indeed, if $\mathbb{P}(A) = 0$, then $\mathbb{E}[\mathbb{1}_A] = 0$, and so $\mathbb{Q}(A) = \mathbb{E}[Z \mathbb{1}_A] = 0$ by MCT.

Conversely, if $\mathbb{Q}(A) = 0$, then $\mathbb{E}[Z \mathbb{1}_A] = 0$, so $\mathbb{P}(Z \mathbb{1}_A = 0) = 1$. But since $\mathbb{P}(Z = 0) = 0$, we must have $\mathbb{P}(\mathbb{1}_A = 0) = 0$, hence $\mathbb{P}(A) = 0$.

In fact, this is the only such example.

Theorem 5.1 (Radon-Nikodym). *$\mathbb{P} \sim \mathbb{Q}$ if and only if there exists random variable Z with respect to $\mathbb{P}$ such that $\mathbb{P}(Z > 0) = 1$, $\mathbb{E}[Z] = 1$ and*

$$\mathbb{Q}(A) = \mathbb{E}[Z \mathbb{1}_A],$$

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

The proof of this is in part III advanced probability.

For such $\mathbb{P}, \mathbb{Q}$, we define

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

This is the *Radon-Nikodym derivative*.

In the first example, $Z(1) = 2/3$, $Z(2) = 4/3$ and $Z(3) = e^2 - \pi$ gives $\mathbb{Q}(A) = \mathbb{P}(Z\mathbb{1}_A)$.

Remark. If $\mathbb{P} \sim \mathbb{Q}$ and $\frac{d\mathbb{Q}}{d\mathbb{P}} = Z$, then $\mathbb{E}_\mathbb{Q}[X] = \mathbb{E}_\mathbb{P}[ZX]$, for any $\mathbb{Q}$-integrable X.

Example 5.2.

Let X be defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and $X \sim \text{Exp}(\lambda)$ under $\mathbb{P}$. Then let

$$Z = \frac{u}{\lambda}e^{(\lambda-u)x}, \qquad \mathbb{E}[Z] = \int_0^\infty \frac{u}{\lambda}e^{(\lambda-u)x}\lambda e^{-\lambda x}\, dx = \int_0^\infty ue^{-ux}\, dx = 1.$$

So if we define $\mathbb{Q}$ by density Z, then we want to find the distribution of X under $\mathbb{Q}$. Note that

$$\mathbb{E}_\mathbb{Q}[f(X)] = \mathbb{E}_\mathbb{P}[Zf(X)] = \int_0^\infty \frac{u}{\lambda}e^{(\lambda-u)x}f(x)\lambda e^{-\lambda x}\, dx$$

$$= \int_0^\infty ue^{-ux}f(x)\, dx,$$

so $X \sim \text{Exp}(u)$ under $\mathbb{Q}$.

Definition 5.2. In a financial model, any measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that

$$S_0 = \frac{\mathbb{E}^\mathbb{Q}[S_1]}{1+r},$$

is called *risk-neutral*.

Remark.

- The definition of a risk-neutral measure does not depend on the agent's utility.

- It does depend on $\mathbb{P}$.

- Agents with equivalent beliefs (i.e. agree on the probability 0 events) agree on the risk-neutrality of a measure.

Theorem 5.2 (Marginal Utility Pricing Z). *Let $X^* \in \mathcal{X}$ maximize $\mathbb{E}[U(X)]$ over $X \in \mathcal{X}$, where U is strictly increasing, differentiable and suitably nice. Let $\mathbb{Q}$ be defined by*

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{U'(X^*)}{\mathbb{E}[U'(X^*)]}.$$

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

We can also say that

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \propto U'(X^*).$$

Proof: Note that

$$\frac{\mathbb{E}_{\mathbb{Q}}[S_1]}{1+r} = \frac{\mathbb{E}_{\mathbb{P}}[U'(X^*)S_1]}{(1+r)\mathbb{E}_{\mathbb{P}}[U'(X^*)]} = S_0,$$

by the first marginal utility pricing theorem.

5.1   Arbitrage

We define the concept of arbitrage, which will be used in the fundamental theorem of asset pricing: there exists a risk-neutral measure if and only if there does not exist an arbitrage.

Definition 5.3. An *arbitrage* $\phi \in \mathbb{R}^d$ is such that $\mathbb{P}(\phi^T(S_1 - (1+r)S_0) \geq 0) = 1$, and $\mathbb{P}(\phi^T(S_1 - (1+r)S_0) > 0) > 0$.

Remark.

- The definition of arbitrage does not depend on U.

- It depends on $\mathbb{P}$, but only through the events of measure 0.

Given $X \in \mathcal{X}$, and an arbitrage ϕ, let

$$X' = X + \phi^T(S_1 - (1+r)S_0) \in \mathcal{X}.$$

Then $U(X') \geq U(X)$ almost surely, and $\mathbb{P}(U(X') > U(X)) > 0$. Hence $\mathbb{E}[U(X')] > \mathbb{E}[U(X)]$, so X' is preferred to X.

Hence there cannot be a maximiser of $\mathbb{E}[U(X)]$. Indeed, if ϕ is an arbitrage, then so is $n\phi$ for all $n \geq 0$.

Hence, if there exists a utility maximiser, then there exists a risk-neutral measure, and if there exists a utility maximiser, then there is no arbitrage.

Theorem 5.3 (Fundamental Theorem of Asset Pricing)**.** *There exists a risk-neutral measure if and only if there does not exist an arbitrage.*

Proof: Suppose there exists a risk-neutral measure. To show there is no arbitrage, let $\phi \in \mathbb{R}^d$ be such that $\phi^T(S_1 - (1+r)S_0) \geq 0$ $\mathbb{P}$-almost surely.

Then by equivalence, $\phi^T(S_1 - (1+r)S_0) \geq 0$ $\mathbb{Q}$-almost surely. But

$$\mathbb{E}_{\mathbb{Q}}[\phi^T(S_1 - (1+r)S_0)] = \phi^T \mathbb{E}_{\mathbb{Q}}[S_1 - (1+r)S_0] = 0,$$

by risk-neutrality. Therefore $\phi^T(S_1 - (1+r)S_0) = 0$ $\mathbb{Q}$-almost surely, and hence $\mathbb{P}$-almost surely. So ϕ is not an arbitrage.

Now we show the backwards direction. Assume there is no arbitrage; then we will find a solution to the utility maximisation problem.

Without loss of generality, assume $\mathbb{E}_{\mathbb{P}}[e^{\theta \cdot S_1}] < \infty$ for all $\theta \in \mathbb{R}^d$. Otherwise, replace $\mathbb{P}$ with $\mathbb{P}'$ where

$$\frac{d\mathbb{P}'}{d\mathbb{P}} \propto e^{-\|S_1\|^2}.$$

Let

$$f(\theta) = \mathbb{E}[e^{-\theta^T Z}] = -\mathbb{E}[U(X)],$$

where $U(x) = -e^{-x}$, $Z = S_1 - (1+r)S_0$ and $X = \theta^T Z$. Let θ_k be a minimizing sequence, so $f(\theta_k) \to \inf f(\theta)$.

1. If (θ_k) is bounded, by Bolzano-Weierstrass, there exists a convergent subsequence. So assume $\theta_k \to \theta^*$. As f is continuous, $f(\theta_k) \to f(\theta^*)$, so θ^* is a minimizer. Then

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \propto e^{-(\theta^*)^T(S_1 - (1+r)S_0)}$$

   is risk-neutral.

2. Now assume all minimising sequences are unbounded. Recall we assumed there was no arbitrage, meaning if $\phi^T Z \geq 0$ almost surely, then $\phi^T Z = 0$ almost surely.

   Now I claim we may assume $Z^1, \ldots, Z^d$ are linearly independent, meaning if $\phi^T Z = 0$ almost surely, then $\phi = 0$. If not, then we can restrict to a linearly independent submarket, then there is still no arbitrage in the submarket.

   Let (θ_k) be a minimising sequence with $\|\theta_k\| \to \infty$. Then let $\phi_k = \theta_k/\|\theta_k\|$, so by Bolzano-Weierstrass, we assume $\phi_k \to \phi_0$, where $\|\phi_0\| = 1$. Now we will show that $\phi_0^T Z \geq 0$ almost surely. Assuming this, no-arbitrage would imply that $\phi_0^T Z = 0$ almost surely, but by linear independence we have $\phi_0 = 0$, which contradicts $\|\phi_0\| = 1$.

We will now prove this claim. Since 0 is not a minimiser, $f(0) = 1 \geq f(\theta_k)$ for k large enough. But

$$f(\theta_k) = \mathbb{E}[e^{-\theta_k^T Z}] \geq \mathbb{E}[e^{-(\phi_k^T Z)\|\theta_k\|} \mathbb{1}_{\{\phi_0^T Z < -\varepsilon, \|Z\| < r\}}]$$

for arbitrary $\varepsilon > 0$, $r > 0$. Now note

$$\phi_k^T Z = (\phi_k - \phi_0)^T Z + \phi_0^T Z \leq \|\phi_k - \phi_0\| r - \varepsilon \leq -\frac{\varepsilon}{2},$$

by choosing k large enough. Hence

$$1 \geq f(\theta_k) \geq e^{\varepsilon\|\theta_k\|/2} \mathbb{P}(\phi_0^T Z < -\varepsilon, \|Z\| < r),$$

for all $k \geq N$. Hence

$$\mathbb{P}(\phi_0^T Z < -\varepsilon, \|Z\| < r) \leq e^{-\varepsilon/2\|\theta_k\|} \to 0$$

as $k \to \infty$. Hence taking $r \to \infty$ and $\varepsilon \to 0$, $\mathbb{P}(\phi_0^T Z < 0) = 0$.

5.2   No-arbitrage Pricing of Contingent Claims

Theorem 5.4. *Given a market with prices S_0, S_1 and interest r, assume there is no arbitrage. Introduce a claim with payout Y and initial price π.*

There is no arbitrage in the augmented market if and only if there exists $\mathbb{Q}$ risk-neutral for the original market such that

$$\pi = \frac{1}{1+r}\mathbb{E}_{\mathbb{Q}}[Y].$$

Then the set of all possible no-arbitrage prices π is an interval.

Proof: The first part is the fundamental theorem of asset pricing.

For the last part, we show the possible set of prices is convex. Let

$$\pi_i = \frac{\mathbb{E}_{\mathbb{Q}_i}[Y]}{1+r},$$

for $i = 0, 1$ and risk-neutral measures $\mathbb{Q}_i$. Let $0 < p < 1$, then let $\mathbb{Q}_p$ have density

$$p\frac{d\mathbb{Q}_1}{d\mathbb{P}} + (1-p)\frac{d\mathbb{Q}_0}{d\mathbb{P}} = \frac{d\mathbb{Q}_p}{d\mathbb{P}}.$$

Note that
$$\frac{\mathbb{E}_{\mathbb{Q}_p}[S_1]}{1+r} = p\frac{\mathbb{E}_{\mathbb{Q}_1}[S_1]}{1+r} + (1-p)\frac{\mathbb{E}_{\mathbb{Q}_0}[S_1]}{1+r} = S_0,$$
so $\mathbb{Q}_p$ is risk-neutral. Then

$$p\pi_1 + (1-p)\pi_0 = \frac{\mathbb{E}_{\mathbb{Q}_p}[Y]}{1+r}$$

is a no-arbitrage price of the claim.

Remark. The marginal utility price is a no-arbitrage price, since $U'(X_1^*)$ is proportional to the density of a risk-neutral measure.

5.3   Attainable Claims

Definition 5.4. A claim is *attainable* if there exists $a \in \mathbb{R}$ and $b \in \mathbb{R}^d$ such that

$$Y = a + b^T S_1 = (1+r)x + b^T(S_1 - (1+r)S_0),$$

where
$$x = \frac{a}{1+r} + b^T S_0.$$

For an attainable claim, the indifference price is x, for all U and X_0. Hence the marginal utility price is the same for all U and X_0.

Theorem 5.5 (Uniqueness of No-Arbitrage Price)**.** *If Y is an attainable claim, then there is a unique no-arbitrage price.*

Conversely, if Y has a unique no-arbitrage price, then it is attainable.

Proof: For the first part, suppose $Y = a + b^T S_1$. Then

$$\frac{1}{1+r}\mathbb{E}_{\mathbb{Q}}[Y] = \frac{a}{1+r} + b^T S_0$$

for all risk-neutral measures $\mathbb{Q}$.

The second part is in the example sheets.

We look at some examples of attainable claims.

Example 5.3.

Let's look at the forward contract, which is the obligation to buy the stock at time 1 for a fixed price K.

Then the payout is $Y = S_1 - K$. Therefore the unique no arbitrage price is

$$S_0 - \frac{k}{1+r}.$$

In reality, we should account for dividends, which are incorporated in the time 0 price, but not the time 1 price.

Moreover, in real markets, K is set so that the time 0 price is $\pi = 0$, i.e. $K = (1+r)S_0$.

Example 5.4.

Let's look at the one-period binomial model, with $-1 < a < b$, $0 < p < 1$ and interest rate r given. Then

$$S_1 = \begin{cases} S_0(1+b) & \text{with probability } p, \\ S_0(1+a) & \text{with probability } 1-p. \end{cases}$$

Consider a claim with payout $Y = g(S_1)$. This is attainable, since if we let

$$X_1 = (1+r)X_0 + \theta(S_1 - (1+r)S_0) = g(S_1),$$

then we must have

$$(1+r)X_0 + \theta S_0(b-r) = g(S_0(1+b)),$$
$$(1+r)X_0 + \theta S_0(a-r) = g(S_0(1+a)).$$

Hence we can find

$$\theta = \frac{g(S_0(1+b)) - g(S_0(1+a))}{S_0(b-a)},$$

and also X_0. There is only no arbitrage in the model if $a < r < b$. Let $\mathbb{Q}$ be a risk-neutral probability with

$$\mathbb{Q}(S_1 = (1+b)S_0) = q = 1 - \mathbb{Q}(S_1 = (1+a)S_0).$$

Then we have

$$qS_0(1+b) + (1-q)S_0(1+a) = \mathbb{E}_{\mathbb{Q}}[S_1] = (1+r)S_0,$$

hence we get
$$q = \frac{r-a}{b-a}.$$

Note that $a < r < b \iff 0 < q < 1$. Now to find X_0, note that

$$\mathbb{E}_{\mathbb{Q}}[g(S_1)] = \mathbb{E}_{\mathbb{Q}}[(1+r)S_0 + \theta(S_1 - (1+r)S_0)] = (1+r)X_0,$$

hence the no-arbitrage price is

$$\pi = \frac{1}{1+r}\mathbb{E}_{\mathbb{Q}}[g(S_1)] = \frac{qg(S_0(1+b)) + (1-q)g(S_0(1+a))}{1+r}.$$

6   Multi-Period Models

The challenge for multi-period models is to keep track of the information that is revealed at each time period. To do this, we use a filtered probability space: we will let $\mathcal{F}_0, \mathcal{F}_1, \ldots = (\mathcal{F}_n)$ be the σ-algebras of information available at time n. We assume that $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$.

To informally motivate why we consider σ-algebras, we let $\mathcal{G}$ be some set of information. Then we can define $\mathbb{P}(A \mid \mathcal{G})$, which is the conditional probability of A happening given $\mathcal{G}$.

We say that A is $\mathcal{G}$-measurable if $\mathbb{P}(A \mid \mathcal{G}) \in \{0, 1\}$.

Example 6.1.

Let $\Omega = \{HH, HT, TH, TT\}$ be the outcome of two coin flips, and $\mathcal{G}$ be the information available after the first flip.

Then $A = \{HH, HT\}$ is $\mathcal{G}$-measurable, because we know whether it happened or not after the first flip, whereas $B = \{HH\}$ is not $\mathcal{G}$-measurable. If the first flip is heads, we still have to know about the second flip.

We can then define our σ-algebra to be the collection of measurable events (we assume it is a σ-algebra).

6.1   Measurability

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G} \subseteq \mathcal{F}$ a sub-σ-algebra.

Definition 6.1. $A \in \mathcal{F}$ is $\mathcal{G}$-measurable $\iff A \in \mathcal{G}$.

A random variable X is $\mathcal{G}$-measurable if and only if $\{X \leq x\}$ is $\mathcal{G}$-measurable for all $x \in \mathbb{R}$.

Remark. If X is $\mathcal{G}$-measurable, then given $\mathcal{G}$ we know the value of X.

Example 6.2.

Let Ω be our collection of two coin flips, and $\mathcal{G} = \{\emptyset, \Omega, \{HH, HT\}, \{TT, TH\}\}$. Define

$$X(\omega) = \begin{cases} 1 & \omega \in \{HH, HT\}, \\ 0 & \text{else.} \end{cases}$$

Then X is $\mathcal{G}$-measurable. However,

$$
Y(\omega) = \begin{cases} 2 & \omega = HH, \\ 1 & \omega \in \{HT, TH\}, \\ 0 & \omega = TT, \end{cases}
$$

then Y is not $\mathcal{G}$-measurable.

Definition 6.2. The σ-algebra generated by X, denoted $\sigma(X)$, is the smallest σ-algebra containing $\{X \leq x\}$ for all $x \in \mathbb{R}$.

Theorem 6.1. *Y is $\sigma(X)$-measurable if and only if $Y = f(X)$ for a measurable function $f : \mathbb{R} \to \mathbb{R}$.*

We will try to model the time t price vector S_t. First we need an increasing collection of information $(\mathcal{F}_t)$, and we want S_t to be $\mathcal{F}_t$-measurable.

Recall, for X taking values in $\{x_1, x_2, \ldots\}$, X is $\mathcal{G}$-measurable $\iff$ $\{X = x_i\} \in \mathcal{G}$. Moreover, X is $\mathcal{G}$-measurable $\iff$ there exists X_n $\mathcal{G}$-measurable discrete random variables, such that $X_n \to X$.

Recall $\sigma(X)$, the σ-algebra generated by X. The smallest such σ-algebra is $\{\emptyset, \Omega\}$.

In discrete time, we will try to extend the fundamental theorem of asset pricing in discrete time: no arbitrage exists if and only if there exists $\mathbb{Q} \sim \mathbb{P}$ such that

$$
\frac{\mathbb{E}^{\mathbb{Q}}[S_t | \mathcal{F}_{t-1}]}{1 + r} = S_{t-1},
$$

for all t. To make sense of this we will need to define conditional expectations.

Recall, that if X is integrable, then for any event A such that $\mathbb{P}(A) > 0$, we can define

$$
\mathbb{E}[X | A] = \frac{\mathbb{E}[X \mathbb{1}_A]}{\mathbb{P}(A)}.
$$

Also recall that if $Y = \{y_1, y_2, \ldots\}$, with $\mathbb{P}(Y = y_i) > 0$, then $\mathbb{E}[X | Y] = f(Y)$, a random variable in terms of Y, where $f(y) = \mathbb{E}[X | Y = y]$, for $y \in \{y_1, y_2, \ldots\}$.

Proposition 6.1. *In the above set up, we have the projection property:*

$$
\mathbb{E}[X \mathbb{1}_G] = \mathbb{E}[f(Y) \mathbb{1}_G],
$$

for any $\sigma(Y)$-measurable event G.

Proof: Let $G = \{Y \in B\}$, for $B \subseteq \{y_1, y_2, \ldots\}$. Then

$$
\begin{aligned}
\mathbb{E}[f(Y)\mathbb{1}_G] &= \mathbb{E}\left[\left(\sum_i \frac{\mathbb{E}[X\mathbb{1}_{\{Y=y_i\}}]}{\mathbb{P}(Y=y_i)}\mathbb{1}_{\{Y=y_i\}}\right)\mathbb{1}_{\{Y\in B\}}\right] \\
&= \sum_i \frac{\mathbb{E}[X\mathbb{1}_{\{Y=y_i\}}]}{\mathbb{P}(Y=y_i)}\mathbb{E}[\mathbb{1}_{\{Y=y_i\}}\mathbb{1}_{\{Y\in B\}}] \\
&= \sum_{y_i \in B} \mathbb{E}[X\mathbb{1}_{\{Y=y_i\}}] = \mathbb{E}\left[X\left(\sum_{y_i \in B}\mathbb{1}_{\{Y=y_i\}}\right)\right] \\
&= \mathbb{E}[X\mathbb{1}_G].
\end{aligned}
$$

This motivates the following.

Definition 6.3. Let X be integrable, and $\mathcal{G}$ a σ-algebra. An integrable $\mathcal{G}$-measurable random variable Z is a *conditional expectation* of X given $\mathcal{G}$ if and only if

$$
\mathbb{E}[X\mathbb{1}_G] = \mathbb{E}[Z\mathbb{1}_G],
$$

for all $G \in \mathcal{G}$.

Theorem 6.2. *Conditional expectations exist and are unique.*

Proof: Existence follows from analysis (the fact L^2 is complete). We prove uniqueness.

Suppose Z_0, Z_1 are two conditional expectations of X given $\mathcal{G}$:

$$
\mathbb{E}[X\mathbb{1}_G] = \mathbb{E}[Z_0\mathbb{1}_G] = \mathbb{E}[Z_1\mathbb{1}_G],
$$

for all $G \in \mathcal{G}$. So subtracting, $\mathbb{E}[(Z_0 - Z_1)\mathbb{1}_G] = 0$. But now we can take anything to be G.

Take first $G = \{Z_0 \geq Z_1\}$, Then $(Z_0 - Z_1)\mathbb{1}_G \geq 0$, but since the expectation is 0, we have $\mathbb{P}(Z_0 \leq Z_1) = 1$. Similarly $\mathbb{P}(Z_1 \leq Z_0) = 1$, so $Z_0 = Z_1$ a.e.

The expectation of X given $\mathcal{G}$ is $\mathbb{E}[X \mid \mathcal{G}]$. We can also define conditional expectations in terms of random variables, as

$$
\mathbb{E}[X \mid Y] = \mathbb{E}[X \mid \sigma(Y)],
$$

which agrees with our example from before.

Moreover, if X is square-integrable, and Y is square-integrable and $\mathcal{G}$-measurable, then the projection property gives

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

Therefore, $\mathbb{E}[X \mid \mathcal{G}]$ minimizes $\mathbb{E}[(X - Z)^2]$ over all $\mathcal{G}$-measurable random variables Z. We can think of this conditional expectation, thus, as the "best guess" for X given our information $\mathcal{G}$.

Proof: Let Z be $\mathcal{G}$-measurable, and $Y = -Z + \mathbb{E}[X \mid \mathcal{G}]$. Then

$$\begin{aligned}
\mathbb{E}[(X - Z)^2] &= \mathbb{E}[(X - \mathbb{E}[X \mid \mathcal{G}] + Y)^2] \\
&= \mathbb{E}[(X - \mathbb{E}[X \mid \mathcal{G}])^2] + \mathbb{E}[Y^2] + 2\underbrace{\mathbb{E}[(X - \mathbb{E}[X \mid \mathcal{G}])Y]}_{0} \\
&\geq \mathbb{E}[(X - \mathbb{E}[X \mid \mathcal{G}])^2].
\end{aligned}$$

Properties of conditional expectation are:

1. Linearity. Given constant a, b and integrable random variables X, Y, then

$$\mathbb{E}[aX + bY \mid \mathcal{G}] = a\mathbb{E}[X \mid \mathcal{G}] + b\mathbb{E}[Y \mid \mathcal{G}].$$

2. Non-negative. If $X \geq 0$ almost surely, then $\mathbb{E}[X \mid \mathcal{G}] \geq 0$.

3. Pulling out what is known. If Y is $\mathcal{G}$-measurable, then

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

4. Independence. If X is independent of $\mathcal{G}$, i.e. $\sigma(X)$ independent of $\mathcal{G}$, then

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

5. Conditional Jensen. If f is convex, then

$$\mathbb{E}[f(X) \mid \mathcal{G}] \geq f(\mathbb{E}[X \mid \mathcal{G}]).$$

6. Holding known random variable fixed and averaging over an independent one. Suppose X is independent of $\mathcal{G}$, and Y is $\mathcal{G}$-measurable. Then,

$$\mathbb{E}[f(X, Y) \mid \mathcal{G}] = \mathbb{E}[f(X, y)]\Big|_{y=Y}.$$

Example 6.3.

Say X, Y are independent $N(0,1)$, and $\mathcal{G} = \sigma(Y)$. Let ϕ be the pdf of a $N(0,1)$ random variable. Then,

$$\mathbb{E}[f(X,Y) \mid \mathcal{G}] = \int f(x,y)\phi(x)\,\mathrm{d}x\bigg|_{y=Y} = \int f(x,Y)\phi(x)\,\mathrm{d}x.$$

The proof of these is using the definitions. For Jensen's, we redo the proof of unconditional Jensen using the supporting line theorem.

7   Martingales

Definition 7.1. A family $(\mathcal{F}_t)$ of σ-algebras, such that $\mathcal{F}_s \subseteq \mathcal{F}_t$ for all $s \leq t$ is called a *filtration*.

In this course, we use the convention that $\mathcal{F}_0 = \{\emptyset, \Omega\}$. If X is $\mathcal{F}_0$-measurable, then X is constant. Also $\mathbb{E}[Y \mid \mathcal{F}_0] = \mathbb{E}[Y]$ for any Y.

Definition 7.2. A *stochastic process* is a collection of random variables (X_t). It is *adapted* to a filtration $(\mathcal{F}_t)$ if and only if X_t if $\mathcal{F}_t$-measurable for all $t \geq 0$.

Definition 7.3. The filtration generated by a process (X_t) is $\mathcal{F}_t = \sigma(X_s, 0 \leq s \leq t)$.

This is the smallest filtration such that the process is adapted.

Definition 7.4. A *martingale* (X_t) with respect to a filtration $(\mathcal{F}_t)$ is an adapted, integrable process such that
$$\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s$$
almost surely, for all $0 \leq s \leq t$.

It is equivalent to check that $\mathbb{E}[X_t - X_s \mid \mathcal{F}_s] = 0$, since X_s is $\mathcal{F}_s$ measurable hence the conditional expectation is itself.

Proposition 7.1 (Tower Property). *If $\mathcal{G} \subseteq \mathcal{H}$, then*
$$\mathbb{E}[\mathbb{E}[Y \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[\mathbb{E}[Y \mid \mathcal{H}] \mid \mathcal{G}] = \mathbb{E}[Y \mid \mathcal{G}].$$

Proposition 7.2. *In discrete time, (X_t) is a martingale with respect to $(\mathcal{F}_t)$ if and only if*
$$\mathbb{E}[X_t \mid \mathcal{F}_{t-1}] = X_{t-1},$$
for all $t \geq 1$.

> **Proof:**   If we know (X_t) is a martingale, then the above holds, using the definition with $s = t - 1$.
>
> Conversely, if the above holds for all $t \geq 1$, then note that
> $$\mathbb{E}[X_{s+k} \mid \mathcal{F}_s] = \mathbb{E}[\mathbb{E}[X_{s+k} \mid \mathcal{F}_{s+k-1}] \mid \mathcal{F}_s] = \mathbb{E}[X_{s+k-1} \mid \mathcal{F}_f],$$
> so the martingale property holds by induction.

Our motivation for this is as follows: the fundamental theorem of asset pricing says that there is no arbitrage if and only if there exists $\mathbb{Q} \sim \mathbb{P}$ such that
$$\mathbb{E}^{\mathbb{Q}}\left[\frac{S_t}{1+r} \mid \mathcal{F}_{t-1}\right] = S_{t-1},$$

which, rephrasing, says that

$$\left(\frac{S_t}{(1+r)^t}\right)$$

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

Example 7.1. (Doob Martingale)

Let Y be an integrable random variable, and given a filtration $(\mathcal{F}_t)$, let $X_t = \mathbb{E}[Y \mid \mathcal{F}_t]$. Then (X_t) is a martingale.

This follows from the tower law:

$$\mathbb{E}[X_t \mid \mathcal{F}_s] = \mathbb{E}[\mathbb{E}[Y \mid \mathcal{F}_t] \mid \mathcal{F}_s] = \mathbb{E}[Y \mid \mathcal{F}_s] = X_s.$$

7.1   Discrete Time Martingales

We look at another example of martingales. Let $X_1, X_2, \ldots$ be independent with $\mathbb{E}[X_n] = 0$ for all n, and integrable.

Let $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. Then define $S_0 = 0$ and $S_n = X_1 + \cdots + X_n$.

We claim that (S_n) is a martingale. These are adapted as S_n is a function of $X_1, \ldots, X_n$, hence $\mathcal{F}_n$-measurable, and integrable as it is a sum of integrable random variables.

Moreover the martingale property is true, since

$$\mathbb{E}[S_n - S_{n-1} \mid \mathcal{F}_{n-1}] = \mathbb{E}[X_n \mid \mathcal{F}_{n-1}] = \mathbb{E}[X_n] = 0,$$

as X_n is independent of all other random variables.

Definition 7.5. A (discrete-time) process (H_n) is *previsible* or *predictable* if H_n is $\mathcal{F}_{n-1}$-measurable.

Definition 7.6. The *martingale transform* of a previsible process (H_n) with respect to an adapted process (X_n) is

$$Y_0 = 0, \qquad Y_n = \sum_{k=1}^{n} H_k(X_k - X_{k-1}).$$

This is also known as a discrete time stochastic integral.

This martingale transform is adapted, as Y_n is a function of $\mathcal{F}_n$-measurable random variable.

Theorem 7.1. *The martingale transform of a bounded previsible process with respect to a martingale is a martingale.*

> **Proof:** We already checked adaptedness. Y_n is integrable since
>
> $$|Y_n| \leq \sum_{k=1}^{n} |H_k||X_k - X_{k-1}| \leq C \sum_{k=1}^{n} |X_k - X_{k-1}|,$$
>
> where $|H_k| \leq C$. Moreover $|X_k|$ is integrable by definition, since it is a martingale. Finally,
>
> $$\mathbb{E}[Y_n - Y_{n-1} \mid \mathcal{F}_{n-1}] = \mathbb{E}[H_n(X_n - X_{n-1}) \mid \mathcal{F}_{n-1}] = H_n \mathbb{E}[X_n - X_{n-1} \mid \mathcal{F}_{n-1}] = 0,$$
>
> since (X_n) is a martingale.

Lets see a couple of applications. The first is from finance. Beginning with the interest rate r, with d assets and where asset i has price S_n^i at time n.

Say that between time $n-1$ and time n the investor holds θ_n^0 in the bank, and θ_n^i shares of asset i. Let X_n be the wealth of the investor at time n. Then,

$$X_{n-1} = \theta_n^0 + \theta_n^T S_{n-1},$$
$$X_n = \theta_n^0(1 + r) + \theta_n^T S_n.$$

This is known as the self-financing condition. Eliminating θ_n^0, we get

$$X_n = (1 + r)X_{n-1} + \theta_n^T(S_n - (1 + r)S_{n-1}),$$

and hence

$$\frac{X_n}{(1 + r)^n} = \frac{X_{n-1}}{(1 + r)^{n-1}} + \theta_n^T \left(\frac{S_n}{(1 + r)^n} - \frac{S_{n-1}}{(1 + r)^{n-1}} \right),$$

therefore we find that

$$\frac{X_n}{(1 + r)^n} = X_0 + \sum_{k=1}^{n} \theta_k^T \left(\frac{S_n}{(1 + r)^n} - \frac{S_{n-1}}{(1 + r)^{n-1}} \right).$$

So the discounted wealth of the investor is the initial wealth, plus the martingale transform of the portfolio process with respect to discounted asset prices (here $X_n/(1 + r)^n$ is the discounted wealth, and $S_n/(1 + r)^n$ is the discounted asset prices).

We will assume that θ_n is $\mathcal{F}_{n-1}$-measurable to avoid clairvoyance, i.e. (θ_n) is previsible.

7.2   Stopping Times

Definition 7.7. A *stopping time* for a filtration $(\mathcal{F}_t)$ is a random variable T defined in $\mathbb{N} \cup \{\infty\}$ for discrete time, or $[0, +\infty]$ for continuous time, such that $\{T \leq t\}$ is $\mathcal{F}_t$-measurable for all $t \geq 0$.

> **Example 7.2.**
>
> In discrete time, let (X_n) be adapted. Let $T = \inf\{n \geq 0 \mid X_n > 0\}$, the first time X_n is positive, with the convention that $\inf \emptyset = \infty$. Then T is a stopping time. Indeed,
>
> $$\{T \leq n\} = \bigcup_{k=0}^{n} \{X_k > 0\} \in \mathcal{F}_n$$
>
> for all n.
>
> A non-example is as follows: let $T = \sup\{n \geq 0 \mid X_n > 0\}$, i.e. the last time that X_n is positive, where $\sup \emptyset = 0$. Then
>
> $$\{T \leq n\} = \bigcap_{k \geq n+1} \{X_k \leq 0\},$$
>
> is not necessarily measurable.

7.3   Optimal Stopping

Definition 7.8. Given an adapted process (X_t) and a stopping time T, the *stopped process* is $(X_{t \wedge T})$, i.e.

$$X_{t \wedge T} = \begin{cases} X_t & \text{for } t \leq T, \\ X_T & \text{for } t \geq T. \end{cases}$$

From now on, time is discrete.

Proposition 7.3. *The stopped process $(X_{n \wedge T})$ is X_0, plus the martingale transform of $(1_{\{n \leq T\}})$ with respect to (X_n).*

> **Proof:**   Indeed, we have
>
> $$X_0 + \sum_{k=1}^{n} 1_{\{k \leq T\}}(X_k - X_{k-1}) = \begin{cases} X_n & \text{if } T > n, \\ X_T & \text{if } T \leq n. \end{cases}$$

We also have to check that the integrand is previsible: indeed, $\{n \leq T\} = \{T \leq n-1\}^c \in \mathcal{F}_{n-1}$, so $(\mathbb{1}_{\{n \leq T\}})$ is previsible.

Theorem 7.2. *Stopped martingales are martingales.*

Proof: Note $(\mathbb{1}_{\{n \leq T\}})$ is bounded, and martingale transforms of a bounded previsible process is a martingale.

Theorem 7.3 (Optimal Stopping Theorem)**.** *Let (X_n) be a martingale and T a stopping time, such that $T < \infty$ almost surely. Suppose that $(X_{n \wedge T})$ is bounded. Then $\mathbb{E}[X_T] = X_0$.*

Note there is nothing to prove if T is not random; this follows directly from the definition of a martingale.

Proof: Note that $\mathbb{E}[X_{n \wedge T}] = X_0$, since $(X_{n \wedge T})$ is a martingale.

By assumption, $|X_{n \wedge T}| < C$ almost surely for all n, where $C > 0$ is a constant. Note $X_{n \wedge T} \to X_T$ almost surely as $n \to \infty$, since $T < \infty$ almost surely, so $|X_T| < C$ almost surely. Then

$$|\mathbb{E}[X_T] - X_0| = |\mathbb{E}[X_T - X_{T \wedge n}]| \leq \mathbb{E}[|X_T - X_{T \wedge n}|]$$
$$\leq 2C\mathbb{P}(T > n) \to 0,$$

as $n \to \infty$.

Remark. The condition $(X_{n \wedge T})$ is bounded can be replaced by $(X_{n \wedge T})$ is uniformly integrable. Another useful criterion is $|X_{n \wedge T}| \leq Z$ almost surely, for all n, where $\mathbb{E}[Z] < \infty$.

Example 7.3.

Let (X_n) be a simple, symmetric random walk on $\mathbb{Z}$, starting at $X_0 = 0$, meaning $X_n = \xi_1 + \xi_2 + \cdots + \xi_n$, where (ξ_n) is iid, with $\mathbb{P}(\xi_n = 1) = \mathbb{P}(\xi_n = -1) = 1/2$.

Fix $a, b > 0$ integers, and let $T = \inf\{n \geq 0 \mid X_n \in \{-a, b\}\}$. This is a stopping time, since it is the first hitting time of an adapted process. Moreover $|X_{n \wedge T}| \leq \max\{a, b\}$ almost surely, and by Markov chains $T < \infty$ almost surely.

So $\mathbb{E}[X_T] = X_0$ by optimal stopping theorem, hence $\mathbb{P}(X_T = -a)(-a) +$

$\mathbb{P}(X_T = b)b = 0$, allowing us to solve

$$\mathbb{P}(X_T = b) = \frac{a}{a + b} = 1 - \mathbb{P}(X_T = -a).$$

For a non-example, in the same setting let $\tau = \inf\{n \geq 0 \mid X_n = -a\}$. Again this is a stopping time, and $\tau < \infty$ almost surely, but

$$\mathbb{E}[X_\tau] = -a \neq 0 = X_0.$$

This is as $(X_{n \wedge \tau})$ is not bounded.

Example 7.4.

We now try to calculate $\mathbb{E}[z^\tau]$, where $0 < z < 1$, and τ is the same stopping time as before. From example sheets, we know $(z^n w^{X_n})$ is a martingale, for some w. Indeed,

$$\mathbb{E}[z^n w^{X_n} \mid \mathcal{F}_{n-1}] = z^n w^{X_{n-1}} \mathbb{E}[w^{\xi_n}] = z \left(\frac{w + 1/w}{2} \right) z^{n-1} w^{X_{n-1}},$$

so w solves $w + 1/w = 2/z$. Hence by optimal stopping theorem,

$$\mathbb{E}[z^\tau w^{X_\tau}] = z^0 w^0 = 1 \implies \mathbb{E}[z^\tau] = w^a.$$

However there are two possible values of w. We want $z^{n \wedge \tau} w^{X_{n \wedge \tau}}$ to be bounded. To do this we use the value of w that is less than 1.

7.4   Super and Sub-Martingales

Definition 7.9. An adapted integrable process (X_t) is a *supermartingale* if and only if $X_s \geq \mathbb{E}[X_t \mid \mathcal{F}_s]$ almost surely, and a *submartingale* if and only if $X_s \leq \mathbb{E}[X_t \mid \mathcal{F}_s]$ almost surely, for all $0 \leq s \leq t$.

Remark. The supermartingale condition is equivalent to $\mathbb{E}[X_t - X_s \mid \mathcal{F}_s] \leq 0$ almost surely, for all $0 \leq s \leq t$.

In discrete time, this is equivalent to $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \leq X_n$ almost surely, for all $n \geq 0$.

Theorem 7.4. *The martingale transform of a non-negative bounded previsible process with respect to a supermartingale is a supermartingale.*

Proof: We sketch the proof. Let

$$Y_k = \sum_{k=1}^{n} H_k(X_k - X_{k-1}),$$

where X is a supermartingale, and H is bounded, non-negative and previsible. Then boundedness is easy, and

$$\mathbb{E}[Y_n - Y_{n-1} \mid \mathcal{F}_n] = \mathbb{E}[H_n(X_n - X_{n-1}) \mid \mathcal{F}_{n-1}] = H_n\mathbb{E}[X_n - X_{n-1} \mid \mathcal{F}_{n-1}] \leq 0.$$

Theorem 7.5. *Let (X_n) be a supermartingale. Let S and T be stopping times such that $0 \leq S \leq T$ almost surely. Then $(X_{n \wedge T} - X_{n \wedge S})$ is a supermartingale.*

Proof: Note that

$$X_{n \wedge T} - X_{n \wedge S} = \sum_{k=1}^{n} \mathbb{1}_{\{S < k \leq T\}}(X_k - X_{k-1}).$$

The indicator functions are bounded, non-negative and $\mathcal{F}_{k-1}$-measurable, so this is a martingale transform, and then the last theorem applies.

Theorem 7.6 (Optional Sampling Theorem). *Let (X_n) be a supermartingale, and S, T be bounded stopping times such that $0 \leq S \leq T$. Then*

$$\mathbb{E}[X_T] \leq \mathbb{E}[X_S].$$

Remark. If X is a martingale, then $\mathbb{E}[X_T] = \mathbb{E}[X_S] = X_0$ for bounded S, T.

Moreover if S, T are not bounded, but $(X_{n \wedge T})$ is bounded, then $\mathbb{E}[X_T] \leq \mathbb{E}[X_S]$ by the argument from the optional stopping theorem.

Proof: Suppose $0 \leq S \leq T \leq N$ almost surely. Let $Y_n = X_{n \wedge T} - X_{n \wedge S}$. Then Y is a supermartingale, from before, but notice that $Y_N = X_T - X_S$.

Then we claim that $\mathbb{E}[Y_n] \leq Y_0 = 0$. But this follows from the definition of a supermartingale, with $s = 0$ and $t = N$.

8   Controlled Markov Processes

Definition 8.1. A process (X_t) is *Markov* for a filtration if (X_t) is adapted, and

$$\mathbb{P}(X_t \in A \mid \mathcal{F}_s) = \mathbb{P}(X_t \in A \mid X_s)$$

almost surely, for all $0 \leq s \leq t$ and measurable subsets A.

In discrete time, it suffices that

$$\mathbb{P}(X_n \in A \mid \mathcal{F}_{n-1}) = \mathbb{P}(X_n \in A \mid X_{n-1}).$$

In discrete time, we have the following equivalent view: there exists a sequence of independent (ξ_n) and a function G such that $X_n = G(n, X_{n-1}, \xi_n)$ for all n.

Definition 8.2. A *controlled Markov process* (X_n^U), given a previsible (U_n) is such that

$$X_n^U = G(n, X_{n-1}^U, U_n, \xi_n),$$

for some function G and independent sequence (ξ_n).

Here U is the control variable.

Example 8.1.

Given X_0 and a time horizon N, we might want to find U to maximize

$$\mathbb{E}\left[g(X_N^U) + \sum_{k=1}^{N} f(k, U_k) \right].$$

The *Bellman equation* associated to this problem is:

$$V(N, x) = g(x),$$
$$V(n-1, x) = \max_u \mathbb{E}\left[V(n, G(n, x, u, \xi_n)) + f(n, u) \right],$$

for all $1 \leq n \leq N$.

Definition 8.3. The *value function* for the problem is the function

$$V(n, x) = \sup_{(U_k)} \mathbb{E}\left[g(X_N^U) + \sum_{k=n+1}^{N} f(k, U_k) \mid X_n^U = x \right].$$

Under certain conditions, the solution to the Bellman equation is the value function.

8.1 Dynamic Programming Principle

Theorem 8.1 (Dynamic Programming Principle). *Let V solve the associated Bellman equations:*

$$V(N, x) = g(x),$$
$$V(n-1, x) = \max_u \mathbb{E}\left[V(n, G(n, x, u, \xi_n)) + f(n, u)\right], \qquad (*)$$

for all x and for all $1 \le n \le N$. Let $u^*(n, x)$ be the maximizer of $(*)$.

Let $X_0^* = X_0$, $U_n^* = u^*(n, X_{n-1}^*)$ and $X_n^* = G(n, X_{n-1}^*, U_n^*, \xi_n)$, so $X_n^{U^*} = X_n^*$.

Then U^* is the optimal control, and V is the value function for the problem.

Proof: We assume that everything is integrable and measurable. To do this, we use the 'martingale principle of optimal control': we find a process which is a martingale for the optimal control, and a supermartingale in general.

In our case, let

$$M_n^U = \sum_{k=1}^{n} f(n, U_k) + V(n, X_n^U).$$

Then we claim M^U is a supermartingale, and M^{U^*} is a martingale. This is as

$$
\begin{aligned}
&\mathbb{E}[M_n - M_{n-1} \mid \mathcal{F}_{n-1}] \\
&\quad = \mathbb{E}[f(n, U_n) + V(n, X_n^U) - V(n-1, X_{n-1}^U) \mid \mathcal{F}_{n-1}] \\
&\quad = \mathbb{E}[f(n, U_n) + V(n, G(n, X_{n-1}^U, U_n, \xi_n)) \mid \mathcal{F}_{n-1}] - V(n-1, X_{n-1}^U) \\
&\quad = \mathbb{E}[f(n, u) + V(n, G(n, x, u, \xi_n))]\Big|_{\substack{u=U_n \\ x=X_{n-1}^U}} - V(n-1, X_{n-1}^U) \\
&\quad \le f(n, u^*(n, x)) + \mathbb{E}[V(n, G(n, x, u^*(n, x), \xi_n))]\Big|_{x=X_{n-1}^U} - V(n-1, X_{n-1}^U) \\
&\quad \le 0,
\end{aligned}
$$

with equality if and only if $U = U^*$. So

$$\mathbb{E}\left[\sum_{k=n+1}^{N} f(k, U_k) + g(X_N^U) \mid X_n^U = x\right]$$

$$= \mathbb{E}\left[M_N^U - \sum_{k=1}^{n} f(k, u_k) \mid X_n^U = x\right]$$

$$= \mathbb{E}\left[\mathbb{E}\left[M_N^U - \sum_{k=1}^{n} f(k, u_k) \mid \mathcal{F}_n\right] \mid X_n^u = x\right]$$

$$\leq \mathbb{E}\left[M_n^U - \sum_{k=1}^{n} f(k, U_k) \mid X_n^U = x\right]$$

$$= \mathbb{E}[V(n, X_n^U) \mid X_n^U = x] = V(n, x),$$

with equality if and only if $U = U^*$.

Example 8.2.

We look at an optimal investment problem. Take $d = 1$ risky asset, and suppose $S_n = S_{n-1} Z_n$, where (Z_n) are independent, with interest rate r.

Between times $n - 1$ and n, the investor consumes c_n between 0 and X_{n-1}, and holds θ_n shares, so the wealth equation is

$$X_n = (1 + r)(X_{n-1} - c_n) + \theta_n(S_n - (1 + r)S_{n-1}) = G(n, X_{n-1}, U_n, \xi_n),$$

where $U_n = (c_n, S_{n-1}\theta_n)$ and

$$G(n, X, (c, \eta), \xi) = (1 + r)(x - c) + \eta(\xi - (1 + r)),$$

where $\eta = S_{n-1}\theta_n$. Our goal is to maximize

$$\mathbb{E}\left[\sum_{k=1}^{n} u(c_k) + u(X_N)\right]$$

over (c_n, θ_n) given X_0, where u is a given utility function. The Bellman equation is then

$$V(N, x) = u(x),$$
$$V(n - 1, x) = \max_{\substack{0 \leq c \leq x \\ \eta}} \left[u(c) + \mathbb{E}\left[V(n, (1 + r)(x - c) + \eta(\xi_n - (1 + r)))\right]\right],$$

for all x and for all $1 \leq n \leq N$. This is generally intractible, but our idea is to let u be a specific utility function, e.g. the CRRA utility.

In this case, we have

$$u(x) = \frac{x^{1-R}}{1-R},$$

for some value $R > 0$, $R \neq 1$. We guess that $V(n,x) = A_n u(x)$, for all $0 \leq n \leq N$, and where A_n is to be determined. This works for $n = N$ with $A_N = 1$. Suppose it works for $n = k$. Then the right hand side is

$$\max_{c,\eta} \left[u(c) + A_k \mathbb{E}\left[u((1+r)(x-c) + \eta(\xi - (1+r))) \right] \right]$$

$$= u(x)(1-R) \max_{c,\eta} \left[u\left(\frac{c}{x}\right) + u\left(1 - \frac{c}{x}\right) A_k(1-R) \right.$$

$$\left. \times \mathbb{E}\left[u\left((1+r) + \frac{\eta}{x-c}(c - (1+r)) \right) \right] \right].$$

Let $\alpha = (1-R) \max \mathbb{E}U(1 + r + t(\xi - (1+r)))$, and t^* be the argmax. Letting $t = \frac{\eta}{x-c}$, and $s = \frac{c}{x}$, then this becomes

$$u(x)(1-R) \max_{0 \leq s \leq 1} [u(s) + u(1-s)A_k \alpha].$$

At the maximum, the derivative will be 0. Hence

$$s^{-R} = (1-s)^{-R}(1-R)A_K \alpha \implies S_k^* = \frac{1}{1 + (\alpha A_k)^{1/k}}.$$

Putting this in again, we get $u(x)(1 + (A_k\alpha)^{1/R})^R$. We want this to equal $u(x)A_{k-1}$. Hence solving

$$A_{k-1} = (1 + (A_k\alpha)^{1/R})^R, \qquad A_N = 1,$$

this gives us

$$A_n = \left(\frac{1 - \alpha^{(N-n+1)/R}}{1 - \alpha^{1/R}} \right)^R,$$

for all $0 \leq n \leq N$. Plugging this back into S_n^*,

$$c_n^* = X_{n-1}^* S_n^* = \frac{X_{n-1}^*}{1 + (A_n\alpha)^{1/R}} = \frac{X_{n-1}^\alpha (1 - \alpha^{1/R})}{1 - \alpha^{(N+n)/R}},$$

which gives us

$$\theta_n^* = \frac{\eta_n^*}{S_{n-1}} = \frac{t^*(X_{n-1}^* - c_n^*)}{S_{n-1}},$$

which is calculable. Therefore, we can calculate the optimal wealth:

$$X_n^* = (1 + r)(X_{n-1}^* - c_n^*) + \eta_n^*(\xi_n - (1 + r))$$

$$= X_{n-1}^* \frac{\alpha^{1/R}(1 - \alpha^{(N-n+1)/R})}{(1 - \alpha^{(N-n+1)/R})}[(1 + r) + t^*(\xi_n - (1 + r))],$$

or something like that.

8.2   Infinite Horizon Problem

We set $X_n = G(X_{n-1}, U_n, \xi_n)$, where (ξ_n) are IID. Notice that there is no n-dependence in this.

Our goal is now to maximize

$$\mathbb{E}\left[\sum_{k=1}^{\infty} \beta^{k-1} f(U_k)\right],$$

where $0 < \beta < 1$ is the *subjective discount factor*. The Bellman equation now relates to

$$V(x) = \max\{f(u) + \beta \mathbb{E} V(G(x, u, \xi))\}.$$

Theorem 8.2. *Suppose $f(u) \geq 0$ for all u, and $u^*(x) = \mathrm{argmax}(f(u) + \beta \mathbb{E} V(x, u, \xi))$, where V is the solution to the Bellman equation.*

Then given X_0, let $U_n^* = u^*(X_{n-1}^*)$, $X_0^* = X_0$, and $X_n^* = G(X_{n-1}^*, U_n^*, \xi_n)$.

Suppose that $\beta^n \mathbb{E} V(X_n^*) \to 0$. Then U^* is optimal, and V is the value function.

To prove this we need the monotone convergence theorem: if (Z_n) is a sequence of random variables such that $0 \leq Z_n \leq Z_{n+1}$ almost surely, then let $Z_\infty = \sup Z_n = \lim Z_n$. Then

$$\mathbb{E}[Z_\infty] = \lim \mathbb{E}[Z_n].$$

Proof: We prove that V is the value function. Pick a control U, and let

$$M_n^U = \sum_{k=1}^{n} \beta^{k-1} f(U_k) + \beta^n V(X_n^U).$$

We claim that this is a supermartingale for all U, and a martingale when

$U = U^*$. Indeed,

$$\mathbb{E}[M_n - M_{n-1} \mid \mathcal{F}_{n-1}] = \beta^{n-1}[f(U) + \beta\mathbb{E}[V(X_n^U) \mid \mathcal{F}_{n-1}] - V(X_{n-1}^U)]$$
$$\leq 0,$$

by the Bellman equation, with equality if $U = U^*$ by the definition of U^*. Hence

$$V(X) = M_0^U \geq \mathbb{E}[M_N^U] = \mathbb{E}\left[\sum_{k=1}^{n} \beta^{k-1} f(U_k)\right] + \beta^n \mathbb{E}[V(X_N^U)],$$

with equality if $U = U^*$. Since $V(x) \geq 0$ for all x, and $f(u) \geq 0$, we can use MCT:

$$V(X) \geq \mathbb{E}\sum_{k=1}^{N} \beta^{k-1} f(U_k) \to \mathbb{E}\left[\sum_{k=1}^{\infty} \beta^{k-1} f(U_k)\right],$$

for $U = U^*$. Hence

$$V(X) = \mathbb{E}\sum_{k=1}^{N} \beta^{k-1} f(U_k^*) + \beta^n \mathbb{E}V(X_N^{U^*}) \to \mathbb{E}\sum_{k=1}^{\infty} \beta^{k-1} f(U_k^*).$$

8.3   Optimal Stopping

Our goal is the following: given a Markov process (X_n), and a horizon $N > 0$, we wish to find a stopping time $T^* \leq N$ to maximize $\mathbb{E}[g(X_T)]$.

Suppose that $X_n = G(n, X_{n-1}, \xi_n)$, for (ξ_n) independent. Then

$$X_{n \wedge T} = \begin{cases} G(n, X_{(n-1) \wedge T}, \xi_N) & \text{if } T \geq n, \\ X_{(n-1) \wedge T} & \text{if } T \leq n-1. \end{cases}$$

So we have $X_{n \wedge T} = H(n, X_{(n-1) \wedge T}, U_n, \xi_n)$, where $U_n = \mathbb{1}_{\{T \leq n-1\}}$. Then,

$$H(n, x, u, \xi) = \begin{cases} G(n, x, \xi) & u = 0, \\ x & u = 1. \end{cases}$$

Theorem 8.3. *Let* $V(N, x) = g(x)$, *and* $V(n-1, x) = \max\{g(x), \mathbb{E}[V(n, G(n, x, \xi_n))]\}$, *for* $1 \leq n \leq N$. *Then*

$$V(n, x) = \max_{\substack{n \leq T \leq N \\ \text{stopping times}}} \mathbb{E}[g(X_T) \mid X_n = x].$$

9   Multiperiod Models

9.1   Multiperiod Arbitrage

Recall that the time n wealth is

$$X_n = (1 + r)X_{n-1} + \theta_n^T(S_n - (1 + r)S_{n-1}).$$

No consumption is equivalent to self-financing. Then

$$\frac{X_n}{(1 + r)^n} = X_0 + \sum_{k=1}^{n} \theta_k^T \left(\frac{S_k}{(1 + r)^k} - \frac{S_{k-1}}{(1 + r)^{k-1}} \right).$$

Definition 9.1. An N-period *arbitrage* is a previsible process (ϕ_n) such that

$$\sum_{n=1}^{n} \phi_n^T \left(\frac{S_n}{(1 + r)^n} - \frac{S_{n-1}}{(1 + r)^{n-1}} \right) \geq 0$$

almost surely, with strict inequality with positive probability.

Definition 9.2. A *risk-neutral measure* $\mathbb{Q}$ is equivalent to $\mathbb{P}$, and such that $(S_n/(1 + r)^n)$ is a $\mathbb{Q}$-martingale, i.e.

$$\frac{\mathbb{E}[S_n \mid \mathcal{F}_{n-1}]}{1 + r} = S_{n-1},$$

for all $n \geq 1$.

Remark. The risk-neutral measures are sometimes called equivalent martingales measures.

Theorem 9.1 (Fundamental Theorem of Asset Pricing)**.** *There is no N-period arbitrage if and only if there exists a risk-neutral measure for (S_n).*

The proof is by essentially reusing the one-period argument N times.

9.2   Binomial Model

Let $d = 1$, and $S_n = S_{n-1}\xi_n$, where $S_0 > 0$ and $\xi_n > 0$ for all n, with

$$0 < \mathbb{P}(\xi_n = 1 + b \mid \mathcal{F}_{n-1}) = 1 - \mathbb{P}(\xi_n = 1 + a \mid \mathcal{F}_{n-1}) < 1$$

almost surely, for all $n \geq 1$, where a, b are given constants. Hence the probability of following any path is positive.

Theorem 9.2. *There exists a risk-neutral measure $\mathbb{Q}$ if and only if $a < r < b$, in which case (ξ_n) are iid under $\mathbb{Q}$, with*

$$\mathbb{Q}(\xi_n = 1 + b) = \frac{r - a}{b - a},$$

for all n. In particular, $\mathbb{Q}$ is unique.

Proof: Suppose $\mathbb{Q}$ exists. Then

$$\frac{\mathbb{E}^{\mathbb{Q}}[S_n \mid \mathcal{F}_{n-1}]}{1 + r} = S_{n-1} = \frac{\mathbb{E}^{\mathbb{Q}}[S_{n-1}\xi_n \mid \mathcal{F}_{n-1}]}{1 + r}$$

$$= \frac{S_{n-1}}{1 + r}\Big(\mathbb{Q}(\xi_n = 1 + b \mid \mathcal{F}_{n-1})(1 + b)$$

$$+ \mathbb{Q}(\xi_n = 1 + a \mid \mathcal{F}_{n-1})(1 + a)\Big).$$

Solving this, we get that

$$\mathbb{Q}(\xi_n = 1 + b \mid \mathcal{F}_{n-1}) = \frac{r - a}{b - a}, \qquad \mathbb{Q}(\xi_n = 1 + a \mid \mathcal{F}_{n-1}) = \frac{b - r}{b - a}.$$

So $\mathbb{Q} \sim \mathbb{P}$ exists if and only if $a < r < b$, since both probabilities must be positive.

Since the conditional distribution of ξ_n given $\mathcal{F}_{n-1}$ is not random, ξ_n is independent of $\mathcal{F}_{n-1}$, and it is the same for all n. So (ξ_n) are iid under $\mathbb{Q}$.

Remark. (S_n) is a $\mathbb{Q}$-Markov process. Moreover, all that we need to show that (ξ_n) is adapted to $(\mathcal{F}_n)$, but we usually take $\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$.

10   Pricing and Hedging Contingent Claims

10.1   European Claims

Definition 10.1. A *European* contingent claim is specified by a non-random maturity date N, and $\mathcal{F}_N$-measurable payout Y.

> **Example 10.1.**
>
> A European *call* with maturity N and strike K has payout
> $$Y = (S_N - K)^+.$$
>
> Similarly, the European *put* has payout
> $$Y = (K - S_N)^+.$$
>
> There are other path-dependent claims: for example the look-back, with payout
> $$Y = S_N - \min_{0 \leq n \leq N} S_n.$$

Let's think about pricing these. To do this, consider the binomial model with $d = 1$, and $S_n = S_{n-1}\xi_n$, where ξ_n takes values in $\{1 + a, 1 + b\}$, for $a < r < b$.

There is a unique risk-neutral measure, so there is a unique no-arbitrage price of the European claim:

$$\pi_n = \frac{1}{(1+r)^{N-n}} \mathbb{E}^{\mathbb{Q}}[Y \mid \mathcal{F}_n] = f_n(S_0, \ldots, S_n),$$

assuming that $\mathcal{F}_n = \sigma(S_0, \ldots, S_n)$, and by measurability, for some $f_n : \mathbb{R}^{n+1} \to \mathbb{R}$.

Theorem 10.1. *Let*

$$\theta_n = \frac{f_n(S_0, \ldots, S_{n-1}, S_{n-1}(1+b)) - f_n(S_0, \ldots, S_{n-1}, S_{n-1}(1+a))}{S_{n-1}(b-a)}.$$

The wealth process starting at $X_0 = \pi_0$, employing trading strategy (θ_n) is such that $X_n = \pi_n$, for $n \leq N - 1$, and $X_N = Y$ almost surely.

> **Proof:**   For each n, there exists a unique $\mathcal{F}_{n-1}$-measurable solution (x_{n-1}, b_n) to
> $$(1+r)x_{n-1} + b_n(S_n - (1+r)S_{n-1}) = \pi_n,$$

since this is equivalent to:

$$(1+r)x_{n-1} + b_n S_{n-1}(b-r) = f_n(S_0, \ldots, S_{n-1}, S_{n-1}(1+b)),$$
$$(1+r)x_{n-1} + b_n S_{n-1}(a-r) = f_n(S_0, \ldots, S_{n-1}, S_{n-1}(1+a)).$$

The solution is $b_n = \theta_n$, and

$$(1+r)_{n-1} = \mathbb{E}^{\mathbb{Q}}[\pi_n \mid \mathcal{F}_{n-1}] = \mathbb{E}^{\mathbb{Q}}[\mathbb{E}^{\mathbb{Q}}[Y \mid \mathcal{F}_n] \mid \mathcal{F}_{n-1}]$$
$$= (1+r)\pi_{n-1},$$

by the tower property. So $x_{n-1} = \pi_{n-1}$.

So starting at $X_0 = \pi_0$, and using strategy (θ_n), this equation says that $X_n = \pi_n$ for all $n \le N$, and $X_N = Y$, by induction.

Remark. If $Y = g(S_N)$ for some function g, the claim is sometimes called a *vanilla claim*. Otherwise it is called *exotic*.

For vanilla claims, we have

$$\pi_n = \frac{\mathbb{E}^{\mathbb{Q}}[g(S_N) \mid \mathcal{F}_n]}{(1+r)^{N-n}} = \frac{\mathbb{E}^{\mathbb{Q}}[g(S_N) \mid S_n]}{(1+r)^{N-n}} = V(n, S_n)$$

for some function V, by the Markov property of (S_N) under $\mathbb{Q}$. In fact, we can say that

$$V(n,s)\frac{1}{(1+r)^{N-n}} \sum_{k=0}^{N-n} \binom{N-n}{k} q^k (1-q)^{N-n-k} g(s(1+b)^k(1+a)^{N-n-k}),$$

where $q = \mathbb{Q}(\xi_1 = 1+b) = (r-a)/(b-a)$. This satisfies

$$(1+r)V(n-1,s) = qV(n, s(1+b)) + (1-q)V(n, s(1+a)),$$

which is an analogue to the Black-Scholes PDE.

10.2   American Claim

Definition 10.2. These are specified by an adapted process $(Y_n)_{0 \le n \le N}$, where Y_n is the payout if the claim is exercised at time n.

The time to exercise is chosen by the holder of the claim.

Consider a binomial model, where $\mathbb{Q}$ is the risk-neutral measure. Let (Y_n) be the payout of the American claim; for example $Y_n = (S_n - K)^+$ for the American call.

Let

$$\pi_n = Y_N,$$

$$\pi_{n-1} = \max\left\{Y_{n-1}, \frac{\mathbb{E}^{\mathbb{Q}}[\pi_n \mid \mathcal{F}_{n-1}]}{1+r}\right\}.$$

Note $\pi_n \geq Y_n$ for all n. Moreover, the price process (π_n) is a supermartingale in general, since

$$\pi_n \geq \frac{\mathbb{E}[\pi_{n+1} \mid \mathcal{F}_n]}{1+r},$$

but moreover we know that $(\pi_n/(1+r)^n)$ is a martingale up to time

$$T^* = \inf\{n \mid \pi_n = Y_n\}.$$

To hedge an American claim, we hedge as if it is a European claim up to time T^*, i.e.

$$\theta_n = \frac{f_n(S_0, \dots, S_{n-1}, S_{n-1}(1+b)) - f_n(S_0, \dots, S_{n-1}, S_{n-1}(1+a))}{S_{n-1}(b-a)},$$

where $f_n(S_0, \dots, S_n) = \pi_n$.

11   Continuous Time Finance

For motivation take a binomial model with $d = 1$, and

$$S_n = S_0 \xi_1 \cdots \xi_n,$$

so that

$$\log S_k = \log S_0 + X_1 + \cdots + X_n,$$

where $X_n = \log \xi_n$. Assume that (X_n) are IID. To change to the continuous time limit, let $t = n\delta$, where δ is a small parameter, so

$$\hat{S}_t = S_n = S_{t/\delta},$$

and hence

$$\log \hat{S}_t = \log S_0 + \mu t + \sigma W_t,$$

where

$$\mu = \frac{\mathbb{E}[X]}{\delta}, \qquad \sigma = \frac{1}{\delta} \operatorname{Var}(X), \qquad W_t = \frac{X_1 + \cdots + X_{t/\delta} - t\mu}{\sigma}.$$

W has the following properties:

- $W_t - W_s$ is independent of $(W_u)_{u \leq s}$.

- $\mathbb{E}[W_t - W_s] = 0$, and $\operatorname{Var}(W_t - W_s) = t - s$.

- $W_t - W_s \sim N(0, t - s)$ as $\delta \to 0$, for fixed s, t, i.e. by taking $n = t/\delta \to \infty$, and $m = s/\delta \to \infty$.

11.1   Brownian Motion

Definition 11.1. A *Brownian motion* is a continuous-time process (W_t) such that:

- $W_0 = 0$.

- $W_t - W_s$ is independent of $(W_u)_{u \leq s}$, for all $0 \leq s \leq t$.

- $W_t - W_s \sim N(0, t - s)$ for all $0 \leq s \leq t$.

- $t \mapsto W_t$ is continuous, i.e. for all $\omega \in \Omega$, $t \mapsto W_t(\omega)$ is continuous.

Theorem 11.1 (Wiener)**.** *Brownian motion exists, i.e. there exists $(\Omega, \mathcal{F}, \mathbb{P})$ and measurable functions $W_t : \Omega \to \mathbb{R}$ satisfying these definitions.*

Proposition 11.1. *Brownian motion is a martingale (on its own filtration).*

Proof: W is adapted to its own filtration. It is integrable since normal random variables are integrable, and

$$\mathbb{E}[W_t - W_s \mid \mathcal{F}_s] = \mathbb{E}[W_t - W_s] = 0,$$

since $W_t - W_s = N(0, t - s)$, which has mean 0.

Proposition 11.2. *Brownian motion is a Markov process.*

Proof: Let g be bounded. Then,

$$
\begin{aligned}
\mathbb{E}[g(W_t) \mid \mathcal{F}_s] &= \mathbb{E}[g(W_s + W_t - W_s) \mid \mathcal{F}_s] \\
&= \mathbb{E}[g(x + W_t - W_s)]\Big|_{x = W_s} \\
&= \mathbb{E}[g(W_s + W_t - W_s) \mid W_s] \\
&= \mathbb{E}[g(W_t) \mid W_s].
\end{aligned}
$$

Definition 11.2. A *Gaussian process* (X_t) is such that $(X_{t_1}, \ldots, X_{t_n})$ is jointly normal for all $0 \le t_1 < t_2 < \cdots < t_n$.

Proposition 11.3. *Brownian motion is Gaussian.*

Proof: Fix $0 \le t_1 < \cdots < t_n$, and constants $a_1, \ldots, a_n$. Then,

$$\sum_{k=1}^{n} a_k W_{t_k} = \sum_{k=1}^{n} b_k (W_{t_k} - W_{t_{k-1}}),$$

where $t_0 = 0$, and $b_k = a_1 + \cdots + a_k$. Then the sum of independent normals is normal.

Proposition 11.4. (W_t) *is a Brownian motion if and only if it is a continuous Gaussian process with* $\mathbb{E}[W_t] = 0$ *for all* t, *and* $\mathbb{E}[W_s W_t] = s$ *for all* $0 \le s \le t$.

Proof: Suppose W is a Brownian motion. Then it s Gaussian and continuous, and $\mathbb{E}[W_t] = 0$, since $W_t \sim N(0, t)$. Moreover,

$$
\begin{aligned}
\mathbb{E}[W_s W_t] &= \mathbb{E}[W_s^2] - \mathbb{E}[W_s(W_t - W_s)] \\
&= s + \mathbb{E}[W_s]\mathbb{E}[W_t - W_s] = s.
\end{aligned}
$$

Now suppose W is continuous, Gaussian, $\mathbb{E}[W_t] = 0$ and $\mathbb{E}[W_s W_t] = s$, for

all $0 \leq s \leq t$.

Then $\operatorname{Var}(W_t) = \mathbb{E}[W_t^2] = t$, $\operatorname{Var}(W_t - W_t s = \operatorname{Var}(W_t) + \operatorname{Var}(W_s) - 2\operatorname{Cov}(W_t, W_s) = t + s - 2s = t - s$, hence $W_t - W_s \sim N(0, t - s)$.

Now for $u < s \leq t$, $\operatorname{Cov}(W_u, W_t - W_s) = \operatorname{Cov}(W_u, W_t) - \operatorname{Cov}(W_u, W_s) = u - u = 0$. Hence by Gaussianity, $W_t - W_s$ is independent of W_u, and as it is jointly Gaussian, it is independent of $(W_u)_{u \leq s}$.

Proposition 11.5. *Let W be a Brownian motion. Then consider:*

(i) $\hat{W}_t = cW_{t/c^2}$, *for any constant $c \neq 0$.*

(ii) $\hat{W}_t = W_{t+T} - W_T$, *for $T \geq 0$ a constant.*

(iii) $\hat{W}_t = tW_{1/t}$, $\hat{W}_0 = 0$.

Then these are all Brownian motions.

These follow from the previous proposition. The last one is as $\mathbb{E}[\hat{W}_t] = t\mathbb{E}[W_{1/t}] = 0$, and

$$\mathbb{E}[\hat{W}_s \hat{W}_t] = st\mathbb{E}[W_{1/s} W_{1/t}] = st/t = s.$$

Theorem 11.2. *Let (W_t) be a Brownian motion, and $T_a = \inf\{t \geq 0 \mid W_t = a\}$. Then $T_a < \infty$ almost surely.*

Proof: Consider the case $a > 0$. We must show that

$$\sup_{t \geq 0} W_t \geq a,$$

almost surely. For any $a, b0$ and $c > 0$,

$$\mathbb{P}(a < \sup_{t \geq 0} W_t < b) = \mathbb{P}(a < \sup_{t \geq 0} cW_{t/c^2} < b)$$
$$= \mathbb{P}(a/c < \sup_{s \geq 0} W_s \leq b/c) \to 0,$$

as $c \to \infty$, and for any $(a, b) \supseteq [0, \infty)$. Hence $Z = \sup_{t \geq 0} W_t \in \{0, \infty\}$ almost surely.

So it satisfies to show that $\sup W_t \neq 0$ almost surely. Let $\hat{Z} = \sup_{t \geq 1}(W_t - W_1) \overset{d}{=} Z$. So $\hat{Z} \in \{0, \infty\}$ almost surely.

Note that $\{\hat{Z} = \infty\}$ if and only if $\{Z = \infty\}$, as $\sup_{0 \leq t \leq 1} W_t$ is finite by the

continuity of Brownian motion. Hence,

$$\mathbb{P}(Z = 0) = \mathbb{P}(Z = 0, \hat{Z} = 0) \leq \mathbb{P}(W_1 \leq 0, \hat{Z} = 0)$$

$$= \mathbb{P}(W_1 \leq 0)\mathbb{P}(\hat{Z} = 0) = \frac{1}{2}\mathbb{P}(Z = 0).$$

Therefore $\mathbb{P}(Z = 0) = 0$, so $\mathbb{P}(Z = \infty) = 1$.

Brownian motion also has the strong Markov property.

Theorem 11.3. *Let (W_t) be a Brownian motion, and T a finite stopping time. The process $(W_{t+T} - W_T)$ is also a Brownian motion, independent of $(W_t)_{0 \leq t \leq T}$.*

The proof of this is omitted.

Theorem 11.4 (Reflection Principle). *Let (W_t) be a Brownian motion, and $a \in \mathbb{R}$. Let*

$$\tilde{W}_t = \begin{cases} W_t & \text{if } 0 \leq t < T_a, \\ 2a - W_t & \text{if } t \geq T_a. \end{cases}$$

Then $(\tilde{W}_t)$ is a Brownian motion.

Proof: Note that since T_a is a stopping time, $W_{s+T_a} - W_{T_a} = W_{s+T_a} - a$ is a Brownian motion independent of (W_t) for $t \leq T_a$. But now notice

$$2a - W_{s+T_a} = a + (a - W_{s+T_a}) \overset{d}{=} a + (W_{s+T_a} - a) = W_{s+T_a},$$

hence $\tilde{W}_t \overset{d}{=} W_t$.

Hence we get the following:

$$\mathbb{P}(\max_{0 \leq s \leq t} W_s \geq a, W_t \leq b) = \mathbb{P}(W_t \geq 2a - b),$$

for $a \geq 0, b \leq a$.

Proof: Note that, taking the same $\tilde{W}_t$ as before,

$$\mathbb{P}(\max_{0 \leq s \leq t} W_s \geq a, W_t \leq b) = \mathbb{P}(\tilde{W}_t \geq 2a - b) = \mathbb{P}(W_t \geq 2a - b).$$

Consider $M_t = \max_{0 \leq s \leq t} W_s$. Then the joint distribution of (M_t, W_t) is

$$F_{M_t, W_t}(a, b) = \mathbb{P}(M_t < a, W_t < b) = \mathbb{P}(W_t < b) - \mathbb{P}(M_t < a, W_t > b)$$

$$= \mathbb{P}(W_t \leq b) - \mathbb{P}(W_t \geq 2a - b) = \Phi\left(\frac{b}{\sqrt{t}}\right) - \left(1 - \Phi\left(\frac{2a - b}{\sqrt{t}}\right)\right)$$

$$= \Phi\left(\frac{b}{\sqrt{t}}\right) + \Phi\left(\frac{2a - b}{\sqrt{t}}\right) - 1.$$

Here $\Phi(x) = \mathbb{P}(Z \leq x)$, where $Z \sim N(0, 1)$. Now we see that

$$f_{M_t, W_t}(a, b) = \frac{\partial^2}{\partial a \partial b} F_{M_t, W_t}(a, b) = \frac{\partial}{\partial b}\left(\frac{\partial}{\partial a}\Phi\left(\frac{2a - b}{\sqrt{t}}\right)\right)$$

$$= \frac{\partial}{\partial b}\left(\frac{2}{\sqrt{t}}\phi\left(\frac{2a - b}{\sqrt{t}}\right)\right) = \frac{2}{\sqrt{t}}\left(-\frac{1}{\sqrt{t}}\right)\phi'\left(\frac{2a - b}{\sqrt{t}}\right)$$

$$= \frac{2(2a - b)}{t^{3/2}}\phi\left(\frac{2a - b}{\sqrt{t}}\right).$$

Proposition 11.6. *For suitable g and $a \geq 0$,*

$$\mathbb{E}[g(W_t)\mathbb{1}_{(M_t \geq a)}] = \mathbb{E}[(g(W_t) + g(2a - W_t))\mathbb{1}_{(W_t \geq a)}].$$

Proof: Note that

$$\mathbb{E}[g(W_t)\mathbb{1}_{(M_t \geq a)}] = \mathbb{E}[g(W_t)\mathbb{1}_{(M_t \geq a, W_t \geq a)}] + \mathbb{E}[g(W_t)\mathbb{1}_{(M_t \geq a, W_t \leq a)}]$$

$$= \mathbb{E}[g(W_t)\mathbb{1}_{(W_t \geq a)}] + \mathbb{E}[g(W_t)\mathbb{1}_{(\tilde{W}_t \geq a)}]$$

$$= \mathbb{E}[g(W_t)\mathbb{1}_{(W_t \geq a)}] + \mathbb{E}[g(2a - W_t)\mathbb{1}_{(W_t \geq a)}].$$

Hence, we see that

$$F_{M_t}(a) = \mathbb{P}(M_t \leq a) = \mathbb{P}(M_t \leq a, W_t \leq a) = 2\Phi\left(\frac{a}{\sqrt{t}}\right) - 1$$

$$= F_{|W_t|}(a),$$

so M_t has the same distribution as $|W_t|$. Hence, for any $a \geq 0$,

$$\mathbb{P}(T_a \leq t) = \mathbb{P}(M_t \geq a) = 1 - F_{M_t}(a)$$

$$= 2\left(1 - \Phi\left(\frac{a}{\sqrt{t}}\right)\right) = 2\Phi\left(\frac{-a}{\sqrt{t}}\right).$$

Hence T_a has density

$$f_{T_a}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(2\Phi\left(-\frac{\alpha}{\sqrt{t}}\right)\right) = \frac{2}{t^{3/2}}\phi\left(-\frac{\alpha}{\sqrt{t}}\right) = \frac{a}{t^{3/2}}\phi\left(\frac{a}{\sqrt{t}}\right).$$

11.2   Cameron-Martin Theorem

To motivate this section, let $X \sim N(0, 1)$. Then,

$$\mathbb{E}[e^{aX - a^2/2} f(X)] = \mathbb{E}[f(X + a)],$$

for $a \in \mathbb{R}$. This was proven in example sheet 1. Now, let $X \sim N_d(0, I_d)$, so multivariate. Then,

$$\mathbb{E}[e^{a^T x - \|a\|^2/2} f(X)] = \mathbb{E}[f(X + a)],$$

for all $a \in \mathbb{R}^d$. We extend this theorem to Brownian motions.

Theorem 11.5 (Cameron-Martin). *Let (W_t) be a Brownian motion, and $f : C[0, T] \to \mathbb{R}$ be suitable. Then,*

$$\mathbb{E}[e^{cW_T - c^2 T/2} f((W_t)_{0 \le t \le T})] = \mathbb{E}[f((W_t + ct)_{0 \le t \le T})].$$

Proof: By monotone class theorem, it is enough to consider functions of the form
$$f(\omega) = g(\omega(t_1), \omega(t_2), \ldots, \omega(t_d)),$$
for some function $g : \mathbb{R}^d \to \mathbb{R}$, since pointwise evaluations generate the σ-algebra on $C[0, T]$. Then,

$$\mathbb{E}[e^{cW_t - c^2 T/2} f((W_t)_{0 \le t \le T})] = \mathbb{E}[e^{cW_T - c^2 T/2} g(W_{t_1}, \ldots, W_{t_d})]$$
$$= \mathbb{E}\left[e^{\sum c\sqrt{t_k - t_{k-1}} Z_k - \sum c^2(t_k - t_{k-1})/2} g\left(\left(\sum \sqrt{t_k - t_{k-1}} Z_k \right) \right) \right].$$

Now let $a_n = c\sqrt{t_n - t_{n-1}}$, for $1 \le n \le d$, and so this equals

$$\mathbb{E}\left[g\left(\sum_{k=1}^{n} \sqrt{t_k - t_{k-1}}(Z_k + c\sqrt{t_k - t_{k-1}}) \right) \right] = \mathbb{E}[g((W t_n + c t_n)_{1 \le n \le d})].$$

> **Example 11.1.**
>
> Consider $\mathbb{P}(\max_{0 \leq t \leq T}(W_t + ct) \leq a)$, a probability of a Brownian motion with drift. Then,
>
> $$\mathbb{P}(\max_{0 \leq t \leq T}(W_t + ct) \leq a) = \mathbb{E}[\mathbb{1}_{(\max_{0 \leq t \leq T}(W_t+ct) \leq a)}]$$
>
> $$= \mathbb{E}[e^{cW_t - c^2 T/2}\mathbb{1}_{(\max_{0 \leq t \leq T} W_t \leq a)}]$$
>
> $$= \mathbb{E}[e^{cW_t - c^2 T/2}\mathbb{1}_{(W_t \leq a)}] - \mathbb{E}[e^{c(2a-W_t) - c^2 T/2}\mathbb{1}_{(W_t \geq a)}]$$
>
> $$= \mathbb{P}(W_T + cT \leq a) - e^{2ca}\mathbb{P}(W_T - cT \geq a)$$
>
> $$= \Phi\left(\frac{a - cT}{\sqrt{T}}\right) - e^{2ca}\Phi\left(\frac{-a - cT}{\sqrt{T}}\right).$$

Here is a reformulation of Cameron-Martin that is useful in the financial setting:

Let (W_t) be a Brownian motion under $\mathbb{P}$, and let

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{cW_t T - c^2 T/2}.$$

Let $\hat{W}_t = W_t - ct$. Then $(\hat{W}_t)$ is a Brownian motion under $\mathbb{Q}$.

> **Proof:**   Let $f : C[0,T] \to \mathbb{R}$ be suitable. Then,
>
> $$\mathbb{E}^{\mathbb{Q}}[f((\hat{W}_t)_{0 \leq t \leq T})] = \mathbb{E}^{\mathbb{P}}[e^{cW_T - c^2 T/2}f((W_t - ct)_{0 \leq t \leq T})]$$
>
> $$= \mathbb{E}^{\mathbb{P}}[f((W_t)_{0 \leq t \leq T})].$$
>
> Hence the law of $\hat{W}$ under $\mathbb{Q}$ is the same as of W under $\mathbb{P}$.

11.3   Brownian Motion and the Heat Equation

Our motivation is as follows. Let (X_n) be a simple, symmetric random walk. Let $u(n, x) = \mathbb{E}[g(X_n + x)]$, for $x \in \mathbb{Z}$. Then,

$$u(n + 1, x) = \frac{1}{2}u(n, x + 1) + \frac{1}{2}u(n, x - 1),$$

from the Markov property of a random walk.

Proposition 11.7. *Let (W_t) be a Brownian motion. Let*

$$u(t, x) = \mathbb{E}[g(W_t + x)] =' int\frac{g(z + x)e^{-z^2/2t}}{\sqrt{2\pi t}}\, dz.$$

Then,

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}.$$

Proof: We can prove this by just differentiating the integral formula. Indeed, this is just

$$\int g(\sqrt{t}z + x)\frac{e^{-z^2/2}}{\sqrt{2\pi}}\,\mathrm{d}z.$$

12   Black-Scholes Model

We are looking at extending the binomial model to continuous time.

To do this, we need to look at interest rate. In the multi-period model, at time n the interest is $(1+r)^n$. If $r = \hat{r}\delta$, for δ the length of the time period, and $n = t/\delta$, this converges to $e^{\hat{r}t}$. We now drop the hat, and view r as the interest per time.

Recall that in the binomial model $S_n = S_0\xi_1 \cdots \xi_n = S_0 e^{\mu t + \sigma W_t} = \hat{S}_t$. In the Black-Scholes model, we drop the hats, so

$$S_t = S_0 e^{\mu t + \sigma W_t},$$

where W is a Brownian motion, and μ, σ are constants with $\sigma > 0$.

Definition 12.1. A *risk-neutral measure* $\mathbb{Q}$ is equivalent to $\mathbb{P}$, and such that $(e^{rt}S_t)_{t \geq 0}$ is a $\mathbb{Q}$-martingale.

We will use the fact that $M_t = e^{cW_t - c^2 t/2}$ is a martingale for any constant c, proven in example sheets.

In the Black-Scholes model,

$$e^{-rt}S_t = S_0 e^{(\mu - r)t + \sigma W_t} = S_0 e^{(\mu - r)t + \sigma(\hat{W}_t + ct)}.$$

Recall Cameron-Martin:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{cW_T - c^2 T/2},$$

then $\hat{W}_t = W_t - ct$ is a $\mathbb{Q}$-Brownian motion. To apply Cameron-Martin, we need $\mu - r + \sigma c = -\sigma^2/2$. Hence, we need

$$c = \frac{r - \mu}{\sigma} - \frac{\sigma}{2}.$$

Then,

$$e^{-rt}S_t = S_0 e^{-\sigma^2 t/2 + \sigma \hat{W}_t},$$

so is a $\mathbb{Q}$-martingale. In fact, it is the unique risk-neutral measure.

12.1   Black-Scholes Pricing

Definition 12.2. The *Black-Scholes price* of a European claim with payout Y is

$$\pi_t = e^{-(T-t)r} \mathbb{E}^{\mathbb{Q}}[Y \mid \mathcal{F}_t],$$

where T is the maturity date of the claim.

Remark. $e^{-rt}\pi_t = \mathbb{E}^{\mathbb{Q}}[e^{-rT}Y \mid \mathcal{F}_t]$ is a $\mathbb{Q}$-martingale.

> **Example 12.1.**
>
> The Black-Scholes formula - this is the price of a European call:
>
> $$\begin{aligned}
> \pi_t &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(S_T - K)^+ \mid \mathcal{F}_t] \\
> &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(S_t e^{(r-\sigma^2/2)(T-T)+\sigma(\hat{W}_t - \hat{W}_t)} - K)^+ \mid \mathcal{F}_t] \\
> &= S_t\mathbb{E}^{\mathbb{Q}}\left[\left(e^{-\sigma^2/2(T-t)+\sigma(\hat{W}_T - \hat{W}_t)} - \frac{K}{S_t}e^{-r(T-t)}\right)^+ \mid \mathcal{F}_t\right].
> \end{aligned}$$
>
> Now define
> $$F(v,m) = \mathbb{E}\left[(e^{-v/2+\sqrt{v}Z} - M)^+\right],$$
>
> where $Z \sim N(0,1)$. Then the Black-Scholes formula says
>
> $$\pi_t = S_t F\left(\sigma^2(T-t), \frac{Ke^{-r(T-t)}}{S_t}\right).$$
>
> Now we simplify F. Note that
>
> $$\begin{aligned}
> F &= \mathbb{E}\left[(e^{-v/2+\sqrt{v}Z} - m)\mathbb{1}_{(Z \geq \log m/\sqrt{v}+\sqrt{v}/2)}\right] \\
> &= \mathbb{E}\left[e^{-v/2+\sqrt{v}Z}\mathbb{1}_{Z \geq \log m/\sqrt{v}+\sqrt{v}/2}\right] - m\mathbb{P}\left(Z > \frac{\log m}{\sqrt{v}} + \frac{\sqrt{v}}{2}\right) \\
> &= \mathbb{P}\left(Z + \sqrt{v} \geq \frac{\log m}{\sqrt{v}} + \frac{\sqrt{v}}{2}\right) - mP\left(Z > \frac{\log m}{\sqrt{v}} + \frac{\sqrt{v}}{2}\right).
> \end{aligned}$$
>
> Let Φ be the CDF of the normal distribution. Note that $\mathbb{P}(Z > x) = \Phi(-x)$, by symmetry. Hence the Black-Scholes formula gives
>
> $$\pi_t = S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2),$$
>
> where
>
> $$d_1 = \frac{-\log(Ke^{-r(T-t)}/S_t)}{\sigma\sqrt{T-t}} + \frac{\sigma\sqrt{T-t}}{2},$$
> $$d_2 = \frac{-\log(Ke^{-r(T-t)}/S_t)}{\sigma\sqrt{T-t}} - \frac{\sigma\sqrt{T-t}}{2}.$$
>
> Here, everything is directly observable, apart from σ, the volatility.

For the European put, the payout is $(K - S_t)^+$. We could do this above derivation again, but instead we will use the following:

Put-call parity: Let P_t be the price of a put at strike K, and expiry T, and C_t the price of the call at the same strike K, and expiry T. Then,

$$P_t - C_t = Ke^{-r(T-t)} - S_t.$$

To prove this, note that

$$P_T - C_T = (K - S_T)^+ - (S_T - K)^+ = K - S_T.$$

Hence,

$$\begin{aligned} P_t - C_t &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[P_T - S_T \mid \mathcal{F}_t] \\ &= Ke^{-r(T-t)} - e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[S_T \mid \mathcal{F}_t] \\ &= Ke^{-r(T-t)} - S_t. \end{aligned}$$

Hence the Black-Scholes put price is

$$\begin{aligned} S_t\Phi(d_1) - Ke^{-r(T_t)}\Phi(d_2) &+ Ke^{-r(T-t)} - S_t \\ &= Ke^{-r(T-t)}\Phi(-d_2) - S_t\Phi(-d_1). \end{aligned}$$

The general formula for the price of a vanilla claim with payout $Y = g(S_T)$ is

$$\pi_t = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[g(S_T) \mid \mathcal{F}_t] = V(t, S_t),$$

where

$$V(t, s) = e^{-r(T-t)}\mathbb{E}[g(se^{(r-\sigma^2/2)(T-t)+\sigma\sqrt{T-t}Z}],$$

where $Z \sim N(0, 1)$.

12.2   Black-Scholes PDE

Note that this function V satisfies the following:

$$V(T, s) = g(s),$$

$$\frac{\partial V}{\partial t} + rs\frac{\partial V}{\partial s} + \frac{1}{2}S^2\sigma^2\frac{\partial^2 V}{\partial S^2} = rV,$$

for all $t < T$. We will derive this. Think of a binomial model approximation, where

$$(1 + r\delta)V(t - \delta, s) = qV(t, s(1 + b)) + (1 - q)V(t, s(1 + a)).$$

Taking the Taylor series (δ small) of the right side, we get this is

$$qV(t, s) + q\frac{\partial V}{\partial s}sb + \frac{1}{2}q\frac{\partial^2 V}{\partial s^2}s^2b^2 + (1 - q)V + (1 - q)\frac{\partial V}{\partial s}sa + (1 - q)\frac{\partial^2 V}{\partial s^2}s^2a^2,$$

plus higher order terms. For the left hand side, the Taylor expansion is

$$V(t, s) + \left(rV - \frac{\partial V}{\partial t}\right)\delta,$$

plus higher order terms. Now notice that $qb + (1 - q)a = r\delta$, and we can call $\delta\sigma^2 = qb^2 + (1 - q)a^2$ to rearrange.

Another derivation is as follows. Let

$$u(\tau, x) = \mathbb{E}[f(x + \sqrt{\tau}Z)],$$

where $Z \sim N(0, 1)$. Then we can prove that

$$\frac{\partial u}{\partial \tau} = \frac{1}{2}\frac{\tau^2 u}{\partial x^2}.$$

Proof: Note that

$$u(\tau, x) = \int f(y)p(\tau, x, y)\, \mathrm{d}y,$$

where

$$p(\tau, x, y) = \frac{1}{\sqrt{2\pi\tau}}e^{-(x-y)^2/2\tau},$$

and we can then show that

$$\frac{\partial p}{\partial \tau} = \frac{1}{2}\frac{\partial^2 p}{\partial x^2},$$

by algebra.

12.3   Hedging Contingent Claims and Greeks

In the binomial model, to hedge a vanilla claim, you must hold

$$\frac{V(n, S_{n-1}(1 + b)) - V(n, S_{n-1}(1 + a))}{S_{n-1}(b - a)}$$

shares at time n. In Black-Scholes, the answer is to hold

$$\frac{\partial V}{\partial S}(t, S_t).$$

Definition 12.3. The *Delta* is $\partial V/\partial S$; this is the sensitivity of the claim price with respect to the underlying asset price.

The *Gamma* is $\partial^2 V/\partial S^2$.

The *Theta* is $\partial V/\partial t$.

The Black-Scholes PDE says that

$$\text{Theta} + rS \times \text{Delta} + \frac{1}{2}S^2\sigma^2 \times \text{Gamma} = r \times \text{Price}.$$

Proposition 12.1. *Suppose g is increasing. Then the delta is non-negative. If g is convex, then gamma is non-negative.*

> **Proof:**
> $$V(t,s) = e^{-r(T-t)}\mathbb{E}[g(se^{(r-\sigma^2/2)(T-t)+\sqrt{T-t}\sigma Z})].$$
>
> We now use the technique of example sheet 1, so if g is increasing the $V(t,\cdot)$ is increasing.

12.4   Barrier Claims

To a vanilla claim, we can modify it by saying it is:

- down and in,
- down and out,
- up and in,
- up and out.

> **Example 12.2.**
>
> For example and up-and-in claim of the payout $g(S_T)$ and barrier B has pay
>
> $$g(S_T)\mathbb{1}_{(\max_{0\leq t\leq T} S_T\geq B)}.$$
>
> An up-and-out claim of the payout has payout
>
> $$g(S_T)\mathbb{1}_{(\max_{0\leq t\leq T} S_t< B)}.$$

Proposition 12.2. *The initial price of an up-and-out claim with payout*

$$g(S_T)\mathbb{1}_{(\max_{0\leq t\leq T} S_t< B)}$$

is the same as the initial price of a vanilla claim with payout

$$g(S_T)\mathbb{1}_{(S_T<B)} - \left(\frac{B}{S_0}\right)^{2r/\sigma^2-1} g\left(\frac{B^2 S_T}{S_0^2}\right) \mathbb{1}_{(S_T \le S_0^2/B)}.$$

Proof: Note that

$$\mathbb{1}_{(\max S_t < B)} = \mathbb{1}_{(S_T < B)} - \mathbb{1}_{(\max S_T \ge B, S_T < B)},$$

and the first term is accounted for. The second term has price

$$\mathbb{E}[g(S_T)\mathbb{1}_{(\max S_t \ge B, S_T < B)}] = \mathbb{E}[g(S_0 e^{\sigma(W_T+cT)})\mathbb{1}_{(\max(W_t+ct)\ge b, W_T+cT \le b)}],$$

where $c = (r - \sigma^2/2)/\sigma$, W is a Brownian motion under $\mathbb{Q}$, and $b = \log(B/S_0)/\sigma$. Then by Cameron-Martin, this is

$$\mathbb{E}[e^{cW_T-c^2T/2}g(S_0 e^{\sigma W_T})\mathbb{1}_{(\max_{0\le t\le T} W_t \ge b, W_T \le b)}]$$
$$= \mathbb{E}[e^{c(2b-W_T)-c^2T/2}g(S_0 e^{\sigma(2b-W_T)}\mathbb{1}_{(W_T \ge b)}]$$
$$= \mathbb{E}[e^{c(2b+W_T)-c^2T/2}g(S_0 e^{\sigma(2b+W_T)})\mathbb{1}_{(W_T \le -b)}]$$
$$= e^{2cb}\mathbb{E}[g(S_0 e^{2b\sigma}e^{\sigma(W_T+cT)})\mathbb{1}_{(W_T+cT \le -b)}].$$

Now note that

$$e^{2cb} = \left(\frac{B}{S_0}\right)^{2r/\sigma^2-1},$$

and remember that $S_T = S_0 e^{\sigma(W_T+cT)}$.

Index