Part II — Stochastic Financial Models

Based on lectures by J. R. Norris

Notes taken by Joseph Tedds using Dexter Chua's header and Gilles Castel's snippets.

Michaelmas 2019

These notes are not endorsed by the lecturers, and I have modified them (often significantly) after lectures. They are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.

Utility and mean-variance analysis
Utility functions; risk aversion and risk neutrality. Portfolio selection with the mean-variance criterion; the efficient frontier when all assets are risky and when there is one riskless asset. The capital-asset pricing model. Reservation bid and ask prices, marginal utility pricing. Simplest ideas of equilibrium and market cleaning. State-price density. [5]

Martingales
Conditional expectation, definition and basic properties. Conditional expectation, definition and basic properties. Stopping times. Martingales, supermartingales, submartingales. Use of the optional sampling theorem. [3]

Dynamic Models
Introduction to dynamic programming; optimal stopping and exercising American puts; optimal portfolio selection. [3]

Pricing contingent conditions
Lack of arbitrage in one-period models; hedging portfolios; martingale probabilities and pricing claims in the binomial model. Extension to the multi-period binomial model. Axiomatic derivation. [4]

Brownian motion
Introduction to Brownian motion; Brownian motion as a limit of random walks. Hitting-time distributions; changes of probability. [3]

Black-Scholes model
The BlackScholes formula for the price of a European call; sensitivity of price with respect to the parameters; implied volatility; pricing other claims. Binomial approximation to BlackScholes. Use of finite-difference schemes to compute prices [6]

Contents

0 Introduction

1 Utility and mean-variance analysis

1.1 Contingency claims and utility functions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A random variable X on Ω, provides a model for an investment which delivers $X(\omega)$ for consumption depending on chance $\omega \in \Omega$.

Definition (Contingent claim)**.** In this context we often use the term *contingent claim* as another name for a random variable.

Definition (Utility function)**.** By a *utility function* we mean any non-decreasing function $U : \mathbb{R} \mapsto [-\infty, \infty)$. Think of $U(x)$ as quantifying the satisfaction obtained on consuming x. Allowing $-\infty$ is a way of saying the value of x that obtains $-\infty$ is unacceptable.

We often assume the investor will act to maximise expected utility. So Y is *preferred* to X iff $\mathbb{E}[U(X)] \leq \mathbb{E}[U(Y)]$. If $\mathbb{E}[U(X)] = \mathbb{E}[U(Y)]$ the investor is said to be *indifferent* between X and Y. We say the investor is *risk averse* if they prefer $\mathbb{E}[X]$ to X for all integrable random variables X. We say *risk neutral* if indifferent between X and $\mathbb{E}[X]$.

Definition. Recall that U is a *concave* function if for all $x, y \in \mathbb{R}$, all $p \in (0, 1)$

$$pU(x) + (1-p)U(y) \leq U(px + (1-p)y).$$

1

Proposition. An investor with utility function U is risk averse if and only if U is concave.

Proof. Suppose risk averse. Consider the contingent claim X taking values x, y with probabilities $p, (1-p)$ respectively. Then,

$$pU(x) + (1-p)U(y) = \mathbb{E}[U(X)] \leq U(\mathbb{E}[X]) = U(px + (1-p)y).$$

Hence U is concave.
Suppose on the other hand U is concave. Let X be an integrable random variable (i.e. $\mathbb{E}[|X|] < \infty$) then by Jensen's inequality

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

Hence, the investor is risk averse. □

2

Example. For $\gamma \in (0, \infty)$ the CARA (constant absolute relative aversion) utility function of parameter γ is given by

$$\mathrm{CARA}_\gamma(x) = -e^{-\gamma x}.$$

3 For $R \in (0,1) \cup (1, \infty)$ the CRRA (constant relative risk aversion) utility function of parameter R is given by

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

Also,

$$\mathrm{CRRA}_1(x) = \begin{cases} \log x & x > 0 \\ -\infty & \text{otherwise} \end{cases}.$$

4

Non-rigorous discussion Let U be concave (note that U is non-decreasing). Consider a small continent claim X. We ask whether we prefer $w + X$ to w for a given constant w. By Taylor's theorem

$$U(w + X) \approx U(w) + X \underbrace{U'(w)}_{>0} + \frac{1}{2} X^2 \underbrace{U''(w)}_{<0}.$$

$$\mathbb{E}\left[U(w + X)\right] \approx U(w) + \mathbb{E}\left[X\right] U'(w) + \frac{1}{2}\mathbb{E}\left[X^2\right] U''(w),$$

so we prefer $w + X$ if

$$2\frac{\mathbb{E}\left[X\right]}{\mathbb{E}\left[X^2\right]} > -\frac{U''(w)}{U'(w)}$$

the Arrow-Pratt coefficient of absolute risk aversion. For CARA_γ this constant is equal to γ.
Similarly, do we prefer $w(1 + X)$ to w? Yes if

$$2\frac{\mathbb{E}\left[X\right]}{\mathbb{E}\left[X^2\right]} > -\frac{wU''(w)}{U'(w)}$$

the Arrow-Pratt coefficient of relative risk aversion. For CRRA_R this constant is equal to R.

1.2 Reservation prices and marginal prices

Consider an investor with concave utility function. Suppose they have available a set $\mathcal{A}$ of contingent claims, and suppose $\mathbb{E}\left[U(X)\right]$ is maximised over $\mathcal{A}$ at $X^* \in \mathcal{A}$. Let Y be another contingent claim. The investor would buy Y for price π if there exists $X \in \mathcal{A}$ such that

$$\mathbb{E}\left[U(X + Y - \pi)\right] > \mathbb{E}\left[U(X^*)\right].$$

The supremum of all such prices $\pi_b(Y)$ is the *(reservation) bid price)* for Y.

The investor would sell Y for price π if there exists $X \in \mathcal{A}$ such that

$$\mathbb{E}\left[U(X - Y + \pi)\right] > \mathbb{E}\left[U(X^*)\right].$$

The infimum of all such prices $\pi_a(Y)$ is the *(reservation) ask price* for Y.

Proposition. (Ask above, bid below) Assume $\mathcal{A}$ is convex. Then $\pi_b(Y) \leq \pi_a(Y)$

Proof. It suffices to show there is no price π at which the investor will both buy and sell. Suppose for a contradiction that there exist X_a, X_b such that

$$\mathbb{E}\left[U(X_a - Y + \pi)\right] > \mathbb{E}\left[U(X^*)\right].$$

$$\mathbb{E}\left[U(X_b + Y - \pi)\right] > \mathbb{E}\left[U(X^*)\right].$$

Now $X = \frac{X_a + X_b}{2} \in \mathcal{A}$ since $\mathcal{A}$ is convex and $U(X) \geq \frac{U(X_a - Y + \pi) + U(X_b + Y - \pi)}{2}$ since U is concave. Then we obtain the following contradiction.

$$\mathbb{E}\left[U(X^*)\right] < \frac{\mathbb{E}\left[U(X_a - Y + \pi)\right] + \mathbb{E}\left[U(X_b + Y - \pi)\right]}{2} \leq \mathbb{E}\left[U(X)\right] \leq \mathbb{E}\left[U(X^*)\right].$$

Hence there is no such π. □

Recall U is concave and non-decreasing. An investor has available a set of of contingent claims $\mathcal{A}$, and seeks to maximise $\mathbb{E}\left[U(X)\right], X \in \mathcal{A}$. Assume $X^* \in \mathcal{A}$ is a maximiser. Suppose Y is another contingent claim. Assume that $\mathcal{A}$ is an affine space and that U is a differentiable and strictly concave.

Definition (Affine space)**.** S is *affine* if $S - S$ is a vector space. This can be visualised as a vector space away from the origin.

Then X^* is unique (or $\frac{X_1^* + X_2^{ast}}{2}$ is better.

Definition (Marginal price)**.** We define the *marginal price* of Y as

$$\pi_m(Y) = \mathbb{E}\left[U'(X^*)Y\right] / \mathbb{E}\left[U'(X^*)\right].$$

<u>Non-rigorous discussion to explain</u> Note that for $\Xi \in \mathcal{A} - \mathcal{A}$ the map $t \mapsto \mathbb{E}\left[U(X^* + t\Xi\right]$ on $\mathbb{R}$ achieves its minimum at $t = 0$. So

$$0 = \frac{\mathrm{d}}{\mathrm{dt}}\bigg|_0 \mathbb{E}\left[U(X^* + t\Xi)\right] = \mathbb{E}\left[U'(X^*)\right].$$

It is plausible that there is a differentiable map $t \mapsto X^*(t) : \mathbb{R} \leftarrow \mathcal{A}$ such that for all t

$$\mathbb{E}\left[U(X^*(t) - tY + \pi_b(tY))\right] = \mathbb{E}\left[U(X^*)\right].$$

Then $X^*(0) = X^*$. Define $\Xi \in \frac{\mathrm{d}}{\mathrm{dt}}\Big|_0 X^*(t), \pi = \frac{\mathrm{d}}{\mathrm{dt}}\Big|_0 \pi_b(tY)$. It is plausible that $\Xi \in \mathcal{A} - \mathcal{A}$. So

$$0 = \frac{\mathrm{d}}{\mathrm{dt}}\bigg|_0 \mathbb{E}\left[U(X^*(t) - tY + \pi_b(tY))\right] = \mathbb{E}\left[U'(X^*)(\Xi - Y + \pi)\right].$$

So we see

$$\pi_m(Y) = \frac{\mathrm{d}}{\mathrm{dt}}\bigg|_0 \pi_b(tY) = \frac{\mathrm{d}}{\mathrm{dt}}\bigg|_0 \pi_a(tY).$$

So marginal price is the price to buy (or sell) a small amount of Y.

1.3 Single period asset price model

Definition (Single period asset price model)**.** By a *single period asset price model*, we mean a pair of random variables (S_0, S_1) in $\mathbb{R}^d$. We write $S_n = (S_n^1, \dots, S_n^d$ with S_n^i the price of asset i at time n.

Definition (Numeraire)**.** By a *numeraire* we mean a pair of random variables $(S_0^0, S_1^0$ in $(0, \infty)$.

Notation. We write

$$\overline{S}_n = (S_n^0, S_n) = (S_n^0, S_n^1, \dots, S_n^{d)}.$$

Call $(\overline{S}_0, \overline{S}_1)$ an *asset price model with numeraire*

Often we take $S_0^0 = 1, S_1^0 = 1 + r$ some constant $r \in (-1, \infty)$, Then S^0 is called a *riskless bond* and r is the *interest rate*. We assume $\overline{S}_0$ is non-random as the default.

In the case without numeraire, an investor with initial wealth w_0 chooses $\theta \in \mathbb{R}^d$ subject to

$$\theta.S_0 = \sum_{i=1}^{d} \theta^i S_0^i = w_0.$$

Then the investor has wealth $\theta.S_1$ at time 1. We call θ the *portfolio*. With numeraire, investor chooses $\overline{\theta} = (\theta^0, \theta$ such that $\overline{\theta}.\overline{S}_0 = w_0$. The wealth at time 1 is $\overline{\theta}.\overline{S}_0$.

It may be that there exists a random variable $\rho \geq 0$ such that $\mathbb{E}\left[\rho S_1^i\right] = S_0^i$ for all i. Then we call ρ a *state price density*

1.4 Portfolio selection using the mean-variance criterion

Let (S_0, S_1) be an asset price model on $\mathbb{R}^d$ with S_0 non-random, S_1 has mean μ, variance V. We assume that V is invertible and S_0, μ are linearly independent. Suppose we are given w_0, w_1. The investor wishes to

$$\begin{aligned} \text{minimise} \quad & \text{var}(\theta.S_1) \\ \text{subject to} \quad & \theta.S_0 = w_0, \\ & \mathbb{E}\left[\theta.S_1\right] = w_1 \end{aligned}$$

Note $\mathbb{E}\left[\theta.S_1\right] = \theta.\mu$, $\text{var}(\theta.S_1) = \theta.(V\theta)$ So our problem is to

$$\begin{aligned} \text{minimise} \quad & \theta.(V\theta) \\ \text{subject to} \quad & \theta.S_0 = w_0, \\ & \theta.\mu = w_1. \end{aligned}$$

Consider $L(\theta, \lambda) = \frac{1}{2}\theta.(V\theta) - \lambda_0 \theta.S_0 - \lambda_1 \theta.\mu$ At minimising θ^*.

$$\begin{aligned} 0 &= \frac{\partial}{\partial \theta^i} L(\theta, \lambda) \\ &= (V\theta)^i - \lambda_0 S_0^i - \lambda_1 \mu^i. \end{aligned}$$

So $\theta^* = \lambda_0 A S_0 + \lambda_1 A \mu, A = V^{-1}$. Now fit the constants

$$\begin{aligned} w_0 &= \theta^*.S_0 = \lambda_0 a + \lambda_1 b \\ w_1 &= \theta^*.\mu = \lambda_0 b + \lambda_1 c \end{aligned}$$

$a = S_0.(AS_0), b = \mu(AS_0) = S_0(A\mu), c = \mu(A\mu)$. Note that $\Delta = ac - b^2 \neq 0$ by linear independence

$$\begin{pmatrix} w_0 \\ w_1 \end{pmatrix} = M \begin{pmatrix} \lambda_0 \\ \lambda_1 \end{pmatrix}, M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

$$M^{-1} = \frac{1}{\Delta} \begin{pmatrix} d & -b \\ -b & a \end{pmatrix}.$$

$$\begin{pmatrix} \lambda_0 \\ \lambda_1 \end{pmatrix} = M^{-1} \begin{pmatrix} w_0 \\ w_1 \end{pmatrix}.$$

So

$$\theta^* = \frac{cw_0 - bw_1}{\Delta} AS_0 + \frac{aw_1 - bw_0}{\Delta} A\mu$$

The minimising variance is

$$\begin{aligned}
\theta^*_{\cdot}(V\theta^*) &= (\lambda_0 AS_1 + \lambda_1 A\mu)_{\cdot}(\lambda_0 S_0 + \lambda_1 \mu) \\
&= (\lambda_0 \lambda_1) M \begin{pmatrix} \lambda_0 \\ \lambda_1 \end{pmatrix} \\
&= (w_0 w_1) M^{-1} \begin{pmatrix} w_0 \\ w_1 \end{pmatrix} \\
&= \frac{cw_0^2 - 2bw_0 w_1 + aw_1^2}{\Delta} = q(w_1)
\end{aligned}$$

We minimise this over w_1

$$w_1^* = \frac{b}{a} w_0, \theta^*_{\text{min}} = \frac{w_0}{a} AS_0.$$

Putting w_1^* back into q, we obtain

$$q(w_1^*) = \frac{acw_0^2 - 2b^2 w_0^2}{a\Delta} + \frac{b^2}{a\Delta} w_0^2 = \frac{w_0^2}{a}$$

Suppose we seek to

$$\begin{aligned}
\text{minimise} \quad & \text{var}(\theta.S_1) \\
\text{subject to} \quad & \theta.S_0 = w_0,
\end{aligned}$$

Consider $L(\theta, \lambda) = \frac{1}{2}\theta.(V\theta) - \lambda\theta.S_0$. At minimiser,

$$0 = \frac{\partial}{\partial \theta^i} L(\theta, \lambda) = (V\theta)^i - \lambda S_0^i.$$

So

$$V\theta^* = \lambda S_0, \quad \theta^* = \lambda AS_0, A = V^{-1}.$$

Use the constraint to find $\lambda : w_0 = \theta^*_{\cdot} S_0 = \lambda \underbrace{a}_{S_0\cdot}(AS_0)$. Hence $\theta^* = \frac{w_0}{a} AS_0 = \theta^*_{\text{min}}$.

Add a riskless bond / bank account.

$$S^0 = 1, S_1^0 = 1 + r > 0.$$

Suppose we seek to

$$\begin{aligned}
\text{minimise} \quad & \text{var}(\overline{\theta}.\overline{S}_1) \\
\text{subject to} \quad & \overline{\theta}.\overline{S}_0 = w0 \\
& \mathbb{E}\left[\overline{\theta}.\overline{S}_1\right] = w_1
\end{aligned}$$

Recalling that $\overline{\theta} = (\theta^0, \theta), \overline{S}_n = (S_n^0, S_n)$. Now $\text{var}(\overline{\theta}.\overline{S}_1 = \theta.(V\theta)$. $\mathbb{E}\left[\overline{\theta}.\overline{S}_1\right] = \theta^0(1+r) + \theta.\mu$. So our problem is to

$$\begin{array}{ll} \text{minimise} & \theta.(V\theta),\ V \text{ invertible} \\ \text{subject to} & \theta^0 + \theta.S_0 = w_0 \quad (1) \\ & \theta^0(1+r) + \theta.\mu = w_1 \quad (2) \end{array}$$

Use (1) to eliminate θ^0 in (2).

$$(w_0 - \theta.S_0)(1+r) + \theta.\mu = w_1.$$

i.e.

$$\theta.(\mu - (1+r)S_0) = w_1 - (1+r)w_0.$$

Set $L(\theta, \lambda) = \frac{1}{\theta.(V\theta) - \lambda\theta.(\mu-(1+r)S_0)}$. At θ^*,

$$0 = \frac{\partial}{\partial\theta^i} L(\theta, \lambda) = (V\theta)^i - \lambda(\mu^i - (1+r)S_0^i).$$

So

$$\theta^* = \lambda \underbrace{(A\mu - (1+r)S_0)}_{\theta_m^* = \theta_{\text{market}}^*}, \quad A = V^{-1}.$$

Find λ using the remaining constraint

$$\lambda \underbrace{(c - 2b(1+r) + (1+r)^2 a)}_{>0 \text{ by Cauchy Schwarz}} = w_1 - (1+r)w_0,$$

where

$$a = S_0.(AS_0), b = \mu(AS_0) = S_0(A\mu), c = \mu(A\mu)$$

as before. So

$$\lambda = \frac{w_1 - (1+r)w_0}{(1+r)^2 a - 2b(1+r) + c}.$$

1.5 Portfolio selection using CARA utility

Take as utility function

$$U(x) = \text{CARA}_\gamma(x) = -e^{-\gamma x} \quad \gamma \in (0, \infty).$$

The investor has available the following set of contingent claims.

$$\mathcal{A} = \{\theta.S_1 : \theta.S_0 = w_0\}.$$

Suppose we seek to

$$\begin{array}{ll} \text{maximise} & \mathbb{E}\left[U(\theta.S_1)\right] \\ \text{subject to} & \theta.S_0 = w0 \end{array}$$

Here, S_1 has mean μ, variance V (invertible) and S_1 is Gaussian.

<u>aside</u>

$$\mathbb{E}\left[\theta.S_1\right] = \theta\mu.$$

$$\mathrm{var}(\theta.S_1) = \theta.(V\theta).$$

$\theta.S_1$ is also Gaussian. $Z \sim N(0,1), \mathbb{E}\left[e^{\lambda z}\right] = e^{-\frac{\lambda^2}{2}}$

$$\begin{aligned}\mathbb{E}\left[e^{\lambda z}\right] &= \int_{\mathbb{R}} e^{\lambda z} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\ &= e^{\frac{\lambda^2}{2}} \int_R \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-\lambda)^2}{2}}\end{aligned}$$

.

Note

$$\begin{aligned}\mathbb{E}\left[U(\theta.S_1)\right] &= -\mathbb{E}\left[e^{-\gamma\theta.S_1}\right] \\ &= -e^{-\gamma\theta. + \frac{1}{2}\gamma^2\theta.(Vg\theta)}\end{aligned}$$

.

So our problem is to Suppose we seek to

$$\begin{array}{ll}\text{maximise} & \mathbb{E}\left[U(\theta.S_1)\right] \\ \text{subject to} & \theta.S_0 = w_0\end{array}$$

Consider $L(\theta, \lambda) = \theta.\mu - \frac{1}{2}\gamma\theta.(V\theta) - \lambda\theta.S_0$ At maximiser θ^*

$$0 = \frac{\partial}{\partial\theta^i} L(\theta, \lambda) = \mu^i - \gamma(V\theta)^i - \lambda S_0^i.$$

So

$$\theta^* = \gamma^{-1}(A\mu - \lambda AS_0).$$

Find λ by

$$w_0 = \theta^*_. S_0 = \gamma^{-1}(b - \lambda a).$$

So $\lambda w_0 = b - \lambda a$. So $\lambda = \frac{b - \gamma w_0}{a}$. So $\theta^* = \underbrace{\frac{w_0}{a} AS_0}_{\theta^*_{\min}} + \gamma^{-1}(A\mu - \frac{b}{a} AS_0)$.

Add riskless bond $S_0^0 = 1, S_1^0 = 1 + r > 0$

$$\overline{\theta}.\overline{S}_0 = \theta^0 + \theta.S_0, \overline{\theta}.\overline{S}_1 = \theta^0(1+r) + \theta.S_1.$$

So

$$\mathbb{E}\left[U(\overline{\theta}.\overline{S}_1\right] = -e^{-\gamma(\theta\mu + \theta^0(1+r)) + \frac{1}{2}\gamma^2\theta.(V\theta)}.$$

$$\begin{array}{rl}\text{with constraint} & \theta.S_0 = w_0 - \theta^0 \\ \text{maximise} & \theta.\mu + \theta^0(1+r) - \dfrac{1}{2}\gamma\theta.(V\theta).\end{array}$$

Using our constraint to eliminate θ^0

$$\theta.\mu + (w_0 - \theta.S_0)(1+r) - \frac{1}{2}\gamma\theta.(V\theta).$$

Maximising θ^* satisfies

$$\mu - (1+r)S_0 = \gamma V\theta^*.$$

So

$$\theta^* = \gamma^{-1} \underbrace{(A\mu - (1+r)AS_0)}_{\theta_{\mathrm{m}}^{ast}=\theta^*_{\mathrm{market}}}.$$

$\gamma >> 1$ means we are highly risk averse.

Critique

- Easy to estimate V, but it is hard to estimate μ 5
- Why do we assume the stock prices are Gaussian? We use Centra Limit Theorem as we can consider them as the sum of random variables, but this relies on variance conditions.
- We've allowed negative asset values, consider $S_1 \sim N(\mu, V)$. More realistically,

  $$S_0 = e^{s_0}, S_1 = e^{s_0+\varepsilon Z} = S_0 e^{\varepsilon Z} \approx S_0(1+\varepsilon Z).$$

  $Z \sim N(\mu, V), \varepsilon$ small.

1.6 Capital-asset pricing model

We have seen $\theta^*_{\mathrm{m}} = A\mu + (1+r)AS_0$ appear twice. Suppose we assume that the market optimises itself. Then, we should be able to observe θ^*_{m}

$$\theta^{*i}_{\mathrm{m}} = \text{ the number of shares of asset } i.$$

$$\theta^{*i}_{\mathrm{m}} S^i_n = \text{ capitalization of asset } i.$$

Notation. Set $S^m_n = \theta^*_{\mathrm{m}}.S_n, n = 0, 1,\ \mu^{\mathrm{m}} = \theta^*_{\mathrm{m}}.\mu$ Define

$$\beta^i = \frac{\mathrm{cov}(S^i_1, S^{\mathrm{m}}_1)}{\mathrm{var}\, S^{\mu}_1}$$

the *beta* or *sensitivity* something we can estimate.

Proposition. For $i = 1, \dots, d$

$$\mu^i = (1+r)S^i_o = \beta^i(\mu^m - (1+r)S^m_0).$$

Proof. For $\theta = A\mu - (1+r)AS_0$, then

$$\mu^m - (1+r)S^m_0 = \theta.(\mu - (1+r)S_0) = \theta.(V\theta) = \mathrm{var}(\theta.S_1) = \mathrm{var}\, S^{\mu}_1.$$

So

$$\begin{aligned}\mu^i - (1+r)S^i_0 &= e_i.(\mu - (1+r)S_0)\\ &= e_i.(V\theta)\\ &= \mathrm{cov}(S^i_1, S^{\mathrm{m}}_1\\ &= \beta^i(\mu^{\mathrm{m}} - (1+r)S^{\mathrm{m}}_0)\end{aligned}$$

.

□

This appears to identify μ^i from the market. Often this pricing formula is written in terms of the returns. Define R^i, R^{m} by $S_1^0 = (1+r)S_0^0$, $S_1^i = (1+R^i)S_0^i, S_1^{\mathrm{m}} = (1+R^m)S_0^m$ Then

$$\mu^i = (1 + \mathbb{E}\left[R^i\right])S_0^i.$$

$$\mu^{\mathrm{m}} = (1 + \mathbb{E}\left[R^{\mathrm{m}}\right])S_0^{\mathrm{m}}.$$

$$\operatorname{var} S_1^{\mathrm{m}} = (S_0^{\mathrm{m}})^2 \operatorname{var}(R^{\mathrm{m}}).$$

$$\operatorname{cov}(S_1^i, S_1^{\mathrm{m}}) = S_0^i S_0^{\mathrm{m}} \operatorname{cov}(R^i, R^{\mathrm{m}}) = \frac{S_0^i S_0^m \operatorname{cov}(R^i, R^{\mathrm{m}})}{(S_0^{\mathrm{m}})^2 \operatorname{var}(R^{\mathrm{m}})}((1+\mathbb{E}\left[R^{\mathrm{m}}\right])S_0^m - (1+r)S_0^{\mathrm{m}}).$$

So

$$\mathbb{E}\left[R^i\right] - r = \hat{\beta}^i(\mathbb{E}\left[R^{\mathrm{m}} - r\right].$$

2 Martingales

2.1 Conditional probabilities and expectations

$(\Omega, \mathcal{F}, \mathbb{P})$, is a probability space. Recall for an event $B \in \mathcal{F}$ with $\mathbb{P}(B) > 0$ we define $\mathbb{P}(\cdot \mid B)$

$$\mathbb{P}\left(A \mid B\right) = \frac{\mathbb{P}\left(A \cap B\right)}{\mathbb{P}\left(B\right)}, A \in \mathcal{F}.$$

Then $\mathbb{P}\left(\cdot \mid B\right)$ has associated expectation written $\mathbb{E}\left[\cdot \mid B\right]$. This satisfies for X a random variable

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

We will need a more general notions of conditional probabilities and expectations associated not with a single event B, but with a sub-σ-algebra $\mathcal{G} \subseteq \mathcal{F}$.

Definition (σ-algebra)**.** We say $\mathcal{G}$ is a σ-*algebra* if

(i) $\emptyset \in \mathcal{G}$

(ii) $A \in \mathcal{G} \implies A^c \in \mathcal{G}$

(iii) $(A_n : n \in \mathbb{N}) \in \mathcal{G} \implies \bigcup_n A \in \mathcal{G}$

Definition (Integrable)**.** We say a random variable X is *integrable* if

$$\mathbb{E}\left[|X|\right] < \infty.$$

Definition ($\mathcal{F}$-measurable)**.** X is $\mathcal{F}$-*measurable* if $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$

Theorem. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$ and let X be an integrable / non-negative random variable. Then there exists an integrable / non-negative random variable Y satisfying

(i) Y is $\mathcal{G}$-measurable

(ii) $\mathbb{E}\left[Y 1_A\right] = \mathbb{E}\left[X 1_A\right], \ \forall\ A \in \mathcal{G}$.

Moreover, if Y' is any integrable / non-negative random variable satisfying (i) and (ii) then $Y' = Y$ almost surely. We call Y (a version of) the *conditional expectation* of X given $\mathcal{G}$ and write

$$Y = \mathbb{E}\left[X|\mathcal{G}\right] \text{ a.s. } .$$

If $\mathcal{G} = \sigma(Z)$ for some random variable Z, we write $Y = \mathbb{E}\left[X|Z\right]$ a.s. If $X = 1_A$ we write $Y = \mathbb{P}\left(A|\mathcal{G}\right)$ a.s.

Proof. Monotonicity and uniqueness argument.
Let X' be another integrable random variable with $X \leq X'$ (pointwise greater) and suppose that Y' is an integrable random variable which satisfies (i) and (ii) with respect to X'. Set $A = \{Y \geq Y'\}$ and consider the non-negative random variable

$$Z = (Y - Y')1_A.$$

Since $A \in \mathcal{G}$

$$\mathbb{E}\left[Y 1_A\right] = \mathbb{E}\left[X 1_A\right] \leq \mathbb{E}\left[X' 1_A\right] = \mathbb{E}\left[Y' 1_A\right].$$

So $\mathbb{E}\left[Z\right] \leq 0$ and so $Z = 0$ a.s. Hence, $Y \leq Y'$ a.s.
In the case that $X = X'$ a.s. we would also have $Y' \leq Y$ a.s. so $Y = Y'$ a.s.
We will omit the existence proof. $\square$

Existence for $\mathcal{G} = \{\bigcup_{n\in I} B_n : I \subseteq \mathbb{N}\}$ with $(B_n : n \in \mathbb{N})$ a partition of Ω by events. Given an integrable random variable X set

$$Y = \sum_n \mathbb{E}\left[X|B_n\right] 1_{B_n},$$

where we set $\mathbb{E}\left[X|B_n\right] = 0$ if $\mathbb{P}\left(B_n\right) = 0$. Since the B_n are disjoint, only one term is non-zero so we need not worry about convergence. Note

$$|Y| = \sum_n |\mathbb{E}\left[X|B_n\right]| 1_{B_n}$$

so by monotone convergence,

$$\begin{aligned}
\mathbb{E}\left[|Y|\right] &\overset{m}{=} \sum_n \mathbb{E}\left[|\mathbb{E}\left[X|B_n\right]| 1_{B_n}\right] \\
&= \sum_n |\mathbb{E}\left[X|B_n\right]| \mathbb{P}\left(B_n\right) \\
&\geq \sum_n \mathbb{E}\left[|X||B_n\right] \mathbb{P}\left(B_n\right) \overset{m}{=} \mathbb{E}\left[|X|\right] < \infty.
\end{aligned}$$

Hence Y is integrable and $\mathcal{G}$-measurable.

Theorem (Monotone convergence theorem)**.** Let $(X_n : n \in N)$ be a sequence of non-negative random variables. Then

$$\mathbb{E}\left[\sum_n X_n\right] = \sum_n \mathbb{E}\left[X_n\right].$$

Also for $I \subseteq \mathbb{N}$, by the dominated convergence theorem,

$$\begin{aligned}
\mathbb{E}\left[Y 1_{\bigcup_{n\in I} B_n}\right] &= \mathbb{E}\left[\sum_{n\in I} Y 1_{B_n}\right] \\
&\overset{D}{\underset{Y}{=}} \sum_{n\in I} \mathbb{E}\left[Y 1_{B_n}\right] \\
&= \sum_{n\in I} \mathbb{E}\left[\mathbb{E}\left[X|B_n\right] 1_{B_n}\right] \\
&= \sum_{n\in I} \mathbb{E}\left[X|B_n\right] \mathbb{P}\left(B_n\right) \\
&= \sum_{n\in I} \mathbb{E}\left[X 1_{B_n}\right] \\
&\overset{D}{\underset{X}{=}} \mathbb{E}\left[X 1_{\bigcup_{n\in I} B_n}\right].
\end{aligned}$$

So Y also satisfies (ii) so $Y = \mathbb{E}\left[X|\mathcal{G}\right]$

Theorem (Dominated convergence thorem)**.** Let $(X_n : n \in \mathbb{N})$ be a sequence of random variables. Suppose $\sum_n |X_n| \leq Z$ for some integrable random variable Z. Then $\sum_n X_n$ is integrable

$$\mathbb{E}\left[\sum_n X_n\right] = \sum_n \mathbb{E}\left[X_n\right].$$

Proposition. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$ and let X and W be integrable random variables. Then

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

(ii) If X is $\mathcal{G}$-measurable then $\mathbb{E}\left[X|\mathcal{G}\right] = X$ a.s.

(iii) If X is independent of $\mathcal{G}$ then $\mathbb{E}\left[X|\mathcal{G}\right] = \mathbb{E}\left[X\right]$ a.s.

(iv) If $X \geq 0$ a.s. , then $\mathbb{E}\left[X|\mathcal{G}\right] \geq 0$ a.s

(v) $\mathbb{E}\left[\alpha X + \beta W|\mathcal{G}\right] = \alpha\mathbb{E}\left[X|\mathcal{G}\right] + \beta\mathbb{E}\left[W|\mathcal{G}\right]$

Proof. (i) Take $A = \Omega$ in (ii)

(ii) Check $Y = X$ works

(iii) (Exercise) $\mathbb{E}\left[X1_A\right] = \mathbb{E}\left[X\right]\mathbb{P}\left(A\right)$, so $Y = \mathbb{E}\left[X\right]$ works.

(iv)

(v) Let Y_1 be a version of $\mathbb{E}\left[X|\mathcal{G}\right]$ and let Y_2 be a version of $\mathbb{E}\left[W|\mathcal{G}\right]$. Set $Y = \alpha Y_1 + \beta Y_2$. Then Y is integrable and $\mathcal{G}$-measurable, and for all $A \in \mathcal{G}$

$$\mathbb{E}\left[Y1_A\right] = \alpha\mathbb{E}\left[Y_1 1_A\right] + \beta\mathbb{E}\left[Y_2 1_A\right] = \alpha\mathbb{E}\left[X1_A\right] + \beta\mathbb{E}\left[W1_A\right] = \mathbb{E}\left[(\alpha X + \beta W)1_A\right].$$

Hence $Y = \mathbb{E}\left[\alpha X + \beta W|\mathcal{G}\right]$ a.s.

□

Proposition. (Tower property)
Let $\mathcal{G}$ and $\mathcal{H}$ be sub-σ-algebras of $\mathcal{F}$ with $\mathcal{H} \subseteq \mathcal{G}$ and let X be an integrable random variable. Then,

$$\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{G}\right]|\mathcal{H}\right] = \mathbb{E}\left[X|\mathcal{H}\right] \text{ a.s. }.$$

This can be visualised as the orthogonal projection onto a subspace $G \subseteq H$. If $X \subset L^2(\mathbb{P}), (\mathbb{E}\left[|X|^2\right] < \infty)$ then $\mathbb{E}\left[X|\mathcal{G}\right] \in L^2(\mathbb{P})$. $X \mapsto \mathbb{E}\left[X|\mathcal{G}\right]$ is an orthogonal projection $L^2(\mathcal{F}, \mathbb{P}) \to L^2(\mathcal{G}, \mathbb{P})$

Proof. Choose a version of $Y = \mathbb{E}\left[\mathbb{E}\left[X|\mathcal{G}\right]\mathcal{H}\right]$. Then Y is integrable, $\mathcal{H}$-measurable and for all $A \in \mathcal{H}$

$$\mathbb{E}\left[Y1_A\right] \underset{A\in\mathcal{H}}{=} \mathbb{E}\left[\mathbb{E}\left[X|\mathcal{G}\right]1_A\right] \underset{A\in\mathcal{G}}{=} \mathbb{E}\left[X1_A\right].$$

Hence $Y = \mathbb{E}\left[X|\mathcal{H}\right]$ a.s. □

Proposition. (Taking out what is known)
Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$ and let X be an integrable random variable. Suppose that Z is $\mathcal{G}$-measurable and ZX is integrable. Then

$$\mathbb{E}\left[ZX|\mathcal{G}\right] = Z\mathbb{E}\left[X|\mathcal{G}\right].$$

Proof. Suppose for now $X \geq 0$. Choose a version of $Y \geq 0$ of $\mathbb{E}[X|\mathcal{G}]$. Consider first the case $Z = 1_B$ for some $B \in \mathcal{G}$. Then for all $A \in \mathcal{G}$

$$\mathbb{E}[ZY1_A] = \mathbb{E}[Y1_{A\cap B}] = \mathbb{E}[X1_{A\cap B}] = \mathbb{E}[ZX1_A].$$

This identity

$$\mathbb{E}[ZY1_A] = \mathbb{E}[ZX1_A]$$

extends to a simple $\mathcal{G}$-measurable Z ($Z = \sum_{i=1}^n a_i 1_{B_i}, B_i \in \mathcal{G}$) by linearity. Now for $Z \geq 0$ consider the $\mathcal{G}$-measurable sets $Z_n = (2^{-n}\lfloor 2^n Z \rfloor) \wedge \wedge$. Then Z_n is simple and Z_n monotonically converges Z as $n \to \infty$. Have

$$\mathbb{E}[Z_n Y 1_A] = \mathbb{E}[Z_n X 1_A] \ \forall \ A \in \mathcal{G},$$

so by monotonice convergence we get $\mathbb{E}[ZY1_A] = \mathbb{E}[ZX1_A]$. For Z integrable, set $Z^{\pm} = (\pm Z) \wedge 0$. Then $Z = Z^+ - Z^-$ and for all $A \in \mathcal{G}$

$$\mathbb{E}\left[Z^{\pm} Y 1_A\right] = \mathbb{E}\left[Z^{\pm} X 1_A\right].$$

Subtract to see $\mathbb{E}[ZY1_A] = \mathbb{E}[ZX1_A]$. Hence $ZY = \mathbb{E}[ZX|\mathcal{G}]$ a.s. □

2.2 Definitions

Let $(\Omega, \mathcal{F})$ be a measurable space.

Definition (Filtration)**.** We say that $(\mathcal{F}_n)_{n\geq 0}$ is a *filtration* if $\mathcal{F}_n$ is a σ-algebra on Ω and $\mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \mathcal{F}$ for all n.

Definition (Random process)**.** We say that $(X_n)_{n\geq 0}$ is a *random process* if X_n is a random variable for all n.

Definition (Adapted)**.** n. We say $(X_n)_{n\geq 0}$ is *adapted* (to $(\mathcal{F}_n)_{n\geq 0}$) if X_n is $(\mathcal{F})_n$-measurable for all n.

Definition (Natural filtration)**.** Given a process $(X_n)_{n\geq 0}$ define $(F_n^X)_{n\geq 0}$ by $\mathcal{F}_n^X = \sigma(X_k : 0 \leq k \leq n)$. We call $(F_n^X)_{n\geq 0}$ the *natural filtration* of $(X_n)_{n\geq 0}$. Filtration gives us some history, so the natural filtration of X gives us the history of X.

Definition (Martingale)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with a filtration $(\mathcal{F}_n)_{n\geq 0}$. We say that a random process $(X_n)_{n\geq 0}$ is a *martingale* if

(i) X_n is $\mathcal{F})_n$-measurable for all n.

(ii) $\mathbb{E}[|X_n|] < \infty$ for all n.

(iii) $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$ a.s. for all n

If (ii) holds, we say $(X_n)_{n\geq 0}$ is integrable. Condition (iii) is called the *martingale property.* So $(X_n)_{n\geq 0}$ is a *martingale* if it is adapted, integrable and satisfies the martingale property.

Remark. – The martingale property is equivalent to

$$\mathbb{E}[(M_{n+1} - M_n)1_A] = 0 \ \forall \ A \in \mathcal{F}_n.$$

– If we take expectations of the martingale property, we get that $\mathbb{E}[M_{n+1}] = \mathbb{E}[M_n]$ so the expected value is constant.

Definition (Supermartingale)**.** If (i) and (ii) hold and also

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] \leq X_n \text{ a.s. },$$

then we say that $(X_n)_{n\geq 0}$ is a *supermartingale*

Definition (submartingale)**.** If (i) and (ii) hold and also

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] \geq X_n \text{ a.s. },$$

then we say that $(X_n)_{n\geq 0}$ is a *submartingale*

2.3 Examples

Let $(X_n)_{n\geq 1}$ be a sequence of iid random variables. Set $\mathcal{F}_0 = \{\emptyset, \Omega\}, \mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. Set $S_0 = 0, Z_0 = 1$ and define

$$S_n = X_1 + \cdots + X_n, \quad Z_n = \prod_{i=1}^n X_i \ n \geq 1.$$

In the case that X_1 is integrable and $\mathbb{E}[X_1] = 0$ the process $(S_n)_{n\geq 0}$ is a martingale - called an additive martingale. In the case that $X_1 \geq 0$ and $\mathbb{E}[X_1] = 1$ the process $(Z_n)_{n\geq 0}$ is a martingale.
Adapted: X_i is $\mathcal{F})_n$-measurable if for all $i \leq n$ so S_n and Z_n are $\mathcal{F}_n$-measurable.
Integrable: Use $|S_n| \leq |X_1| + \cdots + |X_n|$ then $\mathbb{E}[|S_n|] \leq n\mathbb{E}[|X_1|] < \infty$.

$$0 \leq Z_n = \prod_{i=1}^n X_i \text{ so } \mathbb{E}[Z_n] = \prod_{i=1}^n \mathbb{E}[X_1] = 1 < \infty.$$

Martingale property:

$$\begin{aligned}\mathbb{E}[S_{n+1}|\mathcal{F}_n] &= \mathbb{E}[X_{n+1} + S_n|\mathcal{F}_n] \\ &= \mathbb{E}[X_{n+1}|\mathcal{F}_n] + \mathbb{E}[S_n|\mathcal{F}_n] \\ &= \mathbb{E}[X_{n+1}] + S_n \\ &= S_n..\end{aligned}$$

$$\begin{aligned}\mathbb{E}[Z_{n+1}|\mathcal{F}_n] = \mathbb{E}[X_{n+1}Z_n|\mathcal{F}_n]& \\ &= Z_n\mathbb{E}[X_{n+1}|F_n] \\ &= Z_n\mathbb{E}[X_{n+1}] = Z_n \text{ a.s..}\end{aligned}$$

Example. Let $(X_n)_{n\geq 0}$ be a Markov chain with countable state-space S and transition matrix P with $\lambda_x = \mathbb{P}(X_0 = x)$. Set $\mathcal{F} = \sigma(X_0, \ldots, X_n)$. Define for f bounded or non-negative (on S)

$$Pf(x) = \sum_{y\in S} p_{xy}f(y) = \mathbb{E}_x[f(X_1)].$$

Fix $x_0, \dots, x_n \in S$ and set $A = \{X_0 = x_0, \dots, X_n = x_n\}$. Then,

$$\begin{aligned}\mathbb{E}\left[f(X_{n+1})\right] 1_A &= \sum_{y \in S} \lambda_{x_0} p_{x_0 x_1} \cdot p_{x_{n-1} x_n} p_{x_n y} f(y) \\ &= pf(x_n) \mathbb{P}\left(A\right) \\ &= \mathbb{E}\left[pf(X_n) 1_A\right].\end{aligned}$$

So $\mathbb{E}\left[f(X_{n+1} 1_A =\right] = \mathbb{E}\left[pf(X_n)1_A\right]$ for all $A \in \mathcal{F}_n$. We've shown that

$$\mathbb{E}\left[f(X_{n+1})|\mathcal{F}_n\right] = pf(X_n).$$

We say that f is subharmonic if $f(x) \leq Pf(x)\ \forall\ x \in S$. Suppose that f is subharmonic and set $M_n = f(X_n)$. Then,

$$\mathbb{E}\left[M_{n+1}|\mathcal{F}_n\right] = pf(X_n) \text{ a.s } \geq f(X_n) = M_n.$$

So $(M_n)_{n \geq 0}$ is a submartingale.

‘

Example. Continuing with our Markov chain theme, take $A, B \subseteq S$ disjoint. Set $T = \inf\{n \geq 0 : X_n \in A \cup B\}$ and define

$$u(x) = \mathbb{P}_x\left(T < \infty, X_T \in A\right).$$

Then $M_n = U(X_{T \wedge n}$ is a martingale (E_x)

2.4 Optional stopping

Definition (Stopping time)**.** We say that a random time $T : \Omega \to \mathbb{Z}^+ \cup \{\infty\}$ is a *stopping time* if $\{T \leq n\} \in \mathcal{F}_n$ for all n. Equivalently, $\{T = n\} \in \mathcal{F}_n$ for all n.

For a Markov chain, for $A \subseteq S$

$$T_a = \inf\{n \geq 0 : X_n \in A\},\ L_A = \sup\{n \geq 0 : X_n \in A\},$$

T_A is a stopping time, but in general L_A is not.

Theorem (Optional stopping time)**.** Let $(M_n)_{n \geq 0}$ be a martingale and let T be a bounded stopping time. Then $\mathbb{E}\left[M_T\right] = \mathbb{E}\left[M_0\right]$.

Proof. Choose n such that $T \leq n$. Note that

$$M_T = M_0 + (M_1 - M_0) + \cdots + (M_T - M_{T-1}) = M_0 + \sum_{k=1}^{n} 1_{\{k \leq T\}}(M_k - M_{k-1}).$$

Since T is a stopping time $\{k \leq T\} = \{T \leq k-1\}^c \in \mathcal{F}_{k-1}$. Since $(M_n)_{n \geq 0}$ is a martingale

$$\mathbb{E}\left[1_{\{k \leq T\}}(M_k - M_{k-1})\right] 0.$$

So by taking expectations of our sum, we obtain the desired result. □

Theorem. Let $(M_n)_{n \geq 0}$ be a martingale and let T be a stopping time. Suppose that there is a constant $C < \infty$ such that at least one of (i) and (ii) holds

(i) $\mathbb{P}(T < \infty) = 1$ and $|M_n| \leq C$ a.s. for all n.

(ii) $\mathbb{E}[T] < \infty$ and $|M_n - M_{n-1}| \leq C$ a.s. for all $n \leq T$.

Then, $\mathbb{E}[M_T] = \mathbb{E}[M_0]$

Proof. Since T is a stopping time, so is $T \wedge n$ for all n, so $\mathbb{E}[M_{T\wedge n}] = \mathbb{E}[M_0]$ by the previous theorem. So it suffices to show $\mathbb{E}[M_{T\wedge n}] \to \mathbb{E}[M_T]$ as $n \to \infty$. Since $T < \infty$ a.s. $M_{T\wedge n} \to M_T$ a.s. If (i) holds, then $|M_{T\wedge n}| \leq C$ for all n a.s. so $\mathbb{E}[M_{T\wedge n} \to \mathbb{E}[M_T]]$ by bounded convergence. On the other hand, if (ii) holds then $|M_{T\wedge n}| \leq |M_0| + CT$ a.s. so $\mathbb{E}[M_{T\wedge n}] \to \mathbb{E}[M_T]$ by dominated convergence using $|M_0| + CT$ as the dominating random variable. □

Example. Take $(X_n)_{n\geq 1}$ a sequence of iid random variables. Suppose $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = \frac{1}{2}$. Set $S_n = X_1 + \cdots + X_n$, $S_0 = 0$ and $T = \inf\{n \geq 0 : S_n = 1\}$. We know that $T < \infty$ a.s. and that $(S_n)_{n\geq 0}$ is a martingale. But$\mathbb{E}[S_T] = 1 \neq 0 = \mathbb{E}[S_0]$

Example. Suppose that $\mathbb{P}(X_1 = 0) = \mathbb{P}(X_1 = 2) = \frac{1}{2}$. Set $Z_n = \prod_{i=1}^n X_i, Z_0 = 1$.

$$T = \inf\{n \geq 0 : Z_n = 0\}.$$

Then $\mathbb{E}[T] = 2 < \infty$ but $\mathbb{E}[Z_T] = 0 \neq 1 = \mathbb{E}[Z_0]$.

Theorem. Let $(M_n)_{n\geq 0}$ be a martingale and let T be a stopping time. Then, $(M_{T\wedge n})_{n\geq 0}$ is also a martingale.

Proof. We have for all n

$$M_{T\wedge n} = M_0 + \sum_{k=1}^n 1_{\{k\leq T\}}(M_k - M_{k-1}).$$

Since $(M_n)_{n\geq 0}$ is adapted and integrable and $\{k \leq T\} = \{T \leq k-1\}^c \in \mathcal{F}_{k-1}$, $\mathbb{E}[|M_{T\wedge n}|] < \infty$ and $M_{T\wedge n}$ is $\mathcal{F}_n$ measurable for all n. Hence $(M_{T\wedge n})_{n\geq 0}$ is also adapted and integrable.
Now

$$M_{T\wedge(n+1)} - M_{T\wedge n} - 1_{n+1\leq T}(M_{n+1} - M_n).$$

So by taking out what is known

$$\mathbb{E}\left[M_{T\wedge(n+1)} - M_{T\wedge n}|\mathcal{F}_n\right] = \mathbb{E}\left[1_{\{n+1\leq T\}}(M_{n+1} - M_n|\mathcal{F}_n\right] = 1_{n+1\leq T}\mathbb{E}[M_{n+1} - M_n|F_n] = 0 \text{ a.s. }.$$

Hence $(M_{T\wedge n})_{n\geq 0}$ is a martingale. □

Definition (Previsible)**.** We say that a process $(H_n)_{n\geq 1}$ is *previsible* if H_n is $\mathcal{F}_{n-1}$-measurable.

Theorem. Let $(M_n)_{n\geq 0}$ be a martingale and let $(H_n)_{n\geq 1}$ be previsible process. Define

$$Y_n - \sum_{k=1}^n H_k(M_k - M_{k-1}) \text{ and } Y_0 = 0.$$

Suppose that $|H_n| \leq C$ a.s. for all n for some constant $C < \infty$. Then $(Y_n)_{n\geq 0}$ is a martingale.

Proof. We have for all n

$$Y_n = 0 + \sum_{k=1}^{n} H_k(M_k - M_{k-1}).$$

Since $(M_n)_{n\geq 0}$ is adapted and integrable and $(H_n)_{n\geq 1}$ is previsible and bounded we see $\mathbb{E}\left[|Y_n|\right] < \infty$ and Y_n is $\mathcal{F}_n$ measurable for all n. Hence $(Y_n)_{n\geq 0}$ is also adapted and integrable.
Now

$$Y_{n+1} - Y_n = H_n(M_{n+1} - M_n).$$

So by taking out what is known

$$\mathbb{E}\left[Y_{n+1} - Y_n|\mathcal{F}_n\right] = \mathbb{E}\left[H_{n+1}(M_{n+1} - M_n|\mathcal{F}_n\right] = H_{n+1}\mathbb{E}\left[M_{n+1} - M_n|F_n\right] = 0 \text{ a.s. }.$$

Hence $(Y_n)_{n\geq 0}$ is a martingale. □

Financial / gambling interpretations In a casino a martingale is a fair game - given what we know, there is no expected gain or loss. The optional stopping theorem says that

$$\mathbb{E}\left[M_T\right] = \mathbb{E}\left[M_0\right] \;\; T \leq n.$$

Suppose we hold an asset with price $(M_n)_{n\geq 0}$. Our last result says there is no way to invest boundedly in $(M_n)_{n\geq 0}$ to give positive expected reward.

3 Pricing contingent claims

3.1 Multi-period asset price model

Notation. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be equipped with a filtration $(\mathcal{F}_n)_{0\le n\le T}$ (with T a constant $\in \mathbb{N}$). Let $(\overline{S}_n)_{0\le n\le T}$ be a random process in $\mathbb{R}^{d+1}$.

$$\overline{S}_n = (S_n^0, S_n) = (S_n^0, S_n^1, \dots, S_n^d).$$

We assume that $(S_n^0)_{0\le n\le T}$ is a numeraire, that is , $S_n^0 > 0$ for all n. Often interpret $(S_n^0)_{0\le n\le T}$ as a bond or a bank account. Then write

$$S_n^0 = (1 + r_n)S_{n-1}^0, r_n \in (-1, \infty)$$

and call r_n the *interest rate*. We interpret $S_n^i, i = 1, \dots, d$ as the price of the ith *risky assets* at time n. There are T time periods: 0 to 1, 1 to 2, up to $\underbrace{T-1 \text{ to } T}_{T\text{th time period}}$.

Note theta $S_n^i - S_{n-1}^i$ is the price change of the ith asses over the nth time period.
We are not mainly looking at the absolute prices, more the discounted prices

$$X_n^i = \frac{S_n^i}{S_n^0}.$$

Let

$$\overline{X}_n = (X_n^0, X_n^1, \dots, X_n^d) = (1, X_n).$$

Let $(\theta_n)_{1\le n\le T}$ be a random process in $\mathbb{R}^{d+1}$. Write

$$\overline{\theta}_n = (\theta_n^0, \theta_n) = (\theta_n^0, \theta_n^1, \dots, \theta_n^d).$$

Suppose an investor holds θ_n^i units of asset i for the nth time period. The total price of a portfolio at the start of the nth period

$$n = 1 : \sum_{i=0}^{d} \theta_1^i S_0^i = \overline{\theta}_1.\overline{S}_0$$
$$n \ge 2 : \overline{\theta}_n.\overline{S}_{n-1}.$$

The total price of the portfolio at the end of the nth period is $\overline{\theta}_n.\overline{S}_n$.

Definition (Self-financing)**.** We say that $(\overline{\theta}_n)_{1\le n\le T}$ is *self-financing* if $\overline{\theta}_n.\overline{S}_n = \overline{\theta}_{n+1}.S_n$ for $n = 1, \dots, T-1$.

It is natural to assume that $(\overline{\theta}_n)_{1\le n\le T}$ is previsible. The investor choose $\overline{\theta}_n$ given what is known at time $n-1$. We have a value process given by

$$V_0 = \overline{\theta}_1.\overline{X}_0, V_n = \overline{\theta}_n.\overline{X}_n \overset{\text{(self financing)}}{=} \overline{\theta}_{n+1}\overline{X}_n \ n = 1, \dots, T.$$

Proposition. Let $(\theta_n)_{1\le n\le T}$ be previsible process in $\mathbb{R}^d$ and let $V_0 \in \mathbb{R}$. There exists a unique previsible process $(\theta_n^0)_{1\le n\le T} \in \mathbb{R}$ such that (for $\overline{\theta}_n = (\theta_n^0, \theta_n)$) $(\overline{\theta}_n)_{1\le n\le T}$ is a self-financing portfolio, with initial value V_0. Moreover, the associated value process is given by

$$V_n = v_0 + \sum_{k=1}^{n} \theta_k (X_k - X_{k+1}).$$

Proof. The equations $\overline{\theta}_1.\overline{X}_0 = V_0$ and $\overline{\theta}_n.\overline{X}_n = \overline{\theta}_{n+1}.\overline{X}_n$ for $n = 1, \ldots, T$ which express that $(\overline{\theta}_n)_{1 \le n \le T}$ has initial value V_0 and is self-financing. This can be written as

$$\theta_1^0 + \theta_1.X_0 = V_0, \theta_n^0 + \theta_n.Xn = \theta_{n+1}^0 + \theta_{n+1}.X_n.$$

These equations can be solved uniquely for $(\theta_n^0)_{1 \le n \le T}$ which is then previsible. Since $X_n^0 = 1$ for all n,

$$V_n - V_{n-1} = \theta_n.(X_n - X_{n-1}).$$

So by induction, we are done. □

3.2 Examples of contingent claims

Context: asset price model $(\overline{S_n})_{1 \le n \le T}$. Take $\mathcal{F}_T = \sigma(\overline{S}_0, \overline{S}_T)$. By a contingent claim of *maturity* T we mean any $\mathcal{F}_T$-measurable random variable C. Interpret this as a contract which pays C to the investor at time T.

Notation. We write

$$x^+ = \max\{x, 0\} = x \vee 0, \; x^- = \max\{-x, 0\} = (-x) \vee 0.$$

Example. (d = 1)

(i) $(S_T - K)^+$ (European) *call* of *strike price* K. This confers the right but not the obligation (the *option*) to buy one unit of the stock / asset at time T for price K.

(ii) $(S_T - K)^- = (K - S_T)^+$ *put* of K. This is the option to sell one unit at time T for price K.

(iii) Note $(S_T - K)^+ - (S_T - K)^- = S_T - K$ this is a *forward contract* it obliges you to buy a unit of stock at time T at price K.

There are more exotic options which depend not just on the final value S_T, such as *barrier options* which are *knocked in* or *knocked out* when the price crosses a given level.

Example. The *up-and-out call* C given by

$$C = \begin{cases} (S_T - K)^+ & \text{if } \max_{0 \le n \le T} S_n < B \\ 0 & \text{otherwise} \end{cases}.$$

Example. The *down-and-out call* C given by

$$C = \begin{cases} (S_T - K)^- & \text{if } \min_{0 \le n \le T} S_n \le B' \\ 0 & \text{otherwise} \end{cases}.$$

3.3 Arbitrage and completeness

(By an equivalent probability measure $\tilde{\mathbb{P}}$ on $(\Omega, \mathcal{F})$ we mean a probability measure $\tilde{\mathbb{P}}$ such that for some random variable $\rho \ge 0$

(i) $\mathbb{P}(\rho > 0) = 1$

(ii) $\tilde{\mathbb{P}}(A) = \mathbb{E}[\rho 1_A]$ $A \in \mathcal{F}$

Then we write $\tilde{\mathbb{P}} \sim \mathbb{P}$ and $\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \rho$ a.s. and call ρ (a version of) the density of $\tilde{\mathbb{P}}$ with respect to $\mathbb{P}$.