

# 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 clearing. 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

<b>0</b>	<b>Introduction</b>	<b>3</b>
<b>1</b>	<b>Utility and mean-variance analysis</b>	<b>4</b>
1.1	Contingency claims and utility functions . . . . .	4
1.2	Reservation prices and marginal prices . . . . .	5
1.3	Single period asset price model . . . . .	6
1.4	Portfolio selection using the mean-variance criterion . . . . .	7
1.5	Portfolio selection using CARA utility . . . . .	9
1.6	Capital-asset pricing model . . . . .	11
<b>2</b>	<b>Martingales</b>	<b>13</b>
2.1	Conditional probabilities and expectations . . . . .	13
2.2	Definitions . . . . .	16
2.3	Examples . . . . .	17
2.4	Optional stopping . . . . .	18

## 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  $\Omega$ , 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  $\gamma$  is given by

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

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

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

Also,

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

4

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

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

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

so we prefer  $w + X$  if

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

the Arrow-Pratt coefficient of absolute risk aversion. For  $\text{CARA}_\gamma$  this constant is equal to  $\gamma$ .

Similarly, do we prefer  $w(1 + X)$  to  $w$ ? Yes if

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

the Arrow-Pratt coefficient of relative risk aversion. For  $\text{CRR}_R$  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  $\mathbb{E}[U(X)]$  is maximised over  $\mathcal{A}$  at  $X^* \in \mathcal{A}$ . Let  $Y$  be another contingent claim. The investor would buy  $Y$  for price  $\pi$  if there exists  $X \in \mathcal{A}$  such that

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

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

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

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

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  $\pi$  at which the investor will both buy and sell. Suppose for a contradiction that there exist  $X_a, X_b$  such that

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

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

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.

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

Hence there is no such  $\pi$ .  $\square$

Recall  $U$  is concave and non-decreasing. An investor has available a set of contingent claims  $\mathcal{A}$ , and seeks to maximise  $\mathbb{E}[U(X)]$ ,  $X \in \mathcal{A}$ . 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

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

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

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

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

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

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

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

So we see

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

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

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

Call  $(\bar{S}_0, \bar{S}_1)$  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  $\bar{S}_0$  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  $\theta$  the *portfolio*. With numeraire, investor chooses  $\bar{\theta} = (\theta^0, \theta)$  such that  $\bar{\theta}.\bar{S}_0 = w_0$ . The wealth at time 1 is  $\bar{\theta}.\bar{S}_0$ .

It may be that there exists a random variable  $\rho \geq 0$  such that  $\mathbb{E}[\rho S_1^i] = S_0^i$  for all  $i$ . Then we call  $\rho$  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  $\mu$ , variance  $V$ . We assume that  $V$  is invertible and  $S_0, \mu$  are linearly independent. Suppose we are given  $w_0, w_1$ . The investor wishes to

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

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

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

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

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

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

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

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

$$\begin{aligned} \begin{pmatrix} w_0 \\ w_1 \end{pmatrix} &= M \begin{pmatrix} \lambda_0 \\ \lambda_1 \end{pmatrix}, M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}. \\ M^{-1} &= \frac{1}{\Delta} \begin{pmatrix} d & -b \\ -b & a \end{pmatrix}. \\ \begin{pmatrix} \lambda_0 \\ \lambda_1 \end{pmatrix} &= M^{-1} \begin{pmatrix} w_0 \\ w_1 \end{pmatrix}. \end{aligned}$$

So

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

The minimising variance is

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

We minimise this over  $w_1$

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

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

$$\begin{array}{ll} \text{minimise} & \text{var}(\theta \cdot S_1) \\ \text{subject to} & \theta \cdot S_0 = w_0, \end{array}$$

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

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

So

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

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

$\theta_{\min}^*$ .

Add a riskless bond / bank account.

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

Suppose we seek to

$$\begin{array}{ll} \text{minimise} & \text{var}(\bar{\theta} \cdot \bar{S}_1) \\ \text{subject to} & \bar{\theta} \cdot \bar{S}_0 = w_0 \\ & \mathbb{E}[\bar{\theta} \cdot \bar{S}_1] = w_1 \end{array}$$



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

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

Use (1) to eliminate  $\theta^0$  in (2).

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

i.e.

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

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

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

So

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

Find  $\lambda$  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

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

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.

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

Suppose we seek to

$$\begin{aligned} & \text{maximise} && \mathbb{E}[U(\theta \cdot S_1)] \\ & \text{subject to} && \theta \cdot S_0 = w_0 \end{aligned}$$

Here,  $S_1$  has mean  $\mu$ , variance  $V$  (invertible) and  $S_1$  is Gaussian.  
aside

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

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

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

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

Note

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

So our problem is to Suppose we seek to

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

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

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

So

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

Find  $\lambda$  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  $S_0^0 = 1, S_1^0 = 1 + r > 0$

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

So

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

$$\text{with constraint} \quad \theta.S_0 = w_0 - \theta^0$$

$$\text{maximise} \quad \theta.\mu + \theta^0(1 + r) - \frac{1}{2}\gamma\theta.(V\theta).$$

Using our constraint to eliminate  $\theta^0$

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

Maximising  $\theta^*$  satisfies

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

So

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

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

#### Critique

- Easy to estimate  $V$ , but it is hard to estimate  $\mu$
- Why do we assume the stock prices are Gaussian? We use Central 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,

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

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

### 1.6 Capital-asset pricing model

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

$$\theta_m^{*i} = \text{the number of shares of asset } i.$$

$$\theta_m^{*i} S_n^i = \text{capitalization of asset } i.$$

**Notation.** Set  $S_n^m = \theta_m^* S_n$ ,  $n = 0, 1$ ,  $\mu^m = \theta_m^* \mu$ . Define

$$\beta^i = \frac{\text{cov}(S_1^i, S_1^m)}{\text{var } S_1^m}$$

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

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

$$\mu^i - (1+r)S_0^i = \beta^i(\mu^m - (1+r)S_0^m).$$

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

$$\mu^m - (1+r)S_0^m = \theta \cdot (\mu - (1+r)S_0) = \theta \cdot (V\theta) = \text{var}(\theta \cdot S_1) = \text{var } S_1^\mu.$$

So

$$\begin{aligned} \mu^i - (1+r)S_0^i &= e_i \cdot (\mu - (1+r)S_0) \\ &= e_i \cdot (V\theta) \\ &= \text{cov}(S_1^i, S_1^m) \\ &= \beta^i(\mu^m - (1+r)S_0^m) \end{aligned}$$

□

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

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

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

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

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

So

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

## 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  $\mathbb{P}(\cdot | B)$

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

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

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

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

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

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

**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- $\sigma$ -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)  $\mathbb{E}[Y 1_A] = \mathbb{E}[X 1_A], \forall A \in \mathcal{G}$ .

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

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

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

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

So  $\mathbb{E}[Z] \leq 0$  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  $\Omega$  by events. Given an integrable random variable  $X$  set

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

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

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

so by monotone convergence,

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

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

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

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

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

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

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

**Theorem** (Dominated convergence theorem). 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

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

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

- (i)  $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$
- (ii) If  $X$  is  $\mathcal{G}$ -measurable then  $\mathbb{E}[X|\mathcal{G}] = X$  a.s.
- (iii) If  $X$  is independent of  $\mathcal{G}$  then  $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$  a.s.
- (iv) If  $X \geq 0$  a.s. , then  $\mathbb{E}[X|\mathcal{G}] \geq 0$  a.s
- (v)  $\mathbb{E}[\alpha X + \beta W|\mathcal{G}] = \alpha \mathbb{E}[X|\mathcal{G}] + \beta \mathbb{E}[W|\mathcal{G}]$

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

- (ii) Check  $Y = X$  works
- (iii) (Exercise)  $\mathbb{E}[X1_A] = \mathbb{E}[X]\mathbb{P}(A)$ , so  $Y = \mathbb{E}[X]$  works.
- (iv)

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

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

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

□

**Proposition.** (Tower property)

Let  $\mathcal{G}$  and  $\mathcal{H}$  be sub- $\sigma$ -algebras of  $\mathcal{F}$  with  $\mathcal{H} \subseteq \mathcal{G}$  and let  $X$  be an integrable random variable. Then,

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

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

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

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

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

□

**Proposition.** (Taking out what is known)

Let  $\mathcal{G}$  be a sub- $\sigma$ -algebra of  $\mathcal{F}$  and let  $X$  be an integrable random variable. Suppose that  $Z$  is  $\mathcal{G}$ -measurable and  $ZX$  is integrable. Then

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

*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  $n \rightarrow \infty$ . Have

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

so by monotone 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  $A \in \mathcal{G}$

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

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

## 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  $\sigma$ -algebra on  $\Omega$  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  $(\mathcal{F}_n)_n$ -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

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



- 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  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ ,  $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$ . Set  $S_0 = 0, Z_0 = 1$  and define

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

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  $|S_n| \leq |X_1| + \dots + |X_n|$  then  $\mathbb{E}[|S_n|] \leq n\mathbb{E}[|X_1|] < \infty$ .

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

Martingale property:

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

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

**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  $\mathcal{F} = \sigma(X_0, \dots, X_n)$ . 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,

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

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

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

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

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

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

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

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  $T : \Omega \rightarrow \mathbb{Z}^+ \cup \{\infty\}$  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  $A \subseteq S$

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

$T_A$  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  $\mathbb{E}[M_T] = \mathbb{E}[M_0]$ .

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

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

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

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

So by taking expectations of our sum, we obtain the desired result.  $\square$

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

**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  $S_n = X_1 + \dots + X_n$ ,  $S_0 = 0$  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  $\mathbb{E}[Z_T] = 0 \neq 1 = \mathbb{E}[Z_0]$ .