Stochastic Financial Models 1

Michael Tehranchi

6 October 2023

Remark. The Part II course Probability & Measure is listed as desirable for this course. This is because we will be dealing with random variables, and being familiar with some probability theory will be handy. There are essentially three places where we use measure-theoretic probability:

- The convergence theorems will be used to justify statements such as $\lim_n \mathbb{E}(Z_n) = \mathbb{E}(\lim_n Z_n)$.
- The notions of measurability and sigma-algebra to model what information is available in a probabilistic setting
- The monotone class theorem, which says that in order to prove an identity involving expected values, it is usually sufficient check a special case.

However, this course is self-contained, so attending Probability & Measure is absolutely **not** necessary.

1 Standing assumptions and notation

Financial market consists of d risky assets.

- No dividends.
- Infinitely divisibility.
- No bid-ask spread.
- No price impact.
- No transaction costs
- No short selling constraints

The price of asset i at time t will be denoted S_t^i. We will let $S_t = (S_t^1, \ldots, S_t^d)^\top$ be the column vector of prices. In addition, market participants can borrow or lend at a risk-free interest rate r, assumed constant.

2 The one-period set-up

Introduce an investor. Let θ^i be the number of shares of asset i that the investor buys at time $t = 0$. (When $\theta^i < 0$ then the investor shorts $|\theta^i|$ shares of the asset.) Let $\theta = (\theta^1, \ldots, \theta^d)^\top$ be the column vector of portfolio weights. In addition, let θ^0 be the amount of money the investor puts in the bank. The investor's wealth at time t is denoted X_t.

- Initial wealth $X_0 = \theta^0 + \theta^\top S_0$.
- Time-1 wealth $X_1 = \theta^0(1 + r) + \theta^\top S_1$.
- $X_1 = (1 + r)X_0 + \theta^\top [S_1 - (1 + r)S_0]$

We think of the interest rate r and the initial asset prices S_0 as known at time 0. We will model the time-1 asset prices S_1 as a random vector. Moreover, we make the (unrealistically) assumption that we are completely *certain* that we know the *distribution* of S_1. In particular, given the initial wealth X_0 and the portfolio θ, we will model the time-1 wealth X_1 as a random variable with a known distribution.

3 The mean-variance portfolio problem

Mean-variance portfolio problem (Markowitz 1952) Given initial wealth X_0 and target mean m, find the portfolio θ to minimise $\mathrm{Var}(X_1)$ subject to $\mathbb{E}(X_1) \geq m$.

We will assume the random vector S_1 is square-integrable and adopt the notation

- $\mu = \mathbb{E}(S_1)$. We will assume $\mu \neq (1 + r)S_0$.
- $V = \mathrm{Cov}(S_1) = \mathbb{E}[(S_1 - \mu)(S_1 - \mu)^\top]$. Recall that V is automatically symmetric and non-negative definite. We will *assume* that V is positive definite. In particular, the inverse V^{-1} exists.

In this notation, we have

- $\mathbb{E}(X_1) = (1 + r)X_0 + \theta^\top [\mu - (1 + r)S_0]$ and
- $\mathrm{Var}(X_1) = \theta^\top V \theta$

so the mean-variance portfolio problem is to find θ such that

$$\text{minimise } \theta^\top V \theta \quad \text{subject to } \theta^\top [\mu - (1 + r)S_0] \geq m - (1 + r)X_0$$

Theorem (Mean-variance optimal portfolio)**.** *The unique optimal solution to the mean-variance portfolio problem is*

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

where

$$\lambda = \frac{(m - (1 + r)X_0)^+}{[\mu - (1 + r)S_0]^\top V^{-1}[\mu - (1 + r)S_0]}$$

Notation. Here and throughout the course we will use the common notation $x^+ = \max\{x, 0\}$ for a real number x.

Proof. Next lecture.

Stochastic Financial Models 2

Michael Tehranchi

9 October 2023

1 Mean-variance efficiency

To solve the mean-variance portfolio problem, we will use the following lemma

Lemma. *If* $\theta^\top a = b$ *then*

$$\theta^\top V \theta \geq \frac{b^2}{a^\top V^{-1} a}$$

with equality if and only if

$$\theta = \lambda \; V^{-1} a$$

where

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

Proof of lemma. Since V is non-negative definite we have

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

and since V is positive definite there is equality only if

$$\theta = \lambda \; V^{-1} a$$

□

Remark. This proof is secretly using the Lagrangian technique from IB Optimisation or Variational Principles. The constant λ could be thought of as a Lagrange multiplier.

Remark. The lemma is equivalent to

$$(\theta^\top a)^2 \leq (\theta^\top V \theta)(a^\top V^{-1} a).$$

This is just the Cauchy–Schwarz inequality applied to the vectors $V^{1/2}\theta$ and $V^{-1/2}a$.

By applying the lemma with $a = \mu - (1+r)S_0$ and $b = \mathbb{E}(X_1) - (1+r)X_0$, we see that

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

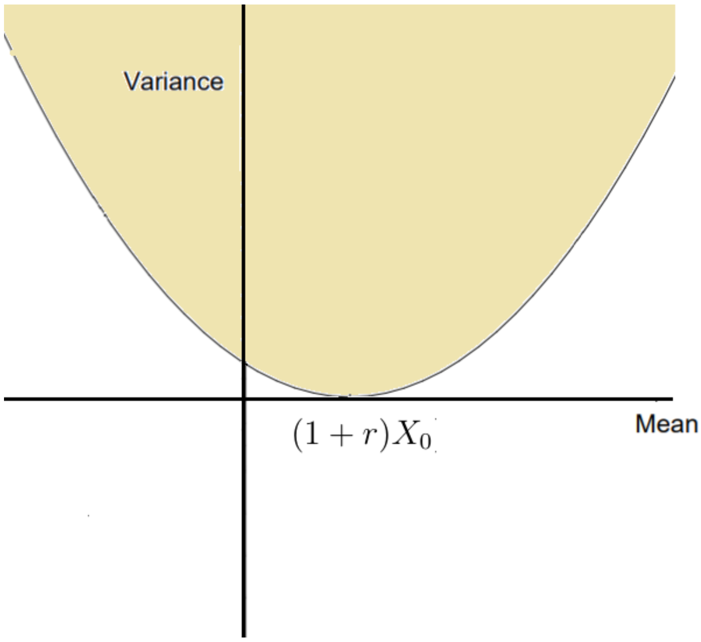
with equality if and only if

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

where

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

When the initial wealth X_0 is fixed, we can plot the set of all possible values of $(\mathbb{E}(X_1), \mathrm{Var}(X_1))$ as we vary the portfolio θ.



Definition. Given X_0, the *mean-variance efficient frontier* is the lower boundary of the set of possible values of $(\mathbb{E}(X_1), \mathrm{Var}(X_1))$; i.e. the set $\{m, (\min_{\mathbb{E}(X_1)=m} \mathrm{Var}(X_1)) : \ m \in \mathbb{R}\}$.

Remark. Note that we have shown that the mean-variance efficient frontier is a parabola.

Proof of mean-variance optimal portfolio. If $m > (1+r)X_0$, then it is optimal to take $\mathbb{E}(X_1) = m$ with portfolio $\theta = \lambda V^{-1}$, since minimised variance increases with $\mathbb{E}(X_1)$.

However, if $m \leq (1+r)X_0$, then the minimised variance decreases with $\mathbb{E}(X_1)$ and hence it is optimal to take $\mathbb{E}(X_1) = (1+r)X_0 \geq m$, with portfolio $\theta = 0$. $\square$

Definition. Given X_0, we say that a portfolio is *mean-variance efficient* iff it is the optimal solution to a mean-variance portfolio problem for *some* target mean m.

Theorem (Mutual fund theorem)**.** *A portfolio θ is mean-variance efficient if and only there exists a scalar $\lambda \geq 0$ such that*

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

Proof. We are given an initial wealth X_0.

Suppose we are given a target mean m. Then the optimal solution of the mean-variance portfolio problem if of the correct form with

$$\lambda = \frac{(m-(1+r)X_0)^+}{[\mu-(1+r)S_0]^\top V^{-1}[\mu-(1+r)S_0]} \geq 0$$

On the other hand, suppose that we are given $\lambda \geq 0$. Then the given portfolio is the optimal solution of the mean-variance portfolio problem for target mean

$$m = (1+r)X_0 + \lambda[\mu-(1+r)S_0]^\top V^{-1}[\mu-(1+r)S_0].$$

□

2 Capital Asset Pricing Model

Theorem (Linear regression coefficients)**.** *Let X and Y be two-square integrable random variables with* $\mathrm{Var}(X) > 0$*. The unique constants a and b such that*

$$Y = a + bX + Z$$

where $\mathbb{E}(Z) = 0$ *and* $\mathrm{Cov}(X,Z) = 0$ *are given by*

$$b = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)} \text{ and } a = \mathbb{E}(Y) - b\mathbb{E}(X).$$

Proof. Let $Z = Y - a - bX$ and note

$$\mathbb{E}(Z) = \mathbb{E}(Y) - a - b\mathbb{E}(X)$$
$$\mathrm{Cov}(X,Z) = \mathrm{Cov}(X,Y) - b\mathrm{Var}(X)$$

The given a and b are the unique solution to the system of equations $\mathbb{E}(Z) = 0$ and $\mathrm{Cov}(X,Z) = 0$. □

Definition. The portfolio

$$\theta_{\mathrm{Mar}} = V^{-1}[\mu - (1+r)S_0]$$

is called the *market portfolio.*

Remark. The name market portfolio is explained below.

Definition. Given initial wealth $X_0 > 0$, the excess return R^{ex} of a portfolio θ is defined by

$$R^{\text{ex}} = \frac{X_1}{X_0} - (1+r) = \frac{1}{X_0}\theta^\top [S_1 - (1+r)S_0]$$

Let $R^{\text{ex}}_{\text{Mar}}$ be the excess return of the market portfolio θ_{Mar}.

Theorem (Alpha is zero)**.** *Fix $X_0 > 0$ and a portfolio θ. Suppose α and β are such that*

$$R^{\text{ex}} = \alpha + \beta R^{\text{ex}}_{\text{Mar}} + \varepsilon$$

where $\mathbb{E}(\varepsilon) = 0$ and $\operatorname{Cov}(R^{\text{ex}}_{\text{Mar}}, \varepsilon) = 0$. Then $\alpha = 0$.

Proof. (next time) Note

$$\begin{aligned}
\operatorname{Cov}(R^{\text{ex}}, R^{\text{ex}}_{\text{Mar}}) &= \frac{1}{X_0^2}\theta^\top \operatorname{Cov}[S_1 - (1+r)S_0]\theta_{\text{Mar}} \\
&= \frac{1}{X_0^2}\theta^\top [\mu - (1+r)S_0] \\
&= \frac{1}{X_0}\mathbb{E}(R^{\text{ex}})
\end{aligned}$$

and hence

$$\begin{aligned}
\operatorname{Var}(R^{\text{ex}}_{\text{Mar}}) &= \operatorname{Cov}(R^{\text{ex}}_{\text{Mar}}, R^{\text{ex}}_{\text{Mar}}) \\
&= \frac{1}{X_0}\mathbb{E}(R^{\text{ex}}_{\text{Mar}}).
\end{aligned}$$

By linear regression, we have

$$\begin{aligned}
\beta &= \frac{\operatorname{Cov}(R^{\text{ex}}, R^{\text{ex}}_{\text{Mar}})}{\operatorname{Var}(R^{\text{ex}}_{\text{Mar}})} \\
&= \frac{\mathbb{E}(R^{\text{ex}})}{\mathbb{E}(R^{\text{ex}}_{\text{Mar}})}
\end{aligned}$$

and

$$\alpha = \mathbb{E}(R^{\text{ex}}) - \beta\mathbb{E}(R^{\text{ex}}_{\text{Mar}}) = 0.$$

□

Stochastic Financial Models 3

Michael Tehranchi

11 October 2023

1 CAPM, continued

Now let's model the entire market. Assumptions:

- There is a total of $n_i > 0$ shares of asset $i = 1, \ldots, d$, and let $n = (n_1, \ldots, n_d)^\top$.
- There are K agents in the market, and agent k holds portfolio θ_k.
- Total supply equals total demand so that

$$\sum_k \theta_k = n.$$

- Each agent's portfolio is mean-variance efficient

By the mutual fund theorem, for each k we have

$$\theta_k = \lambda_k \theta_{\mathrm{Mar}}$$

where $\lambda_k \geq 0$. Hence,

$$n = \Lambda \, \theta_{\mathrm{Mar}}$$

where $\Lambda = \sum_k \lambda_k$. Since $n \neq 0$, it follows $\Lambda > 0$. That is the say, in this model, the entire market is just some positive scalar multiple of the market portfolio (explaining the name).

A prediction of the CAPM is that when the excess returns of a portfolio are statistically regressed against the excess returns of a broad market index (such as the FTSE or S&P) then you should find $\alpha = 0$.

Remark. Markowitz and Sharpe shared the 1990 Nobel Prize in Economics for studying mean-variance efficiency and the CAPM.

2 Expected utility hypothesis

Up to now, given two random payouts X and Y we have implicitly assumed that an agent prefers X over Y if either

- $\mathbb{E}(X) > \mathbb{E}(Y)$ and $\mathrm{Var}(X) = \mathrm{Var}(Y)$, or
- $\mathbb{E}(X) = \mathbb{E}(Y)$ and $\mathrm{Var}(X) < \mathrm{Var}(Y)$

This is rather crude. Here is a historical example that illustrates one of the issues.
Aside: historical origin of expected utility hypothesis (not lectured). Consider the *St Petersburg paradox*: You and I play a game. I toss a coin repeatedly until it comes up heads. If toss the coin a total of n times, I will pay you 2^n pounds. How much would you pay me to play this game? This problem was invented by Nicolaus Bernoulli in 1713. The issue is that according to N Bernoulli's intuition, the answer should be the expected value of the payout $\sum_n 2^n \times 2^{-n} = \infty$, but he thought no sensible person would pay more than 20 pounds. His cousin Daniel Bernoulli proposed in 1738 that people don't care about the expected payout *per se*, but instead the relevant quantity is the expected *utility* of the payout.

Definition. The *expected utility hypothesis* says that each agent has a function U (called the *utility function*) such that the agent prefers random payout X to Y if and only if

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

In the case $\mathbb{E}[U(X)] = \mathbb{E}[U(Y)]$ the agent is said to be *indifferent* between X and Y.

Remark. If $\tilde{U}(x) = a + b\,U(x)$ with $b > 0$, then $\tilde{U}$ gives rise to the same expected utility preferences as U.

Remark. In 1947, von Neumann–Morgenstern axioms derived a short list of properties of an agent's preferences which are equivalent to the assumption that the agent's preferences are derived from expected utility.

3 Risk-aversion and concavity

Once we've assumed the expected utility hypothesis, there are two additional properties we will assume of the agent's utility function:

- *(Strictly) increasing.* $x > y$ implies $U(x) > U(y)$.
- *(Strictly) concave.*

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

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

Remark. Note that if $X \geq Y$ almost surely, then $X \succeq Y$. Furthermore, if $\mathbb{P}(X > Y) > 0$ then $X \succ Y$.

Remark. Recall Jensen's inequality:

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

whenever the expectations are defined. Hence $\mathbb{E}(X) \succeq X$ for any random payout X. If X is not constant, then $\mathbb{E}(X) \succ X$.

Stochastic Financial Models 4

Michael Tehranchi

13 October 2023

1 Properties of concave functions

We will nearly always assume our agent's utility function U is strictly increasing and strictly concave. If U is differentiable (always assumed), the gradient U' is called the *marginal utility.*

- $U'(x) > 0$ measures how much the utility increases at x
- $U''(x) < 0$ measures the concavity of the utility at x

Definition. The (Arrow–Pratt) *coefficient of absolute risk aversion* is

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

The (Arrow–Pratt) *coefficient of relative risk aversion* for $x > 0$ is

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

Examples

- *exponential or CARA.* $U(x) = -e^{-\gamma x}$ with $\gamma > 0$ the constant coefficient of absolute risk aversion
- *power or CRRA.* $U(x) = \frac{1}{1-R}x^{1-R}$, $x > 0$, with $R > 0, R \neq 1$, modelling the constant coefficient of relative risk aversion
- *logarithmic.* $U(x) = \log x$, $x > 0$ with constant coefficient of relative risk aversion $R = 1$.
- *risk-neutral.* $U(x) = x$ so the coefficient of risk aversion is zero. Note that this function is concave, but not strictly concave, so we won't use it as a utility function!

Remark. To be really technically accurate, we should talk about the *domain* of a concave function, i.e. the set where the function is finite-valued.

Theorem (Concave functions are continuous, and their graphs lie above their tangents)**.** *Let U be concave. Then U is continuous. If U is differentiable, then for any x, y we have*

$$U(y) \le U(x) + U'(x)(y - x).$$

Proof. Fix x and $0 < \varepsilon < \ell$. We have

$$\begin{aligned}\frac{\varepsilon}{\ell}(U(x) - U(x - \ell)) &\ge U(x) - U(x - \varepsilon) \\ &\ge U(x + \varepsilon) - U(x) \\ &\ge \frac{\varepsilon}{\ell}(U(x + \ell) - U(x))\end{aligned}$$

This is proven by looking each inequality one at a time, and rearranging the definition of concavity. For instance, note

$$x - \varepsilon = \frac{\varepsilon}{\ell}(x - \varepsilon) + (1 - \frac{\varepsilon}{\ell})x$$

so by concavity

$$U(x - \varepsilon) \ge \frac{\varepsilon}{\ell}U(x - \ell) + (1 - \frac{\varepsilon}{\ell})U(x)$$

This is equivalent to the first inequality.

Sending $\varepsilon \to 0$ shows continuity. Now assuming differentiability, dividing by ε and taking the limit yields

$$U(x) - U(x - \ell) \ge \ell\, U'(x) \ge U(x + \ell) - U(x)$$

as claimed by letting $y = x + \ell$ or $x - \ell$. □

Theorem (Increasing concave functions are unbounded on the left)**.** *Suppose U is increasing and concave, but not constant. Then $U(x) \to -\infty$ as $x \to \infty$.*

Proof. Let $x < a < b$, where $U(a) < U(b)$. Then using $a = (\frac{b-a}{b-x})x + (\frac{a-x}{b-x})b$ in the definition yields

$$U(x) \le U(a) + \frac{x - a}{b - a}(U(b) - U(a))$$

from which the conclusion follows. □

2 Optimal investment and marginal utility

In this section we assume that U is strictly increasing, concave and differentiable.

Theorem (Marginal utility pricing)**.** *Suppose U is suitably nice*[1]*, and let θ^* maximise the expected utility $\mathbb{E}[U(X_1)]$ where $X_1 = (1 + r)X_0 + \theta^\top[S_1 - (1 + r)S_0]$. Then*

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

[1]That is, it satisfies a technical condition that allows the formal calculation to go through, but the condition is uninteresting for the main focus of this course. In this case, we assume $U(X_1)$ is integrable for all portfolios θ then the formal calculation is justified by the dominated convergence theorem of Probability & Measure.

where $X_1^* = (1+r)X_0 + (\theta^*)^\top [S_1 - (1+r)S_0]$ *is the optimal time-1 wealth.*

Proof. Let

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

We can differentiate inside the expectation yielding

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

where $X_1 = (1+r)X_0 + \theta^\top [S_1 - (1+r)S_0]$. Since by calculus, at the maximising portfolio θ^* the gradient vanishes $Df(\theta^*) = 0$, the conclusion follows upon rearrangement. $\square$

Stochastic Financial Models 5

Michael Tehranchi

16 October 2023

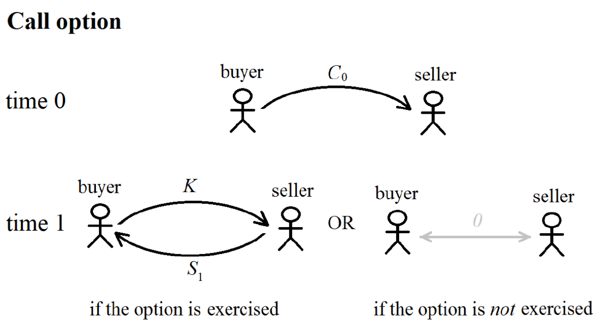
1 Contingent claims

In the context of a one-period model a *contingent claim* is just another name for an asset with a random payout at time 1.

- interest rate r and d risky assets with time t price vector S_t, for $t \in \{0, 1\}$. These are thought of as 'fundamental' assets.
- We introduce a $(d+1)$st risky asset with time-1 payout Y.
- Often $Y = g(S_1)$ for some function g, but not always.
- The problem is to find a 'reasonable' time-0 price for the claim

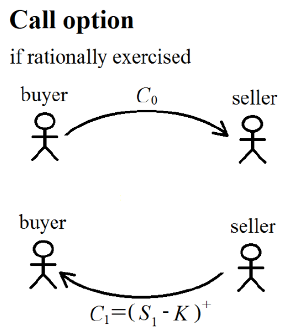
Example

Definition. A *call option* is the right, but not the obligation, to buy a certain asset at a certain price (called the strike) at a certain time in the future (the maturity date).



- If $S_1 > K$ it is rational to receive the payout $S_1 - K$.
- If $S_1 \leq K$ it is rational to let the call expire unexercised.
- The payout is $(S_1 - K)^+$

- notation: $x^+ = \max\{x, 0\}$ is the positive part of the real number x.



2 Indifference pricing

Consider an investor with initial wealth X_0 and concave, increasing utility function U. She is offered to buy a contingent claim with payout Y. How much should she pay?

- Let
$$\mathcal{X} = \{(1+r)X_0 + \theta^\top[S_1 - (1+r)S_0] : \theta \in \mathbb{R}^d\}$$
be the set of time-1 wealths attainable from trading the original market.

- The agent would prefer to buy one share of the contingent claim with time-1 payout Y for time-0 price π iff there exists an $X^* \in \mathcal{X}$ such that
$$\mathbb{E}[U(X^* + Y - (1+r)\pi)] \geq \mathbb{E}[U(X)]$$
for all $X \in \mathcal{X}$.

Assumption. In the examples from this course, we will assume that the data of the problem is such that any given utility maximisation problem has a solution.

Definition. An *indifference* (or *reservation*) price of the claim with payout Y is any solution π of
$$\max_{X \in \mathcal{X}} \mathbb{E}[U(X + Y - (1+r)\pi)] = \max_{X \in \mathcal{X}} \mathbb{E}[U(X)]$$

3 Properties of indifference prices

Theorem. *Under our assumptions*[1]*, indifference prices exist and are unique.*

[1]For the technically minded, we will assume the random variable $U(X + Y + x)$ is integrable for all $X \in \mathcal{X}$, $x \in \mathbb{R}$, and possible payouts Y, and that for x, Y there exists $X^* \in \mathcal{X}$ such that $\mathbb{E}[U(X^* + Y + x)] = \max_{X \in \mathcal{X}} \mathbb{E}[U(X + Y + x)]$

Proof. Next time.

Notation. For a fixed initial wealth X_0 and utility function U, we will let $\pi(Y)$ denote the (unique) indifference price of a contingent claim with payout Y.

Theorem (Indifference prices are increasing)**.** *If* $Y_0 \leq Y_1$ *almost surely with* $\mathbb{P}(Y_0 < Y_1) > 0$ *then*

$$\pi(Y_0) < \pi(Y_1)$$

Proof. Next time.

Theorem (Indifference prices are concave)**.** *Given random variable* Y_0, Y_1 *and* $0 < p < 1$, *we have*

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

Proof. Next time.

Definition. The marginal utility price of a claim with payout Y is

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

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

Note that our first marginal utility pricing theorem (from last time) says

$$\pi_0(a + b^\top S_1) = \frac{a}{1+r} + b^\top S_0$$

for any $a \in \mathbb{R}$ and $b \in \mathbb{R}^d$.

Theorem (Marginal utility price is larger than indifference price)**.**

$$\pi(Y) \leq \pi_0(Y)$$

Proof. Next time.

Theorem (Convergence of indifference prices to marginal utility prices)**.**

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

Proof. Next time.

Stochastic Financial Models 6

Michael Tehranchi

18 October

1 Proofs of indifference pricing properties

To prove the properties listed last time, it is convenient to define for a any suitable random variable Z the *indirect utility*

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

In this notation, π is an indifference price for the claim with payout Y iff

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

We prove two lemmata:

Lemma (Indirect utility is strictly increasing)**.** *If $Z_0 \leq Z_1$ almost surely with $\mathbb{P}(Z_0 < Z_1) > 0$ then*

$$V(Z_1) > V(Z_0)$$

Proof of lemma. Let X^i be the maximiser for the two problems, i.e.

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

for $i = 0, 1$. Then

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

□

Lemma (Indirect utility is concave)**.** *Given random variable Z_0, Z_1 and $0 < p < 1$. Then*

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

Proof of lemma. Let X^i be the maximiser for the two problems for $i = 0, 1$.

Now noting that $pX^1 + (1-p)X^0 \in \mathcal{X}$ yields

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

□

Proof of existence and uniqueness of indifference prices. By our assumption of the existence of a maximiser, we have $V(0) = \mathbb{E}[U(X^*)]$ for some $X^* \in \mathcal{X}$. In particular we have that $U(-\infty) < V(0) < U(\infty)$.

For fixed Y, we will show that the function $x \mapsto V(Y + x)$ is a bijection from $(-\infty, \infty)$ to $(U(-\infty), U(\infty))$. This would imply that there is a unique solution x to $V(Y+x) = V(0)$. The indifference price is uniquely defined by $\pi(Y) = -\frac{1}{1+r}x$.

Note the function $x \mapsto V(Y + x)$ is strictly increasing, and hence an injection. To complete the proof, we need only show its range is the interval $(U(-\infty), U(\infty))$.

The function is concave, hence continuous, so its range is an interval. Since strictly increasing concave functions are unbounded from the left, we have

$$V(Y + x) \downarrow -\infty = U(-\infty) \text{ as } x \downarrow -\infty.$$

Also

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

by a form of the monotone convergence theorem from Probability & Measure (this step is not examinable). This shows $x \mapsto V(Y + x)$ is a bijection. □

Proof that indifference prices are increasing. Suppose $Y_0 \leq Y_1$ a.s. and $\mathbb{P}(Y_0 < Y_1) > 0$. Note

$$\begin{aligned} V(Y_1 - (1+r)\pi(Y_1)) &= V(0) \\ &= V(Y_0 - (1+r)\pi(Y_0)) \\ &< V(Y_1 - (1+r)\pi(Y_0)). \end{aligned}$$

Since $x \mapsto V(Y_1 + x)$ is strictly increasing, we have $-(1+r)\pi(Y_1) < -(1+r)\pi(Y_0)$ as desired. □

Proof of concavity of indifference prices. Given Y_0, Y_1 and $0 < p < 1$, let $Y_p = pY_1 + (1-p)Y_0$ and $\pi_i = \pi(Y_i)$ for $i = 0, p, 1$. By definition of indifference price and concavity of V we have

$$\begin{aligned} V(Y_p - (1+r)\pi_p) =& V(0) \\ =& V(Y_1 - (1+r)\pi_1) \\ =& V(Y_0 - (1+r)\pi_0) \\ =& pV(Y_1 - (1+r)\pi_1) + (1-p)V(Y_0 - (1+r)\pi_0) \\ \leq& V(Y_p - (1+r)(p\pi_1 + (1-p)\pi_0)) \end{aligned}$$

Since $x \mapsto V(Y_p + x)$ is strictly increasing, we have $-(1+r)\pi_p \leq -(1+r)(p\pi_1 + (1-p)\pi_0)$. $\square$

Proof that marginal utility price is larger than indifference price. Let X^* be the optimiser without the claim, and X^1 be the optimiser with the claim. Using the supporting line property of the concave function U we have

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

where we have used the fact that

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

The conclusion follows upon rearranging. $\square$

Stochastic Financial Models 7

Michael Tehranchi

20 October 2023

1 Proof of the convergence of indifference to marginal utility price

Fix Y and let

$$\pi_t = \frac{\pi(tY)}{t}$$

and $p = \sup_{t>0} \pi_t$. Example sheet: $t \mapsto \pi_t$ decreasing. [Hint: use $\pi(0) = 0$ and concavity] Hence $\pi_t \uparrow p$ as $t \downarrow 0$. We must show $p = \pi_0(Y)$.

From last time $\pi_t \leq \pi_0(Y)$ for all $t > 0$ so $p \leq \pi_0(Y)$. It remains to show the reverse inequality.

Now by definition of $X^* \in \mathcal{X}$ as maximiser of $\mathbb{E}[U(X)]$ we have

$$\begin{aligned}
0 &= \frac{1}{t}[V(tY - (1+r)t\pi_t) - V(0)] \\
&\geq \mathbb{E}\left[\frac{U(X^* + tY - (1+r)t\pi_t) - U(X^*)}{t}\right] \\
&\geq \mathbb{E}\left[\frac{U(X^* + tY - (1+r)tp) - U(X^*)}{t}\right] \text{ since } p \geq \pi_t \\
&\to \mathbb{E}\{U'(X^*)[Y - (1+r)p]\}
\end{aligned}$$

(by the dominated convergence theorem from Probability & Measure) Rearranging yields $p \geq \pi_0(Y)$. □

2 Risk neutral measures

- Given an probability space $(\Omega, \mathcal{F}, \mathbb{P})$
- Let Z be a positive random variable such that $\mathbb{E}^{\mathbb{P}}(Z) = 1$.
- We can define a probability new measure $\mathbb{Q}$ by the formula

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

  for any event A.

- By measure theory, $\mathbb{E}^{\mathbb{Q}}(X) = \mathbb{E}^{\mathbb{P}}(ZX)$ for any $\mathbb{Q}$-integrable random variable X.
- Notation $Z = \frac{d\mathbb{Q}}{d\mathbb{P}}$
- $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is called the *density* or *likelihood ratio* of $\mathbb{Q}$ with respect to $\mathbb{P}$.
- Important point: $\mathbb{Q}(A) = 0$ if and only if $\mathbb{P}(A) = 0$ by the pigeon-hole principle.

Definition. Let $\mathbb{P}$ and $\mathbb{Q}$ be probability measures defined on the same measurable space $(\Omega, \mathcal{F})$. The measures are said to be *equivalent* if they have the property that $\mathbb{Q}(A) = 0$ if and only if $\mathbb{P}(A) = 0$.

Theorem (Radon–Nikodym theorem)**.** *Let $\mathbb{P}$ and $\mathbb{Q}$ be probability measures defined on the same measurable space $(\Omega, \mathcal{F})$. There exists a $\mathbb{P}$-a.s. positive random variable Z such that*

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

for any event A if and only if $\mathbb{P}$ and $\mathbb{Q}$ are equivalent.

Remark. We don't need this theorem, but is only stated for mathematical context.

Example

- Let $\Omega = \{\omega_1, \omega_2, \ldots\}$
- $\mathbb{P}\{\omega_i\} = p_i > 0$ for all i
- $\mathbb{Q}\{\omega_i\} = q_i > 0$ for all i
- $Z(\omega_i) = q_i/p_i$ for all i.
- Then $Z = \frac{d\mathbb{Q}}{d\mathbb{P}}$.

Example

- Let X be defined on $(\Omega, \mathcal{F}, \mathbb{P})$, and μ, λ positive constants.
- $X \sim \exp(\lambda)$ under $\mathbb{P}$.
- Let $Z = \frac{\mu}{\lambda} e^{(\lambda-\mu)X}$. Note Z is positive and

$$\mathbb{E}^{\mathbb{P}}(Z) = \int_0^\infty \frac{\mu}{\lambda} e^{(\lambda-\mu)x} \lambda e^{-\lambda x} dx = \int_0^\infty \mu e^{-\mu x} dx = 1.$$

- Let $\mathbb{Q}$ have density Z with respect to $\mathbb{P}$. Then for any bounded function f we have

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}[f(X)] &= \mathbb{E}^{\mathbb{P}}[Zf(X)] \\ &= \int_0^\infty \frac{\mu}{\lambda} e^{(\lambda-\mu)x} f(x) \lambda e^{-\lambda x} dx \\ &= \int_0^\infty f(x) \mu e^{-\mu x} dx\end{aligned}$$

- That is, the distribution of X under $\mathbb{Q}$ is $\exp(\mu)$

Now consider the one-period model set-up defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

- interest rate r
- d risky assets with time t price vector S_t.

Definition. A *risk-neutral measure* is any probability measure $\mathbb{Q}$, equivalent to $\mathbb{P}$, such that

$$S_0 = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}(S_1)$$

The probability measure $\mathbb{P}$ is called the *objective* or *statistical* measure.

Theorem (Marginal utility pricing 2)**.** *Consider the problem of maximising* $\mathbb{E}^{\mathbb{P}}[U(X)]$ *over*

$$X \in \mathcal{X} = \{(1+r)X_0 + \theta^\top(S_1 - (1+r)S_0) : \theta \in \mathbb{R}^d\}$$

where U is strictly increasing, and assume there exists a maximiser $X^* \in \mathcal{X}$. Define the equivalent probability measure $\mathbb{Q}$ with density $\frac{d\mathbb{Q}}{d\mathbb{P}} \propto U'(X^*)$. Then $\mathbb{Q}$ is risk-neutral.

Proof. Let

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

Note that $Z > 0$ and $\mathbb{E}^{\mathbb{P}}(Z) = 1$. By assumption $\frac{d\mathbb{Q}}{d\mathbb{P}} = Z$. But we already know from the first marginal utility pricing theorem (Lecture 4) that

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

$\square$

Stochastic Financial Models 8

Michael Tehranchi

23 October 2023

1 Arbitrage

Recall the set-up

- one risk-free asset with interest rate r
- d risky assets with time-t price S_t for $t \in \{0, 1\}$

Definition. An *arbitrage* is a portfolio $\varphi \in \mathbb{R}^d$ such that

$$\varphi^\top [S_1 - (1+r)S_0] \geq 0 \text{ almost surely}$$

and

$$\mathbb{P}\big(\varphi^\top [S_1 - (1+r)S_0] > 0\big) > 0.$$

Arbitrage and utility maximisation

Fix initial wealth X_0 and strictly increasing utility function U, consider the problem

$$\text{maximise } \mathbb{E}[U(X)] \text{ over } X \in \mathcal{X}$$

where

$$\mathcal{X} = \{(1+r)X_0 + \theta^\top [S_1 - (1+r)S_0] : \quad \theta \in \mathbb{R}^d\}$$

- Suppose φ is an arbitrage.
- Given $X \in \mathcal{X}$ consider
$$X^* = X + \varphi^\top [S_1 - (1+r)S_0]$$
- Note $X^* \in \mathcal{X}$ also, but
$$U(X^*) \geq U(X) \text{ almost surely}$$
and
$$\mathbb{P}\big(U(X^*) > U(X)\big) > 0$$
- Hence
$$\mathbb{E}[U(X^*)] > \mathbb{E}[U(X)]$$

- Since $X \in \mathcal{X}$ was arbitrary, there cannot be a maximiser!

Why arbitrages are bad for theory

- Suppose φ is an arbitrage.
- From above, an investor would prefer the portfolio $(n+1)\varphi$ to $n\varphi$ for any n.
- As n gets large, the assumption that an agent can trade with no price impact becomes more and more unrealistic.

Comments

- The definition of arbitrage does not depend on the agent's initial wealth X_0 or utility function U.
- However, it does depend on the agent's *beliefs* through the probability measure $\mathbb{P}$.
- Agents with equivalent beliefs will agree on the set of arbitrage portfolios.

2 Fundamental theorem of asset pricing

Things we know so far

- If there exists an optimal solution to a utility maximisation problem, then there exists risk-neutral measure.
- If there exists an optimal solution to a utility maximisation problem, then there exists no arbitrage.

Theorem (FTAP). *A market model has no arbitrage if and only if there exists a risk-neutral measure.*

Proof of the easy direction. Let φ be such that

$$\mathbb{P}\big(\varphi^\top[S_1 - (1+r)S_0] \geq 0\big) = 1.$$

Suppose there exists a risk-neutral measure $\mathbb{Q}$. By equivalence

$$\mathbb{Q}\big(\varphi^\top[S_1 - (1+r)S_0] \geq 0\big) = 1.$$

However

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\{\varphi^\top[S_1 - (1+r)S_0]\} &= \varphi^\top \mathbb{E}^{\mathbb{Q}}[S_1 - (1+r)S_0] \\ &= 0\end{aligned}$$

by the definition of risk-neutrality.

By the pigeon-hole principle

$$\mathbb{Q}\big(\varphi^\top[S_1 - (1+r)S_0] > 0\big) = 0.$$

Again by equivalence

$$\mathbb{P}\big(\varphi^\top[S_1 - (1+r)S_0] > 0\big) = 0.$$

Hence φ is not an arbitrage. $\square$

Proof of the harder direction of the FTAP. Assume that there is no arbitrage. For easier notation, let $\xi = S_1 - (1+r)S_0$.

We also assume without loss that

$$\mathbb{E}[e^{-\theta^\top \xi}] < \infty$$

for all $\theta \in \mathbb{R}^d$. (Otherwise, we replace $\mathbb{P}$ with the equivalent measure $\widetilde{\mathbb{P}}$ with density

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} \propto e^{-\|\xi\|^2}$$

and note by equivalence there is no $\widetilde{\mathbb{P}}$-arbitrage.)

Consider the problem of maximising $\mathbb{E}[U(\theta^\top \xi)]$ and $U(x) = -e^{-x}$. We will show that the assumption of no arbitrage implies that there exists an optimal solution.

Let $(\theta_n)_n$ be a sequence such that

$$\mathbb{E}[U(\theta_n^\top \xi)] \to \sup\{\mathbb{E}[U(\theta^\top \xi)] : \quad \theta \in \mathbb{R}^d\}$$

Case: $(\theta_n)_n$ is bounded. Then by the Bolzano–Weierstrass theorem, there exists a convergent subsequence. By passing to that subsequence, we assume $\theta_n \to \theta_0$.

By continuity

$$\mathbb{E}[U(\theta_n^\top \xi)] \to \mathbb{E}[U(\theta_0^\top \xi)]$$

Hence θ_0 is a maximiser. We are done since $U'(\theta_0^\top \xi)$ is proportional to the density of a risk-neutral measure.

Case: every maximising sequence $(\theta_n)_n$ is unbounded. (next time)

Stochastic Financial Models 9

Michael Tehranchi

25 October 2023

1 Harder direction of FTAP continued

We will assume without loss that the random variables $\{\xi^1, \ldots, \xi^d\}$ are linearly independent. (Otherwise, we could consider a sub-market where the asset prices are linearly independent. Since there is no arbitrage in the given market, there is no arbitrage in the sub-market.)

We may assume $\|\theta_n\| \uparrow \infty$. Let

$$\varphi_n = \frac{\theta_n}{\|\theta_n\|}$$

Note $(\varphi_n)_n$ is bounded, so by the Bolzona–Weierstrass theorem, there exists a convergent subsequence. By passing to that subsequence, we assume $\varphi_n \to \varphi_0$. Note $\|\varphi_0\| = 1$.

We will show that $\varphi_0^\top \xi \geq 0$ almost surely. By no arbitrage, this will imply that $\varphi_0^\top \xi = 0$ almost surely. And by linear independence, this would show that $\varphi_0 = 0$, contradicting $\|\varphi_0\| = 1$.

Now to show $\varphi^\top \xi \geq 0$ almost surely, that is $\mathbb{P}(\varphi_0^\top \xi < 0) = 0$. By the continuity, it is enough to show $\mathbb{P}(\varphi_0^\top \xi < -\varepsilon, \|\xi\| < r) = 0$ for every $\varepsilon > 0, r > 0$. So fix ε, r. We can pick N such that $\|\varphi_n - \varphi_0\| \leq \frac{\varepsilon}{2r}$ for $n \geq N$. Note on the event $\{\varphi_0^\top \xi < -\varepsilon, \|\xi\| < r\}$ for $n \geq N$ we have

$$\begin{aligned}
\varphi_n^\top \xi &\leq \|\varphi_n - \varphi_0\| \|\xi\| + \varphi_0^\top \xi \\
&\leq -\frac{\varepsilon}{2}
\end{aligned}$$

by Cauchy–Schwarz.

Since $\theta = 0$ is not optimal we have for $n \geq N$ that

$$\begin{aligned}
1 = F(0) &\geq F(\theta) \\
&= \mathbb{E}[e^{-\theta_n^\top \xi}] \\
&\geq \mathbb{E}[(e^{-\varphi_n^\top \xi})^{\|\theta_n\|} \mathbb{1}_{\{\varphi_0^\top \xi < -\varepsilon, \|\xi\| < r\}}] \\
&\geq e^{\frac{1}{2}\|\theta_n\| \varepsilon} \mathbb{P}(\varphi_0^\top \xi < -\varepsilon, \|\xi\| < r)
\end{aligned}$$

so $\mathbb{P}(\varphi_0^\top \xi < -\varepsilon, \|\xi\| < r) \leq e^{-\frac{1}{2}\|\theta_n\| \varepsilon} \to 0$

$\square$

Remark on examining. The details of the above proof should individually be accessible to someone in Part II, and could be examined. However, the proof in its entirety is bit longer than usual bookwork questions for this course, so don't worry too much about memorising it.

2 No-arbitrage pricing

Given a market of tradable assets and a contingent claim with payout Y, how can you assign an initial price π? Possible solutions

- Given U and X_0, find the indifference price.
- Given U and X_0, find the marginal utility price.
- Pick π such that the augmented market (consisting of the original market and the contingent claim) has no arbitrage.

Theorem. *Suppose that the original market has no arbitrage. There is no arbitrage in the augmented market if and only if there exists a risk-neutral measure for the original market such that*

$$\pi = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}(Y)$$

In particular, the set of no-arbitrage prices of the claim is an interval.

Proof. The first part is just the fundamental theorem of asset pricing. The second part. Fix two risk neutral measures $\mathbb{Q}_0$ and $\mathbb{Q}_1$ and let $\mathbb{Q}_p$ have density

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

where $0 \leq p \leq 1$. Note that $\frac{d\mathbb{Q}_p}{d\mathbb{P}}$ is strictly positive, so $\mathbb{Q}_p$ is equivalent to $\mathbb{P}$. Also

$$\mathbb{E}^{\mathbb{Q}_p}(S_1) = p\mathbb{E}^{\mathbb{Q}_1}(S_1) + (1-p)\mathbb{E}^{\mathbb{Q}_0}(S_1) = (1+r)S_0$$

and hence $\mathbb{Q}_p$ is a risk-neutral measure. Hence for any $0 \leq p \leq 1$ the expression

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

is a no-arbitrage price of the claim. This shows that the set of no-arbitrage prices is an interval. $\square$

Remark. Note that the *marginal utility price* of a claim

$$\pi_0(Y) = \frac{\mathbb{E}[U'(X_1^*)Y]}{(1+r)\mathbb{E}[U'(X_1^*)]}$$

is also a no-arbitrage price since $U'(X_1^*)$ is proportional to the density of a risk-neutral measure. However, in general we cannot say that an *indifference price* is a no-arbitrage prices, but since $\pi(Y) \leq \pi_0(Y)$, we can say it is bounded from above by a no-arbitrage price.

3 Attainable claims

Definition. A contingent claim with payout Y is *attainable* iff $Y = a + b^\top S_1$ for some $a \in \mathbb{R}$ and $b \in \mathbb{R}^d$.

Remark. We can equivalently write

$$a + b^\top S_1 = (1+r)x + b^\top [S_1 - (1+r)S_0]$$

with

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

- Attainable claims have indifference prices independent of U and X_0 (example sheet)
- Attainable claims have marginal utility prices independent of U and X_0
- Attainable claims have unique no-arbitrage prices (today)

Theorem (Attainable claims have unique no-arbitrage prices)**.** *Suppose that our given market of tradable assets has no arbitrage. If a contingent claim is attainable then there is unique initial price such that the augmented market has no arbitrage.*

Proof. Suppose

$$Y = (1+r)x + b^\top [S_1 - (1+r)S_0]$$

To show: the unique no arbitrage price is $\pi = x$.
Method 1. Use the FTAP (in lecture) The only possible no arbitrage prices of the claim are of the form

$$\pi = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}(Y) = x + \frac{b^\top}{1+r}\mathbb{E}^{\mathbb{Q}}[S_1 - (1+r)S_0] = x$$

where $\mathbb{Q}$ is a risk-neutral measure. Since the answer is always x, the no-arbitrage price is unique.
Method 2. Use the definition of arbitrage (not lectured) First, suppose $\pi = x$. Let $(\varphi^\top, \phi)^\top$ be a candidate arbitrage:

$$\varphi^\top [S_1 - (1+r)S_0] + \phi[Y - (1+r)x] \geq 0 \text{ almost surely}$$

This means

$$(\varphi + \phi b)^\top [S_1 - (1+r)S_0] \geq 0 \text{ almost surely}$$

Since the original market has no arbitrage, the almost sure inequalities are almost sure equalities. So there is no arbitrage in the augmented market. So $\pi = x$ is a no-arbitrage price.

Now suppose $\pi > x$. Note

$$b^\top [S_1 - (1+r)S_0] - [Y - (1+r)\pi] = (1+r)(\pi - x) > 0$$

so the portfolio $(b^\top, -1)^\top \in \mathbb{R}^{d+1}$ is an arbitrage in the augmented market. Otherwise, if $\pi < x$ the portfolio $(-b^\top, +1)^\top$ is an arbitrage. Hence there is exactly one price such that the augmented market has no arbitrage. $\square$

Theorem (Claims with unique no-arbitrage prices are attainable)**.** *Suppose that our given market of tradable assets has no arbitrage. A contingent claim is attainable if there is unique initial price such that the augmented market has no arbitrage.*

Proof. Use the FTAP. Details are on the example sheet. □

Stochastic Financial Models 10

Michael Tehranchi

27 October 2023

1 Examples of attainable claims

Example 1: Forward contract. A forward contract is the right *and the obligation* to buy a given asset at fixed price K (the strike) at time 1. When $d = 1$, the payout of a forward on the risky asset is given by $Y = S_1 - K$. Note that this is attainable by holding 1 share and borrowing $K/(1+r)$ from the bank. Hence the unique no-arbitrage initial price of the forward is $\pi = S_0 - K/(1+r)$

[The strike of a forward contract is usually chosen such that the initial price of the forward is zero. That is $K = (1+r)S_0$. This is called the forward price of the asset.]

Example 2: one-period binomial model. Suppose $d = 1$ as before and that S_1 can take exactly two values with $\mathbb{P}(S_1 = S_0(1+b)) = p = 1 - \mathbb{P}(S_1 = S_0(1+a))$, for constants $-1 < a < b$, where $0 < p < 1$.

First we find the risk-neutral measures. Let $\mathbb{Q}(S_1 = S_0(1+b)) = q = 1 - \mathbb{Q}(S_1 = S_0(1+a))$. Then

$$S_0 = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}(S_1) = \frac{1}{1+r}S_0(1+b)q + \frac{1}{1+r}S_0(1+a)(1-q)$$

so

$$q = \frac{r-a}{b-a} \text{ and } 1-q = \frac{b-r}{b-a}$$

Thus we learn that there exists a risk-neutral measure iff

$$a < r < b$$

in which case the risk-neutral measure is unique. This means that every contingent claim is attainable! Consider a claim with payout $Y = g(S_1)$. We need only check that the unique solution (x, θ) to

$$(1+r)x + \theta[S_1 - (1+r)S_0] = g(S_1)$$

that is, the system of equations

$$\begin{aligned}(1+r)x + \theta S_0(b-r) &= g(S_0(1+b))\\ (1+r)x + \theta S_0(a-r) &= g(S_0(1+a))\end{aligned}$$

is

$$\theta = \frac{g(S_0(1+b)) - g(S_0(1+a))}{S_0(b-a)}$$
$$x = \frac{1}{(1+r)(b-a)}[(r-a)g(S_0(1+b)) + (b-r)g(S_0(1+a))] = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}[g(S_1)]$$

2 Multi-period models

Motivating discussion

- In a one period model, we think of S_0 as constant but S_1 as random
- In a two period model, S_0 is constant, but S_1 and S_2 are random, at least as observed at time 0.
- But at time 1, we can think of both S_0 and S_1 as constant, and only S_2 is random

flow of information

- Initially, an agent has information $\mathcal{F}_0$
- at time 1, has information $\mathcal{F}_1$
- and at time 2, has information $\mathcal{F}_2$.
- Naturally, we should have $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2$
- We also want, for instance, S_0 and S_1 (but not S_2) to be $\mathcal{F}_1$-'measurable'.
- But what is information?

Given $(\Omega, \mathcal{F}, \mathbb{P})$, and a 'set of information' $\mathcal{G}$, an event $A \in \mathcal{F}$ is $\mathcal{G}$-measurable intuitively iff

$$\mathbb{P}(A|\mathcal{G}) \text{ is always either 0 or 1}$$

Example.

- Imagine flipping a coin two times.
- Let $\mathcal{G}$ be knowledge of the result of the first flip.
- $\mathbb{P}(\{HH, HT\}|\mathcal{G}) = 1$ if the first flip is heads and 0 otherwise. So $\{HH, HT\}$ is $\mathcal{G}$-measurable. That is to say, knowing $\mathcal{G}$, you can always measure whether the outcome is in $\{HH, HT\}$ or not.
- $\mathbb{P}(\{TT\}|\mathcal{G}) = 1/2$ if the first flip is tails, so $\{TT\}$ is not $\mathcal{G}$ measurable. That is, even knowing $\mathcal{G}$, sometimes you cannot perfectly measure whether the outcome is TT or not.

3 Measurability

Idea: Identify the information $\mathcal{G}$ with the collection of all $\mathcal{G}$-measurable events.
What kind of collection of events should it be?

Definition. Given a set Ω, a non-empty collection $\mathcal{G}$ of subsets of Ω is called a *sigma-algebra* iff

- $A \in \mathcal{G}$ implies $A^c \in \mathcal{G}$
- $A_1, A_2, \ldots \in \mathcal{G}$ implies $\cup_n A_n \in \mathcal{G}$.

Example. Consider tossing a coin twice. Let $\Omega = \{HH, HT, TH, TT\}$. The information measurable after the first coin toss is $\{\emptyset, \Omega, \{HH, HT\}, \{TH, TT\}, \}$

Definition. Given a sigma-algebra $\mathcal{G}$, a random variable X is $\mathcal{G}$-measurable iff the event $\{X \le x\}$ is in $\mathcal{G}$ for all $x \in \mathbb{R}$.

Remark. Intuitively, knowing the information in $\mathcal{G}$ allows you measure the value of X.
Remark. If X is $\mathcal{G}$-measurable, then the event $\{X \in B\}$ is in $\mathcal{G}$ for all 'nice' (for the measure theory specialists: Borel) subsets $B \subseteq \mathbb{R}$.
Remark. If X takes values in the countable set $\{x_1, x_2, \ldots\}$ then X is $\mathcal{G}$-measurable iff $\{X = x_i\} \in \mathcal{G}$ for all i.
Exercise. Show that if X is measurable with respect to the trivial sigma-algebra $\{\emptyset, \Omega\}$ then X is equal to a constant.

Definition. The sigma-algebra *generated* by a random variable X is the sigma-algebra $\mathcal{G}$ containing all events of the form $\{X \in B\}$ where for 'nice' subsets $B \subseteq \mathbb{R}$. Notation: $\mathcal{G} = \sigma(X)$

Theorem (Sometimes called factorisation lemma)**.** *A random variable Y is measurable respect to $\sigma(X)$ if and only if there is a 'nice' function f such that $Y = f(X)$.*

Stochastic Financial Models 11

Michael Tehranchi

October 30, 2023

1 Conditional expectation

Set up: Probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Given a sigma-algebra $\mathcal{G} \subseteq \mathcal{F}$, how to define $\mathbb{E}(X|\mathcal{G})$?
Motivation. Conditional expectation given an event

$$\mathbb{E}(X|G) = \frac{\mathbb{E}(X\mathbb{1}_G)}{\mathbb{P}(G)}$$

where X is integrable (i.e. $\mathbb{E}(|X|) < \infty$) and $\mathbb{P}(G) > 0$.
Motivation. Conditional expectation given a discrete random variable.

Suppose Y takes values in $\{y_1, y_2, \ldots\}$ and X in integrable. Let

$$f(y) = \mathbb{E}(X|Y = y)$$

Then we define

$$\mathbb{E}(X|Y) = f(Y)$$

Note that in this set-up $\mathbb{E}(X|Y)$ is $\sigma(Y)$-measurable. Also, it satisfies the *Projection property:* For any $\sigma(Y)$-measurable random event G we have

$$\mathbb{E}[X\mathbb{1}_G] = \mathbb{E}[\mathbb{E}[X|Y]\mathbb{1}_G].$$

Proof of the projection property for conditional expectation given a discrete random variable: By measurability there exists a subset $B \subseteq \{y_1, y_2, \ldots\}$ such that $G = \{Y \in B\}$. By the law of total probability

$$\begin{aligned}
\mathbb{E}[f(Y)\mathbb{1}_{\{Y\in B\}}] &= \sum_i \mathbb{P}(Y = y_i)\mathbb{E}(X|Y = y_i)\mathbb{1}_{\{y_i\in B\}} \\
&= \sum_{i:y_i\in B} \mathbb{E}(X\mathbb{1}_{\{Y=y_i\}}) \\
&= \mathbb{E}[X\mathbb{1}_{\{Y\in B\}}]
\end{aligned}$$

since

$$\sum_{i:y_i\in B} \mathbb{1}_{\{Y=y_i\}} = \mathbb{1}_{\{Y\in B\}}$$

We now use the projection property as defining property of conditional expectation:

Definition. The conditional expectation of an integrable random variable X given a sigma-algebra $\mathcal{G}$ is any $\mathcal{G}$-measurable integrable random variable Z such that

$$\mathbb{E}(X\mathbb{1}_G) = \mathbb{E}(Z\mathbb{1}_G)$$

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

Proposition (Existence and uniqueness of conditional expectations)**.** *Let X be integrable and $\mathcal{G}$ be a sigma-algebra. There exists a unique conditional expectation of X given $\mathcal{G}$.*

Proof. Existence requires some analysis. But uniqueness is straight-forward. Let Z_0, Z_1 be two conditional expectations of X given $\mathcal{G}$. By definition, this means for all $G \in \mathcal{G}$ we have

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

Now note $\{Z_0 < Z_1\}$ is in $\mathcal{G}$ since Z_1 and Z_0 are both $\mathcal{G}$-measurable by definition. Of course

$$(Z_1 - Z_0)\mathbb{1}_{\{Z_0<Z_1\}} \geq 0$$

But by equation $(*)$ we have

$$\mathbb{E}[(Z_1 - Z_0)\mathbb{1}_{\{Z_0<Z_1\}}] = 0$$

so by the pigeon-hole principle we have $(Z_1 - Z_0)\mathbb{1}_{\{Z_0<Z_1\}} = 0$ almost surely. That is to say, we have $Z_1 - Z_0 \leq 0$ almost surely. Now by symmetry we also have $Z_1 - Z_0 \geq 0$ almost surely, and hence $Z_1 = Z_0$ almost surely as claimed. □

Notation: The conditional expectation of X given $\mathcal{G}$ is denoted $\mathbb{E}(X|\mathcal{G})$. In the special case where $\mathcal{G} = \sigma(Y)$ for a random variable Y, we write $\mathbb{E}(X|Y)$ for $\mathbb{E}(X|\sigma(Y))$.

Remark. Note that we have already checked that our new definition of $\mathbb{E}(X|Y)$ agrees with our old definition in the case where Y is discrete.

The following gives an interpretation of conditional expectation given a sigma-algebra:

Proposition (Mean squared error minimisation)**.** *Suppose X is square-integrable and $\mathcal{G}$ a sigma-algebra. Then $\mathbb{E}(X|\mathcal{G})$ minimises the quantity*

$$\mathbb{E}[(X - Z)^2]$$

among all $\mathcal{G}$-measurable square-integrable Z.

Sketch of proof. By measure theory, the following extended projection property holds true. For any square-integrable $\mathcal{G}$-measurable random variable Y we have

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

Now given Z, let $Y = \mathbb{E}[X|\mathcal{G}] - Z$.

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

since Y is $\mathcal{G}$-measurable, where we have used the extension of the projection property discussed above. □

Remark. The above proof may look familiar – this is exactly how the Rao–Blackwell theorem from IB Statistics is proven.

Stochastic Financial Models 12

Michael Tehranchi

November 1, 2023

1 Properties of conditional expectations

Theorem. *Supposing all conditional expectations are defined:*

- *additivity:* $\mathbb{E}(X+Y|\mathcal{G}) = \mathbb{E}(X|\mathcal{G}) + \mathbb{E}(Y|\mathcal{G})$
- *'Pulling out a known factor': If X is $\mathcal{G}$-measurable, then* $\mathbb{E}(XY|\mathcal{G}) = X\mathbb{E}(Y|\mathcal{G})$.
- *tower property: If $\mathcal{H} \subseteq \mathcal{G}$ then*
$$\mathbb{E}[\mathbb{E}(X|\mathcal{G})|\mathcal{H}] = \mathbb{E}[\mathbb{E}(X|\mathcal{H})|\mathcal{G}] = \mathbb{E}(X|\mathcal{H})$$
- *If X is independent of $\mathcal{G}$ then* $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$.
- *positivity: If $X \geq 0$, then* $\mathbb{E}(X|\mathcal{G}) \geq 0$.
- *Jensen's inequality: If f is convex, then* $\mathbb{E}[f(X)|\mathcal{G}] \geq f[\mathbb{E}(X|\mathcal{G})]$
- *'Fix known quantity and average independent one': If X is independent of $\mathcal{G}$ and Y is $\mathcal{G}$-measurable, then*
$$\mathbb{E}[f(X,Y)|\mathcal{G}] = \mathbb{E}[f(X,y)|\mathcal{G}]\big|_{y=Y}$$

Example. Suppose X, Y are independent $N(0,1)$ random variables, and let $\mathcal{G} = \sigma(Y)$. Then

$$\mathbb{E}[f(X,Y)|\mathcal{G}] = \int f(x,Y)\varphi(x)dx$$

where φ is the probability density function of $N(0,1)$.

2 Filtrations, adaptedness and martingales

Definition. A *filtration* is a family $(\mathcal{F}_t)_{t\geq 0}$ of sigma-algebras such that $\mathcal{F}_s \subseteq \mathcal{F}_t$ for all $0 \leq s \leq t$.

Convention for this course: Unless otherwise specified, we will assume that $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

Definition. A *stochastic process* is a family $(X_t)_{t\geq 0}$ of random variables.

Definition. A stochastic process $(X_t)_{t\geq 0}$ is *adapted* to a filtration $(\mathcal{F}_t)_{t\geq 0}$ iff X_t is $\mathcal{F}_t$ measurable for all $t \geq 0$. The process is *integrable* if $\mathbb{E}(|X_t|) < \infty$ for all $t \geq 0$.

Remark. By our convention, if $(X_t)_{t\geq 0}$ is adapted to $(\mathcal{F}_t)_{t\geq 0}$, then X_0 is a constant, that is, not random.

The following definition is will be useful for examples.

Definition. The filtration $(\mathcal{F}_t)_{t\geq 0}$ *generated* by a process $(X_t)_{t\geq 0}$ is $\mathcal{F}_t = \sigma(X_s :\ 0 \leq s \leq t)$ for all $t \geq 0$. (i.e. the smallest fitration such that the process is adapted)

Definition. An adapted, integrable process $(X_t)_{t\geq 0}$ is a *martingale* with respect to a filtration $(\mathcal{F}_t)_{t\geq 0}$ iff

$$\mathbb{E}(X_t|\mathcal{F}_s) = X_s \text{ for all } 0 \leq s \leq t$$

Remark. By the rules of conditional expectations, an equivalent definition is this: An adapted, integrable process $(X_n)_{n\geq 0}$ is a martingale iff

$$\mathbb{E}(X_t - X_s|\mathcal{F}_s) = 0 \text{ for all } 0 \leq s \leq t.$$

Theorem. *An adapted, integrable discrete-time process $(X_n)_{n\geq 0}$ is a martingale with respect to a filtration $(\mathcal{F}_n)_{n\geq 0}$ iff*

$$\mathbb{E}(X_n|\mathcal{F}_{n-1}) = X_{n-1} \textit{ for all } n \geq 1.$$

Proof. If $(X_n)_{n\geq 0}$ is a martingale, then we can use the definition with $s = n-1$ and $t = n$.

Now suppose the given condition holds for all $n \geq 1$. Note that for $k \geq 0$ we have

$$\begin{aligned}\mathbb{E}(X_{s+k}|\mathcal{F}_s) &= \mathbb{E}[\mathbb{E}(X_{s+k}|\mathcal{F}_{s+k-1})|\mathcal{F}_s] \\ &= \mathbb{E}[X_{s+k-1}|\mathcal{F}_s]\end{aligned}$$

by the tower property. Hence the martingale property is proven fixing s and using induction in t. $\square$

Example. Given a filtration $(\mathcal{F}_t)_{t\geq 0}$ and an integrable random variable Y. Let $X_t = \mathbb{E}(Y|\mathcal{F}_t)$ for $t \geq 0$. Then $(X_t)_{t\geq 0}$ is a martingale.

- That X_t is integrable and $\mathcal{F}_t$-measurable is from the definition of conditional expectation.
- and $\mathbb{E}(X_t|\mathcal{F}_s) = \mathbb{E}[\mathbb{E}(Y|\mathcal{F}_t)|\mathcal{F}_s] = \mathbb{E}(Y|\mathcal{F}_s) = X_s$ by the tower property.

Stochastic Financial Models 13

Michael Tehranchi

November 3, 2023

1 Discrete-time martingales

Example.

- Let $X_1, X_2, \ldots$ be independent with $\mathbb{E}(X_n) = 0$ for all n.
- Let $S_0 = 0$ and $S_n = X_1 + \ldots + X_n$.

Then $(S_n)_{n \geq 0}$ is a martingale in the filtration generated by $(X_n)_{n \geq 1}$ since

- S_n is integrable: $\mathbb{E}(|S_n|) \leq \mathbb{E}(|X_1|) + \ldots + \mathbb{E}(|X_n|) < \infty$
- S_n is clearly $\mathcal{F}_n$ measurable (since it is a function of $X_1, \ldots, X_n$)
- $\mathbb{E}(S_n - S_{n-1}|\mathcal{F}_{n-1}) = \mathbb{E}(X_n|\mathcal{F}_{n-1}) = \mathbb{E}(X_n) = 0$ by the independence of X_n and $\mathcal{F}_{n-1}$.

Note that in this example $(S_n)_{n \geq 0}$ and $(X_n)_{n \geq 1}$ generate the same filtration

Definition. A discrete-time process $(H_n)_{n \geq 1}$ is *previsible* (or *predictable*) with respect to a filtration $(\mathcal{F}_n)_{n \geq 0}$ iff H_n is $\mathcal{F}_{n-1}$-measurable for all $n \geq 1$.

Remark. The index set for a previsible process is usually $\{1, 2, \ldots\}$.
Remark. Let $X_n = H_{n+1}$. Then $(H_n)_{n \geq 1}$ is previsible if and only if $(X_n)_{n \geq 0}$ is adapted.

Definition. The *martingale transform* of a previsible process $(H_n)_{n \geq 1}$ with respect to an adapted process $(X_n)_{n \geq 0}$ is the process defined by

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

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

Proof. Let $(H_n)_{n\geq 1}$ be bounded and previsible and $(X_n)_{n\geq 0}$ a martingale, and let $(Y_n)_{n\geq 0}$ be the martingale transform. Note that $(Y_n)_{n\geq 0}$ is adapted since each term of the formula defining Y_n is $\mathcal{F}_n$-measurable by the adaptedness of (X_n) and the previsibility of (H_n). Integrability follows from by the triangle inequality

$$\mathbb{E}(|Y_n|) \leq \mathbb{E}\left(\sum_{k=1}^{n} |H_k||X_k - X_{k-1}|\right) \leq C \sum_{k=1}^{n} \mathbb{E}(|X_k - X_{k-1}|) < \infty$$

and the integrability of (X_n) (from the definition of martingale), where $C > 0$ is the constant such that $|H_k| \leq C$ a.s. for all k (from the assumption of boundedness of (H_n))

Now

$$\begin{aligned}\mathbb{E}(Y_n - Y_{n-1}|\mathcal{F}_{n-1}) &= \mathbb{E}[H_n(X_n - X_{n-1})|\mathcal{F}_{n-1}] \\ &= H_n \mathbb{E}(X_n - X_{n-1}|\mathcal{F}_{n-1}) \\ &= 0\end{aligned}$$

by taking out what is known, and the martingale property of $(X_n)_{n\geq 0}$. $\square$

Important example from finance. Consider a market

- with a risk-free asset with interest rate r
- and d risky assets with time n prices $(S_n)_{n\geq 0}$.

and investor who

- holds the portfolio $\theta_n \in \mathbb{R}^d$ of risky assets during the time interval $(n-1, n]$,
- and the rest of his wealth is held in the risk-free asset.
- Suppose the investor is *self-financing*: his changes in wealth are explained by the changes in asset prices (but not by consumption or non-market income)

$$X_n = (1+r)X_{n-1} + \theta_n^\top [S_n - (1+r)S_{n-1}]$$

Definition. The investor's *discounted* wealth at time n is $\frac{X_n}{(1+r)^n}$. The *discounted* asset prices at time n are $\frac{S_n}{(1+r)^n}$.

Proposition. *A self-financing investor's discounted wealth is the initial wealth plus the martingale transform of the portfolio process with respect to the discounted risky asset prices.*

Proof. It is easy to see by induction that

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

$\square$

2 Stopping times

Definition. A *stopping time* for a filtration $(\mathcal{F}_t)_{t\geq 0}$ is a random variable T valued in $\{0, 1, 2, \ldots, +\infty\}$ (discrete-time) or $[0, +\infty]$ (continuous time) such that

$$\{T \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0$$

Example.

- Let $(X_n)_{n\geq 0}$ be a discrete-time adapted process.
- Let $T = \inf\{n \geq 0 : X_n > 0\}$
- Convention: $\inf \emptyset = \infty$.
- Then T is a stopping time.

Note $\{T \leq n\} = \cup_{k=0}^{n} \{X_k > 0\} \in \mathcal{F}_n$ since $\{X_k > 0\} \in \mathcal{F}_k \subseteq \mathcal{F}_n$ for all $k \leq n$. (Recall that the sigma-algebra $\mathcal{F}_n$ is closed under finite unions.)

Possible counter-example.

- Let $(X_n)_{n\geq 0}$ be an adapted process.
- Let $T = \sup\{n \geq 0 : X_n > 0\}$
- Then T is a *not* a stopping time in general.

Note $\{T \leq n\} = \cap_{k=n+1}^{\infty} \{X_k \leq 0\}$ so the event $\{T \leq n\}$ generally contains information about the future.

Stochastic Financial Models 14

Michael Tehranchi

6 November 2023

1 Optional sampling theorem

Definition. Let $(X_t)_{t\geq 0}$ be an (either discrete- or continuous-time) adapted process and T a stopping time. The *stopped process* $(X_{t\wedge T})_{t\geq 0}$ is defined by

$$X_{t\wedge T} = \begin{cases} X_t & \text{if } t \leq T \\ X_T & \text{if } t > T \end{cases}$$

Remark. Recall the notation $a \wedge b = \min\{a, b\}$ for real numbers a, b.

For the rest of the lecture, **time is discrete.**

Proposition. *Let $(X_n)_{n\geq 0}$ be an adapted process and and T a stopping time. Then the stopped process $(X_{n\wedge T})_{n\geq 0}$ is X_0 plus a martingale transform.*

Proof. Note that

$$X_{n\wedge T} = X_0 + \sum_{k=1}^{n} \mathbb{1}_{\{k\leq T\}}(X_k - X_{k-1})$$

Since $\{k \leq T\} = \{T \leq k-1\}^c \in \mathcal{F}_{k-1}$ for all $k \geq 1$ the process $(\mathbb{1}_{\{n\leq T\}})_n$ is previsible. $\square$

Corollary. *A stopped martingale is a martingale.*

Proof. This follows from the theorem that says the martingale transform of a bounded previsible process with respect to a martingale is again a martingale. $\square$

Theorem (Optional stopping theorem)**.** *Let T be a stopping time and $(X_n)_{n\geq 0}$ be a martingale such that $(X_{n\wedge T})_n$ bounded and $T < \infty$ almost surely. Then*

$$\mathbb{E}(X_T) = X_0$$

Remark. Recall our convention that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ so X_0 is constant.

Proof. Let $M_n = X_{n\wedge T}$. Note $(M_n)_{n\geq 0}$ is a martingale so that

$$\mathbb{E}(X_{n\wedge T}) = \mathbb{E}(M_n|\mathcal{F}_0) = M_0 = X_0$$

for all non-random n, by the definition of martingale and the convention on $\mathcal{F}_0$.

Now by assumption there exists a constant $C > 0$ such that $|X_{n\wedge T}| \leq C$ a.s. for all n. Also, since T is a.s. finite we have $X_{n\wedge T} \to X_T$ a.s., and hence $|X_T| \leq C$ a.s. In particular, we have

$$|X_{n\wedge T} - X_T| \leq 2C\mathbb{1}_{\{T>n\}}$$

by the triangle inequality.

Combining the two observations above,

$$\begin{aligned}
|\mathbb{E}(X_T) - X_0| &= |\mathbb{E}(X_T - X_{n\wedge T})| \\
&\leq \mathbb{E}(|X_T - X_{n\wedge T}|) \\
&\leq 2C\mathbb{P}(T > n) \\
&\to 0
\end{aligned}$$

□

Remark. It turns out that we do not need to assume that T is finite nor do we need to assume that $(X_{n\wedge T})_n$ is bounded to get the conclusion. A much weaker version of the OST is

Theorem. (A more general optional stopping theorem). Let $(X_n)_n$ be a martingale and T a stopping time such that $(X_{n\wedge T})_n$ is uniformly integrable. Then $\mathbb{E}(X_T) = X_0$.

2 Examples of the optional stopping theorem

Let $(S_n)_{n\geq 0}$ be a simple symmetric random walk starting from $S_0 = 0$, i.e. $S_n = \xi_1 + \ldots + \xi_n$ where $(\xi_n)_{n\geq 1}$ are IID $\mathbb{P}(\xi_n = \pm 1) = \frac{1}{2}$.

Example 1.

- Fix integers $a, b > 0$ and let $T = \inf\{n \geq 0 : S_n \in \{-a, b\}$.
- By Markov Chains, $T < \infty$ almost surely.
- Let $p = \mathbb{P}(S_T = -a)$ and $q = \mathbb{P}(S_T = b)$.
- By optional stopping $S_0 = 0 = \mathbb{E}(S_T) = -ap + bq$
- $p = \frac{b}{a+b}$ and $q = \frac{a}{a+b}$
- Optional stopping is justified since $|S_{T\wedge n}| \leq \max\{a, b\}$ for all n.

Counterexample 2.

- Now let $\tau = \inf\{n \geq 0 : S_n = -a\}$.
- By Markov Chains, $\tau < \infty$ almost surely. So $S_\tau = -a$.
- $\mathbb{E}(S_\tau) = -a \neq 0 = S_0$ in apparent contradiction to the optional stopping theorem.
- But note that $S_{n\wedge\tau}$ is not bounded from above, so there is no a priori reason to believe that the optional stopping theorem is applicable.

Example 3. Our goal is to find the probability generating function $\mathbb{E}(z^\tau)$ for fixed $0 < z < 1$. *Claim*: the process $w^{S_n} z^n$ is a martingale iff $w + w^{-1} = 2z^{-1}$. Indeed, note

$$\frac{\mathbb{E}(w^{S_n} z^n | \mathcal{F}_{n-1})}{w^{S_{n-1}} z^{n-1}} = z\mathbb{E}(w^{\xi_n}) = \frac{z}{2}(w + w^{-1})$$

Let $M_n = w^{S_n} z^n$ where $w + w^{-1} = 2z^{-1}$. This is a martingale with $M_\tau = w^{-a} z^\tau$. We want to apply the optional stopping theorem to conclude

$$\mathbb{E}(M_\tau) = w^{-a}\mathbb{E}(z^\tau) = M_0 = 1$$

or

$$\mathbb{E}(z^\tau) = w^a.$$

But which value of w makes the above identity true? Given z, there are two possible solutions

$$w_\pm = \frac{1 \pm \sqrt{1 - z^2}}{z}$$

and $0 < w_- < 1$ while $w_+ > 1$. In particular, since $S_{n\wedge\tau} \geq -a$ for all n and $z < 1$, then

$$w_-^{S_{n\wedge\tau}} z^{n\wedge T} \leq w_-^{-a} \text{ for all } n$$

Hence the OST is applicable and the correct formula is with $w = w_-$, i.e.

$$\mathbb{E}(z^\tau) = w_-^a = \left(\frac{1 - \sqrt{1 - z^2}}{z}\right)^a.$$

Stochastic Financial Models 15

Michael Tehranchi

8 November 2023

1 Submartingales and supermartingales

Definition. An integrable adapted process $(X_t)_{t\geq 0}$ with respect to a filtration $(\mathcal{F}_t)_{t\geq 0}$ (in either discrete- or continuous-time) is a submartingale if and only if

$$\mathbb{E}(X_t|\mathcal{F}_s) \geq X_s \text{ for all } 0 \leq s \leq t$$

An integrable adapted $(X_t)_{t\geq 0}$ is called *supermartingale* with respect to a filtration iff $(-X_t)_{t\geq 0}$ is a submartingale.

For the rest of the time, we work in **discrete-time**.

Remark. In discrete-time, a *submartingale* is an integrable adapted process $(X_n)_{n\geq 0}$ such that

$$\mathbb{E}(X_n|\mathcal{F}_{n-1}) \geq X_{n-1} \text{ for all } n \geq 1$$

by the tower property and the posivity of conditional expectation.

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

Proof. Let $(H_n)_{n\geq 1}$ be non-negative, bounded and previsible, and $(X_n)_{n\geq 0}$ a submartingale, and let $(Y_n)_{n\geq 0}$ be the martingale transform. Integrability of $(Y_n)_n$ follows from the boundedness of $(H_n)_n$ and integrability of $(X_n)_n$. The adaptedness of $(Y_n)_n$ follows from the adaptedness of both $(H_n)_n$ and $(X_n)_n$.

Now

$$\begin{aligned}
\mathbb{E}(Y_n - Y_{n-1}|\mathcal{F}_{n-1}) &= \mathbb{E}[H_n(X_n - X_{n-1})|\mathcal{F}_{n-1}] \\
&= H_n\mathbb{E}(X_n - X_{n-1}|\mathcal{F}_{n-1}) \\
&\geq 0
\end{aligned}$$

by taking out what is known, and the submartingale property of $(X_n)_{n\geq 0}$. □

Theorem. *Let $(X_n)_{n\geq 0}$ be a submartingale and $S \leq T$ are stopping times. Let*

$$M_n = X_{n\wedge T} - X_{n\wedge S}.$$

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

Proof. Note

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

Also $H_n = \mathbb{1}_{\{S<k\leq T\}} = \mathbb{1}_{\{S\leq n-1}\} - \mathbb{1}_{\{T\leq n-1\}}$ is bounded and $\mathcal{F}_{n-1}$-measurable. Hence $(M_n)_n$ is the martingale transform of a non-negative bounded previsible process with respect to a submartingale. □

Theorem (Optional sampling theorem)**.** *Let $(X_n)_{n\geq 0}$ be a submartingale and $S \leq T$ are bounded stopping times, then*

$$\mathbb{E}(X_T) \geq \mathbb{E}(X_S)$$

Proof. Let $M_n = X_{n\wedge T} - X_{n\wedge S}$. Now pick a constant N such that $T \leq N$ a.s. The conclusion follows from $\mathbb{E}(M_N) \geq M_0 = 0$ since $M_N = X_T - X_S$. □

2 Controlled Markov processes

Definition. A *Markov process* $(X_t)_{t\geq 0}$ with respect to a filtration $(\mathcal{F}_t)_{t\geq 0}$ (in either discrete- or continuous-time) is an adapted process such that

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

for all $0 \leq s \leq t$, (measurable) sets A, and where $(\mathcal{F}_t)_{t\geq 0}$.

We now work in discrete time. To check that a process $(X_n)_n$ is a martingale, we need only check

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

for all $n \geq 1$.

A useful way of to think about a Markov process is as **random dynamical system.** A Markov process valued in $\mathcal{X}$ can be constructed with

- Initial condition $X_0 = x$
- A function $G : \mathbb{N} \times \mathcal{X} \times \mathcal{V} \to \mathcal{X}$
- An sequence $(\xi_n)_{n\geq 1}$ of independent $\mathcal{V}$-valued random variable
- Then we construct the process recursively

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

  for $n \geq 1$.

Example. A simple symmetric random walk on $\mathbb{Z}$ starting at $X_0 = 0$ can be constructed as follows

- Let $\mathcal{V} = \{-1, 1\}$
- Let $(\xi_n)_{n\geq 1}$ be an IID sequence such that $\mathbb{P}(\xi_n = \pm 1) = 1/2$.
- Let $G(n, x, v) = x + v$ for all n.
- Then $X_n = G(n, X_{n-1}, \xi_n)$ for $n \geq 1$.

A **controlled Markov process** is built from

- Initial condition $X_0 = x$
- A previsible process $(U_n)_{n\geq 1}$
- A function $G : \mathbb{N} \times \mathcal{X} \times \mathcal{U} \times \mathcal{V} \to \mathcal{X}$
- A sequence $(\xi_n)_{n\geq 1}$ of independent $\mathcal{V}$-valued random variables
- Then we construct the process recursively
$$X_n^U = G(n, X_{n-1}^U, U_n, \xi_n)$$
for $n \geq 1$.

3 Stochastic optimal control

A typical problem that we will encounter is this. Given a controlled Markov process $(X_n^U)_{n\geq 0}$ and a (non-random) time horizon N we wish to

$$\text{maximise } \mathbb{E}\left[\sum_{k=1}^{N} f(k, U_k) + g(X_N^U) \middle| X_0 = x\right]$$

over previsible controls $(U_k)_{1\leq k\leq N}$, where the controlled Markov process evolves as $X_n = G(n, X_{n-1}, U_n, \xi_n)$ for $n \geq 1$ for a given function G and independent sequence $(\xi_n)_n$.

Definition. The system of equations

$$\begin{aligned} V(N, x) &= g(x) \quad \text{for all } x \\ V(n-1, x) &= \sup_u \{f(n, u) + \mathbb{E}[V(n, G(n, x, u, \xi_n))]\} \quad \text{for all } x, 1 \leq n \leq N \end{aligned}$$

is called the *Bellman equation* for the problem.

Definition. The *value function* for the problem is

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

The dynamic programming principle: Under some assumptions, the solution to the Bellman equation is the value function. (details in next lecture)