Stochastic Financial Models 1

Michael Tehranchi

October 7, 2022

1 Standing assumptions and notation

Financial market consists of d risky assets.

- No dividends.
- Infinitely divisible.
- 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.

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.

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.

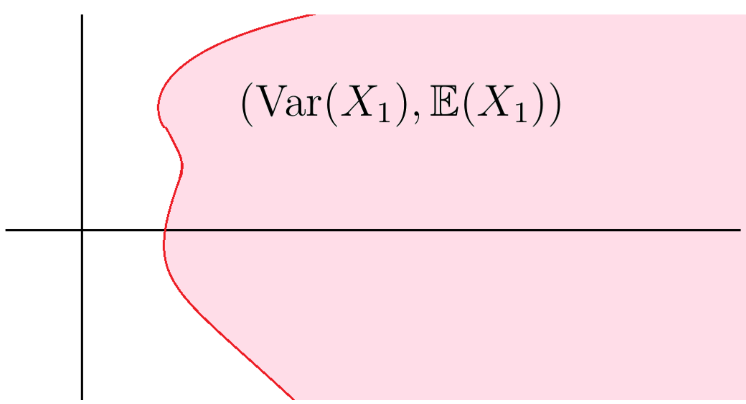
However, this course is self-contained, so attending Probability & Measure is absolutely *not necessary*.

3 Mean-variance analysis

The first problem we consider is

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

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



Definition. The *mean-variance efficient frontier* is the left-hand boundary[1] of the set of possible values of $(\mathrm{Var}(X_1), \mathbb{E}(X_1))$.

[1] i.e. the set $\{(\inf_{\mathbb{E}(X_1)=m} \mathrm{Var}(X_1), m) : m \in \mathbb{R}\}$

Definition. We say that a portfolio is *mean-variance efficient* iff it is the optimal solution to a mean-variance portfolio problem for some x and m.

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

- $\mu = \mathbb{E}(S_1)$
- $V = \mathrm{Cov}(S_1) = \mathbb{E}[(S_1 - \mu)(S_1 - \mu)^\top]$

We also will assume that $\mu \neq (1+r)S_0$ and that V is positive definite.

Theorem. *The unique optimal solution to the mean-variance problem is*

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

where

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

Notation. For a real number a, the positive part of a is defined to be $a^+ = \max\{a, 0\}$.

Stochastic Financial Models 2

Michael Tehranchi

October 10, 2022

1 The market portfolio

Definition. The portfolio

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

is called the *market portfolio.*

Given the optimal solution to the mean-variance portfolio problem from last time we immediately have:

Theorem (Mutual fund theorem)**.** *A portfolio is mean-variance efficient if and only if it a non-negative scalar multiple of the market portfolio.*

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

$$\sum_k \theta_k = n.$$

Now, if we further suppose that each agent's portfolio is mean-variance efficient, then by the mutual fund theorem $\theta_k = \lambda_k V^{-1}[\mu - (1+r)S_0]$ for some $\lambda_k \geq 0$. Hence,

$$n = \Lambda \; \theta_{\text{Mar}}$$

where $\Lambda = \sum_k \lambda_k$. That is the say, the entire market is just some positive scalar multiple of the market portfolio. We will explore the implications of this below.

In the notation introduced last time, we have

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

and the mean-variance portfolio problem becomes

$$\text{minimise } \frac{1}{2}\theta^\top V\theta \qquad \text{subject to } \theta^\top[\mu - (1+r)S_0] \geq m - (1+r)x$$

(the extra factor of 1/2 in the objective function neatens the calculation)

Proof of optimality. Let

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

where

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

Note

- $(\theta^*)^\top[\mu - (1+r)S_0] = [m-(1+r)x]^+ \geq m - (1+r)x$ (primal feasibility)
- $\lambda \geq 0$ (dueal feasibility)
- $\lambda\big(m - (1+r)x - (\theta^*)^\top[\mu - (1+r)S_0]\big)$ complementary slackness

Now for any feasible θ we have

$$\begin{aligned}
\frac{1}{2}\theta^\top V\theta &\geq \frac{1}{2}\theta^\top V\theta + \lambda\big(m - (1+r)x - \theta^\top[\mu - (1+r)S_0]\big) \\
&= \frac{1}{2}(\theta-\theta^*)^\top V(\theta-\theta^*) + \lambda\big(m-(1+r)x\big) + \frac{1}{2}(\theta^*)^\top V\theta^* \\
&\geq \lambda\big(m-(1+r)x\big) - \frac{1}{2}(\theta^*)^\top V\theta^*
\end{aligned}$$

with equality if and only if $\theta = \theta^*$.

2 Capital Asset Pricing Model

Probability fact. Let X and Y be two-square integrable random variables with $\mathrm{Var}(X) > 0$. There exist unique constants a and b such that

$$Y = a + bX + Z$$

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

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 unique solution of $\mathbb{E}(Z) = 0$ and $\mathrm{Cov}(X,Z) = 0$ is

$$b = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}$$

and $a = \mathbb{E}(Y) - b\mathbb{E}(X)$. □

Notation. Let $X_1^{\mathrm{Mar}} = (1+r)X_0^{\mathrm{Mar}} + (\theta_{\mathrm{Mar}})^\top [S_1 - (1+r)S_0]$.

Theorem. *Fix X_0 and $\theta \in \mathbb{R}^d$ and let $X_1 = (1+r)X_0 + \theta^\top [S_1 - (1+r)S_0]$. There exists a constant b such that*

$$X_1 - (1+r)X_0 = b[X_1^{\mathrm{Mar}} - (1+r)X_0^{\mathrm{Mar}}] + Z$$

where $\mathbb{E}(Z) = 0$ and $\mathrm{Cov}(Z, S_1^{\mathrm{Mar}}) = 0$.

Proof. Note

$$\begin{aligned}\mathbb{E}[X_1^{\mathrm{Mar}} - (1+r)X_0^{\mathrm{Mar}}] &= (\theta_{\mathrm{Mar}})^\top \mathbb{E}[S_1 - (1+r)S_0] \\ &= [\mu - (1+r)S_0]^\top V^{-1}[\mu - (1+r)S_0] \\ &= (\theta_{\mathrm{Mar}})^\top V(\theta^{\mathrm{Mar}} \\ &= \mathrm{Var}[X_1^{\mathrm{Mar}} - (1+r)X_0^{\mathrm{Mar}}]\end{aligned}$$

and

$$\begin{aligned}\mathrm{Cov}[X_1 - (1+r)X_0, X_1^{\mathrm{Mar}} - (1+r)X_0^{\mathrm{Mar}}] &= \theta^\top V \theta^{\mathrm{Mar}} \\ &= \theta^\top [\mu - (1+r)S_0] \\ &= \mathbb{E}[X_1 - (1+r)X_0]\end{aligned}$$

Set

$$\begin{aligned}b &= \frac{\mathrm{Cov}[X_1 - (1+r)X_0, X_1^{\mathrm{Mar}} - (1+r)X_0^{\mathrm{Mar}}]}{\mathrm{Var}[X_1^{\mathrm{Mar}} - (1+r)X_0^{\mathrm{Mar}}]} \\ &= \frac{\mathbb{E}[X_1 - (1+r)X_0]}{\mathbb{E}[X_1^{\mathrm{Mar}} - (1+r)X_0^{\mathrm{Mar}}]}\end{aligned}$$

and hence

$$\begin{aligned}a &= \mathbb{E}[X_1 - (1+r)X_0] - b\mathbb{E}[X_1^{\mathrm{Mar}} - (1+r)X_0^{\mathrm{Mar}}] \\ &= 0\end{aligned}$$

□

Capital Asset Pricing Model (Sharpe 1964) Assumptions:

- Every investor agrees on the mean $\mathbb{E}(S_1) = \mu$ and $\mathrm{Cov}(S_1) = V$, assumed positive definite.
- Every investor chooses a mean-variance efficient portfolio.
- As discussed above, the entire market is then a posisitve scalar multiple of the market portfolio.

Define the return of an investement strategy as

$$R = \frac{X_1}{X_0} - 1$$

and let R_{Mar} be defined similarly

Theorem (Alpha is zero in a mean-variance efficient market)**.** *Suppose α and β are such that*

$$R - r = \alpha + \beta(R_{\mathrm{Mar}} - r) + \varepsilon.$$

where $\mathbb{E}(\varepsilon) = 0$ and $\mathrm{Cov}(R_{\mathrm{Mar}}, \varepsilon) = 0$. Then $\alpha = 0$.

Remark. Markowitz and Sharpe shared the 1990 Nobel Prize in Economics

Stochastic Financial Models 3

Michael Tehranchi

October 12, 2022

1 The efficient frontier

By reasoning similar to our previous analysis, the optimal solution to

$$\text{minimise } \mathrm{Var}(X_1) \text{ subject to } X_0 = x \text{ and } \mathbb{E}(X_1) = m$$

is

$$\theta^* = \lambda\, \theta_{\mathrm{Mar}}$$

where

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

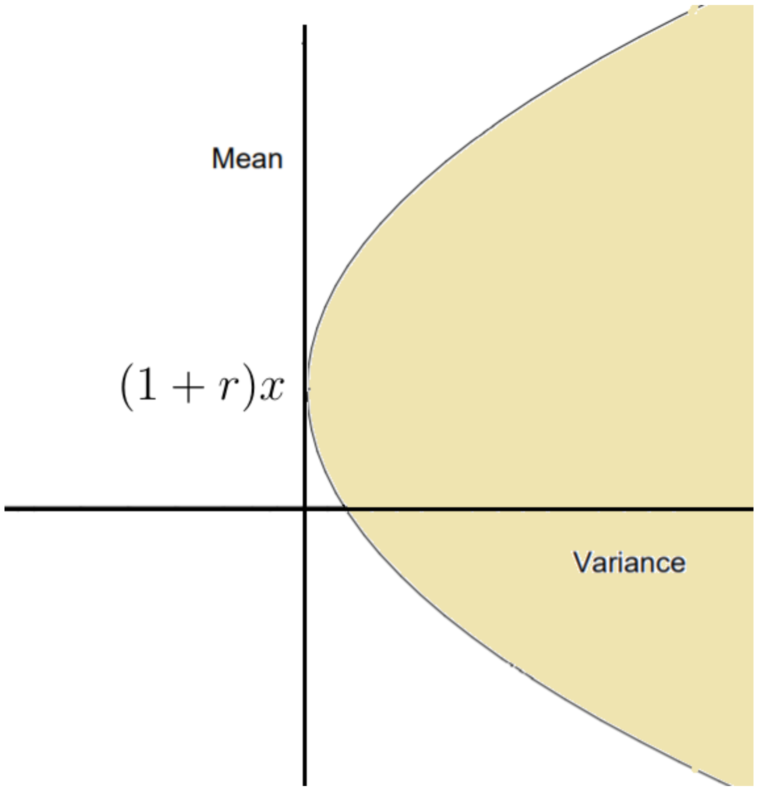
as before, but this time

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

The minimised variance is

$$\min \mathrm{Var}(X_1) = \frac{(m - (1+r)x)^2}{[\mu - (1+r)S_0]^\top V^{-1}[\mu - (1+r)S_0]}$$

So the efficient frontier is a parabola.



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. 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 the expected payout is not the relevant quantity, but 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.

Remarks

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

- Expected utility preferences are determined by the marginal distributions of the possible payouts:X is preferred to Y if and only if

$$\int U(x) dF_X(x) > \int U(x) dF_Y(x)$$

  where F_X and F_Y are the distribution functions. The full joint distribution $F_{X,Y}$ is not needed.

Aside: von Neumann–Morgenstern axioms Let X, Y, Z be random payouts, and let A be an event independent of X, Y, Z.

- Either $X \succ Y$ or $Y \prec X$ or $X \sim Y$. (completeness)

- If $X \succ Y$ and$Y \succ Z$ then $X \succ Z$.(transitivity)

- $X \succeq Y$ if and only if $\mathbb{1}_A X + \mathbb{1}_{A^c} Z \succeq \mathbb{1}_A Y + \mathbb{1}_{A^c} Z$ (independence)

- If $X \succeq Y \succeq Z$ then there exists $p \in [0,1]$ such that if $\mathbb{P}(A) = 1$ then $Y \sim \mathbb{1}_A X + \mathbb{1}_{A^c} Z$. (continuity)

If an agent has expected utility preferences, then her preference relation $\succ$ satisfies the vN–M axioms. Conversely, if the agent's preference relation satisfies the vN–M axioms plus a certain technical continuity axiom, then in 1947 von Neumann–Morgenstern proved that the preferences are derived from expected utility.

3 Risk-aversion

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

- Increasing. $x > y$ implies $U(x) > U(y)$ Note if U is increasing and $X \geq Y$ almost surely, then $X \succeq Y$.

- *Concave.*

$$U(px + qy) \geq pU(x) + qU(y)$$

  for any $x, y \in \mathbb{R}$ and$0 \leq p = 1 - q \leq 1$. Recall Jensen's inequality:

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

  whenever the expectations are defined. Hence if U is concave, $\mathbb{E}(X) \succeq X$ for any payout X.

Assume U is increasing and concave

- $U'(x) > 0$ measures how much 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* is

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

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 $R = 1$.

Stochastic Financial Models 4

Michael Tehranchi

October 14, 2022

1 State price densities

Utility maximisation problem. Given a concave, increasing utility function U and initial wealth $X_0 = x$, maximise

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

Theorem. *Suppose U is suitably nice, then*

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

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

Proof. This is just the first order condition for a maximum, where we have passed the derivative inside the expectation by the assumed niceness of U. □

Remark. A sufficient condition for 'niceness' is that U is differentiable, and that $U\big(x(1+r) + \theta^\top[S_1 - (1+r)S_0]\big)$ is integrable for all θ in a neighbourhood of θ^*.

Notice that

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

Definition. Given a market model, a *state price density* or *pricing kernel* is a positive random variable Z such that

$$\mathbb{E}[Z] = \frac{1}{1+r}$$

and

$$\mathbb{E}[ZS_1] = S_0$$

The theorem says that the marginal utility of maximised wealth is proportional to a state price density for the model, i.e. if

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

then Z is a state price density.

Aside: origin of name

- Suppose the sample space consists of d possible states, labelled $\omega_1, \ldots, \omega_d$.
- Suppose the payout of asset i is the indicator function of the event $\{\omega_i\}$:

$$S_1^i = \begin{cases} 1 & \text{if } \omega = \omega_i \\ 0 & \text{otherwise} \end{cases}$$

  These are called Arrow–Debreu securities.
- If Z is a state price density then

$$S_0^i = \mathbb{E}[ZS_1^i] = Z(\omega_i)\mathbb{P}\{\omega_i\}$$

- Rearranging

$$Z(\omega_i) = \frac{S_0^i}{\mathbb{P}\{\omega_i\}}$$

  that is, $Z(\omega_i)$ is the ratio of the price of the indicator of $\{\omega_i\}$ to the probability of the event $\{\omega_i\}$.

2 Risk neutral measures

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

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

  for any event A.
- By measure theory, $\mathbb{E}^{\mathbb{Q}}(X) = \mathbb{E}^{\mathbb{P}}(YX)$ for any $\mathbb{Q}$-integrable random variable X.
- Notation $Y = \frac{d\mathbb{Q}}{d\mathbb{P}}$
- $\frac{d\mathbb{Q}}{d\mathbb{P}}$ 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 positive random variable Y such that*

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

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

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
- $Y(\omega_i) = q_i/p_i$ for all i.
- Then $Y = \frac{d\mathbb{Q}}{d\mathbb{P}}$.

Example

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

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

- Let $\mathbb{Q}$ have density Y with respect to $\mathbb{P}$. Then

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

- 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.

By the Radon–Nikodym theorem, every risk neutral measure $\mathbb{Q}$ is related to a state price density Z by the formula

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = (1+r)Z$$

Stochastic Financial Models 5

Michael Tehranchi

October 18, 2022

1 Utility maximisation in one-period binomial model

- Let $d = 1$. Let S_1 take two values $\mathbb{P}\{S_1 = (1+b)S_0\} = p = 1 - \mathbb{P}\{S_1 = (1+a)S_0\}$ for constants $a < b$.
- Risk-free interest rate r.
- To find a risk-neutral measure, we must find q such that

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

- $q(1+b) + (1-q)(1+a) = 1 + r \Rightarrow q = \frac{r-a}{b-a}$
- A risk-neutral measure exists if and only if $a < r < b$, in which case it is unique. Assume $a < r < b$ from now on.
- Consider the problem to maximise $\mathbb{E}[U(X_1)]$ where

$$X_1 = (1+r)x + \theta[S_1 - (1+r)S_0].$$

- Assume U is increasing, concave and continuously differentiable.
- Let θ^* be an optimal portfolio and X_1^* optimised time-1 wealth.
- There exists a constant $\lambda > 0$ such that

$$U'(X_1*) = \lambda \frac{d\mathbb{Q}}{d\mathbb{P}}. \tag{*}$$

Equation (*) is equivalent to

$$\begin{aligned} U'\big((1+r)x + \theta^* S_0(b-r)\big) &= \frac{\lambda(r-a)}{p(b-a)} \\ U'\big((1+r)x - \theta^* S_0(r-a)\big) &= \frac{\lambda(b-r)}{(1-p)(b-a)} \end{aligned}$$

There are two unknowns θ^* and λ. There are at least two ways to solve this system.

Method 1. (not lectured) The optimal stock holding θ^* solves

$$\frac{U'\big((1+r)x + \theta^* S_0(b-r)\big)}{U'\big((1+r)x - \theta^* S_0(r-a)\big)} = \frac{(1-p)(r-a)}{p(b-r)}$$

This is a non-linear equation, so we can't proceed further in general.

Method 2. (lectured) Let $I = (U')^{-1}$ be the inverse function of the decreasing continuous function U'. (This means that the inverse marginal utility function I is also continuous and decreasing.) Equation (*) becomes

$$X_1^* = I\left(\lambda \frac{d\mathbb{Q}}{d\mathbb{P}}\right)$$

Compute the expected value of both sides with respect to $\mathbb{Q}$ to get

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}(X_1^*) &= (1+r)x + \theta^* \mathbb{E}^{\mathbb{Q}}[S_1 - (1+r)S_0] \\
&= (1+r)x \\
&= \mathbb{E}^{\mathbb{Q}}\left[I\left(\lambda \frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] \\
&= \frac{r-a}{b-a} I\left(\frac{\lambda(r-a)}{p(b-a)}\right) + \frac{b-r}{b-a} I\left(\frac{\lambda(b-r)}{(1-p)(b-a)}\right)
\end{aligned}$$

This is also a non-linear equation for λ. We see that the right-hand side is decreasing and continuous in λ so there is a unique solution. Plugging this solution into equation (*) yields the unique θ^* that we were looking for.

With either method, our problem simplifies considerably if we assume that the utility function has constant relative risk aversion (CRRA), that is, $U(x) = \frac{1}{1-R} x^{1-R}$ for $x > 0$, where $R > 0, R \neq 1$ is a given constant. In this case the marginal utility is $U'(x) = x^{-R}$ and so the inverse marginal utility is $I(y) = y^{-1/R}$. When all the dust clears, the optimal solution is (not done in lectures)

$$\theta^* = \frac{x(1+r)[p^{1/R}(b-r)^{1/R} - (1-p)^{1/R}(r-a)^{1/R}]}{S_0[(r-a))[p^{1/R}(b-r)^{1/R} + (b-r)(1-p)^{1/R}(r-a)^{1/R}]}$$

It is an example sheet question to show that θ^* has the same sign as $\mathbb{E}(S_1) - (1+r)S_0$.

2 Contingent claims

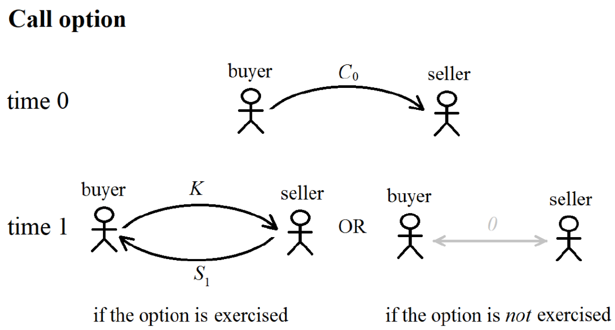
In the context of a one-period model a *contingent claim* is just another name for an asset with a random payout at time 1.

- 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.

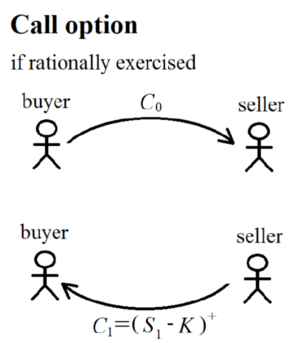
- We introduce a $(d+1)$st risky asset with time-1 payout Y.
- Often $Y = g(S_1)$ for some function g, but not always.
- The problem is to find a 'reasonable' time-0 price for the claim

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



3 Indifference pricing

- Let

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

be the set of time-1 wealths attainable from trading the market with initial wealth x.

- An agent with initial wealth x and utility function U would be willing to buy one share of the contingent claim with time-1 payout Y for time-0 price π iff there exists an $X^* \in \mathcal{X}(x)$ such that

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

for all $X \in \mathcal{X}(x)$.

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.
Assumption. In the examples from this course, we assume that the quantity

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

is a decreasing, continuous function of π.

Definition. The *indifference* (or *reservation*) price of the claim with payout Y is the unique solution $\pi(Y)$ of

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

Stochastic Financial Models 6

Michael Tehranchi

October 19, 2022

1 Concavity of indifference price

We will assume

- the utility function U is differentiable, concave and increasing.
- the payouts Y of the contingent claims that we want to price are such that for all π the problem to maximise $\mathbb{E}[U(X+Y-(1+r)\pi)]$ over $X \in \mathcal{X}(x)$ has an optimal solution.

We assumed last time that the function $\pi \mapsto \max_{X\in\mathcal{X}(x)} \mathbb{E}[U(X+Y-(1+r)\pi)]$ is decreasing and continuous so that indifference prices are uniquely defined. As a warm-up for some arguments we'll make in this lecture, here is a proof.

Fix π and $\varepsilon > 0$, let X_ε be the maximiser of $\mathbb{E}[U(X+Y-(1+r)(\pi+\varepsilon))]$.

$$\begin{aligned}
\max_{X\in\mathcal{X}(x)} \mathbb{E}[U(X+Y-(1+r)(\pi+\varepsilon))] &= \mathbb{E}[U(X_\varepsilon+Y-(1+r)(\pi+\varepsilon))] \\
&< \mathbb{E}[U(X_\varepsilon+Y-(1+r)\pi)] \text{ since } U \text{ is increasing} \\
&\le \max_{X\in\mathcal{X}(x)} \mathbb{E}[U(X+Y-(1+r)\pi)]
\end{aligned}$$

Continuity follows from the concavity of the function $\pi \mapsto \max_{X\in\mathcal{X}(x)} \mathbb{E}[U(X+Y-(1+r)\pi)]$). The proof of concavity is omitted here, but follows from a similar argument to one appearing below.

Theorem. *The indifference pricing function $Y \mapsto \pi(Y)$ is concave.*

Proof of theorem. Fix Y_0, Y_1 and let $\pi_i = \pi(Y_i)$. For $0 \le p = 1-q \le 1$, let $Y_p = pY_1 + qY_0$, let $\pi_p = \pi(Y_p)$.

Let $X_i \in \mathcal{X}(x)$ be the maximiser for $i = 0, 1$ so that

$$\begin{aligned}
\max_{X\in\mathcal{X}(x)} \mathbb{E}[U(X+Y_p-(1+r)\pi_p)] = \max_{X\in\mathcal{X}(x)} \mathbb{E}[U(X)] \\
&= p\mathbb{E}[U(X_1+Y_1-(1+r)\pi_1)] \\
&\quad + q\mathbb{E}[U(X_0+Y_0-(1+r)\pi_0)] \text{ by def. of } \pi_0, \pi_1 \\
&\le \mathbb{E}[U(pX_1+qX_0+Y_p-(1+r)(p\pi_1+q\pi_0))] \text{ by concavity of } U \\
&\le \max_{X\in\mathcal{X}(x)} \mathbb{E}[U(X+Y_p-(1+r)(p\pi_1+q\pi_0))]
\end{aligned}$$

since $pX_1 + qX_0 \in \mathcal{X}(x)$. Hence $\pi_p \geq p\pi_1 + q\pi_0$. $\square$

2 Marginal utility pricing

Fix Y and let

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

Example sheet: $t \mapsto \pi_t$ decreasing. [Hint: use $\pi(0) = 0$ and concavity]

Let

$$\pi_0 = \sup_{t>0} \pi_t = \lim_{t \downarrow 0} \pi_t.$$

Theorem.

$$\pi_0 = \frac{\mathbb{E}[U'(X_0)Y]}{(1+r)\mathbb{E}[U'(X_0)]}$$

where X_0 is the optimal solution to the problem

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

Interpretation:

An investor's average indifference price for a very small number of shares of a contingent claim is

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

where $\mathbb{Q}$ is the risk-neutral measure whose density is proportional to the marginal utility $U'(X_0)$ of optimised time-1 wealth in a world where there is no contingent claim available to buy.

Proof. Let X_t be the maximiser of $\mathbb{E}[U(X + tY - (1+r)t\pi_t)]$ over $X \in \mathcal{X}(x)$.

$$\begin{aligned}
0 &= \frac{1}{t}\mathbb{E}[U(X_t + tY - (1+r)t\pi_t)] - U(X_0)] \\
&\geq \frac{1}{t}\mathbb{E}[U(X_0 - t\pi_t(1+r) + tY) - U(X_0)] \text{ since } X_t \text{ is maximiser} \\
&\geq \mathbb{E}\left[\frac{U(X_0 - t\pi_0(1+r) + tY) - U(X_0)}{t}\right] \text{ since } \pi_0 \geq \pi_t \\
&\to \mathbb{E}\{U'(X_0)[Y - (1+r)\pi_0]\}
\end{aligned}$$

So

$$\pi_0 \geq \frac{\mathbb{E}[U'(X_0)Y]}{(1+r)\mathbb{E}[U'(X_0)]}$$

For the reverse inequality,

$$\begin{aligned}
0 &= \frac{1}{t}\mathbb{E}[U(X_t + tY - (1+r)t\pi_t)] - U(X_0)] \\
&\leq \frac{1}{t}\mathbb{E}[U'(X_0)(X_t + tY - (1+r)t\pi_t - X_0)] \\
&= \mathbb{E}\{U'(X_0)[Y - (1+r)\pi_t]\}
\end{aligned}$$

So

$$\pi_t \leq \frac{\mathbb{E}[U'(X_0)Y]}{(1+r)\mathbb{E}[U'(X_0)]}$$

for all t. Taking the limit as $t \downarrow 0$ completes the proof. $\square$

Remark. A similar argument works for $t < 0$.
Remark. For the second inequality in the above proof, we have used

Lemma (Supporting hyperplane). *For any x, y we have*

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

Proof. For $x < x + \varepsilon < y$ we have

$$\begin{aligned}
\frac{U(y) - U(x)}{y - x} &\leq \frac{U(x+\varepsilon) - U(x)}{\varepsilon} \\
&\to U'(x) \text{ as } \varepsilon \downarrow 0
\end{aligned}$$

by the definition of concavity with $p = \frac{\varepsilon}{y-x}$. The case $y < x$ is similar. $\square$

Stochastic Financial Models 7

Michael Tehranchi

October 21, 2022

(continued from last time.) For the reverse inequality, we will use

Lemma (Supporting hyperplane). *For any x, y we have*

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

Proof. For $x < x + \varepsilon < y$ we have

$$\begin{aligned}\frac{U(y)-U(x)}{y-x} &\leq \frac{U(x+\varepsilon)-U(x)}{\varepsilon} \\ &\to U'(x) \text{ as } \varepsilon \downarrow 0\end{aligned}$$

by the definition of concavity with $p = \frac{\varepsilon}{y-x}$. The case $y < x$ is similar. $\square$

Proof of upper bound.

$$\begin{aligned}0 &= \frac{1}{t}\mathbb{E}[U(X_t + tY - (1+r)t\pi_t)] - U(X_0)] \\ &\leq \frac{1}{t}\mathbb{E}[U'(X_0)(X_t + tY - (1+r)t\pi_t - X_0)] \\ &= \mathbb{E}\{U'(X_0)[Y - (1+r)\pi_t]\}\end{aligned}$$

where we have used the fact that $\mathbb{E}[U'(X_0)X] = (1+r)x$ for any $X \in \mathcal{X}(x)$ and in particular $\mathbb{E}[U'(X_0)(X_t - X_0)] = 0$. So

$$\pi_t \leq \frac{\mathbb{E}[U'(X_0)Y]}{(1+r)\mathbb{E}[U'(X_0)]}$$

for all t. Taking the limit as $t \downarrow 0$ completes the proof.

A similar argument establishes the case $t < 0$. $\square$

1 Arbitrage

Recall the set-up

- one risk-free asset with interest rate r

- d risky assets with time-t price S_t.

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 $X_0 = x$ and increasing utility function U, consider the problem

$$\text{maximise } \mathbb{E}[U(X_1)] \text{ over } X_1 \in \mathcal{X}(x)$$

where

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

- Suppose φ is an arbitrage.
- Given $X \in \mathcal{X}(x)$ consider

$$X^* = X + \varphi^\top [S_1 - (1+r)S_0]$$

- Note $X^* \in \mathcal{X}(x)$ 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 was arbitrary, there cannot be a maximiser!

Why arbitrages are bad for theory

- Suppose φ is an arbitrage.
- From above, the portfolio $\theta + \varphi$ is better than θ.
- But $\theta + (n+1)\varphi$ is better than $\theta + n\varphi$.
- 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 or utility function U.
- However, it does depend on the agent's *beliefs* through the probability measure $\mathbb{P}$.
- Nevertheless, if one agent believes φ is an arbitrage, then another agent with *equivalent* beliefs (i.e. they agree on the almost sure events) would also believe that φ is an arbitrage.

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. $\square$

Stochastic Financial Models 8

Michael Tehranchi

October 24, 2022

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 Bolzono–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. 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

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

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

□

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.

1 No-arbitrage pricing

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

- Given U and x, find the indifference price.
- Given U and x, 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

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

where $0 \leq p = 1 - q \leq 1$. Note that $\mathbb{Q}_p$ has the same sets of measure zero as $\mathbb{Q}_0$ and $\mathbb{Q}_1$. 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. Let

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

so π_p is an no-arbitrage price for the claim. But

$$\pi_p = p\pi_1 + (1-p)\pi_0$$

so the set of no-arbitrage prices is an interval. □

Remark. Note that the marginal utility price of a claim

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

is also a no-arbitrage price since $U'(X_1^*)$ is proportional to the density of a risk-neutral measure. Also, since $\pi(tY)/t$ is decreasing, we have $\pi(Y) \leq \pi_0$. So, in general, all we can say is

$\pi(Y) \leq \sup\{\pi :$ the market augmented has no arbitrage, where π is the initial price of the claim$\}$

Stochastic Financial Models 9

Michael Tehranchi

October 26, 2022

1 Attainable claims

If one starts with θ^0 in the bank and a portfolio θ of risky assets, the value of these assets is $\theta^0(1+r)+\theta^\top S_1$.

Definition. A contingent claim with payout Y is *attainable* iff $Y=\theta^0(1+r)+\theta^\top S_1$ for some $\theta^0\in\mathbb{R}$ and $\theta\in\mathbb{R}^d$.

Remark. We can equivalently write

$$\theta^0(1+r)+\theta^\top S_1=(1+r)y+\theta^\top[S_1-(1+r)S_0]$$

with

$$\theta^0=y-\theta^\top S_0.$$

- Attainable claims have indifference prices independent of U and x (example sheet)
- Attainable claims have marginal utility prices independent of U and x
- 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.*

'Primal' proof. Suppose

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

Claim: the unique no arbitrage price is $\pi=y$.

First, suppose $\pi=y$. Let $(\varphi^\top,\phi)^\top$ be a candidate arbitrage:

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

This means

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

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 $\pi = y$ is a no-arbitrage price.

Now suppose $\pi > y$. Consider the portfolio $(\theta^\top, -1)$ and note

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

This is an arbitrage. Otherwise, if $\pi < x$ the portfolio $(-\theta^\top, +1)$ is an arbitrage. Hence there is exactly one price such that the augmented market has no arbitrage. □

'Dual' proof. Suppose the augmented market has no arbitrage and such that the claim has initial price π. By the FTAP, there exists a risk neutral measure $\mathbb{Q}$ such that

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

But

$$\begin{aligned}\frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}(Y) &= \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}((1+r)\theta^0 + \theta^\top S_1)\\ &= \theta^0 + \theta^\top S_0\end{aligned}$$

so there is only one possible value for π. □

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. □

Example 1. 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. Consider the one-period binomial model from lecture 5. A claim with payout $Y = g(S_1)$ is attainable for any function g. Indeed, we need only check that there is a solution (θ^0, θ) to the system of equation

$$\begin{aligned}(1+r)\theta^0 + \theta S_0(1+b) &= g(S_0(1+b))\\ (1+r)\theta^0 + \theta S_0(1+a) &= g(S_0(1+a))\end{aligned}$$

Of course, the solution is

$$\begin{aligned}\theta &= \frac{g(S_0(1+b)) - g(S_0(1+a))}{S_0(b-a)}\\ \theta^0 &= \frac{(1+b)g(S_0(1+a)) - (1+a)g(S_0(1+b))}{(1+r)(b-a)}\end{aligned}$$

Check that

$$\theta^0 + \theta S_0 = \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}[g(S_1)]$$

where $\mathbb{Q}$ is the unique risk-neutral measure.

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.

Stochastic Financial Models 10

Michael Tehranchi

October 28, 2022

1 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 $\{X \leq x\}$ 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. *A random variable Y is measurable respect to $\sigma(X)$ if and only if there is a 'nice' function f such that $Y = f(X)$.*

2 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 bounded $\sigma(Y)$-measurable random variable Z we have

$$\mathbb{E}[XZ] = \mathbb{E}[\mathbb{E}[X|Y]Z].$$

Proof of the projection property: By measurability there exists a bounded function g such that $Z = g(Y)$. By the law of total probability

$$\begin{aligned}
\mathbb{E}[f(Y)g(Y)] &= \sum_i \mathbb{P}(Y = y_i)\mathbb{E}(X|Y = y_i)g(y_i) \\
&= \sum_i \mathbb{E}(X\mathbb{1}_{\{Y=y_i\}})g(y_i) \\
&= \mathbb{E}[Xg(Y)]
\end{aligned}$$

since

$$\sum_i \mathbb{1}_{\{Y=y_i\}}g(y_i) = g(Y)$$

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

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

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

Proposition (Uniqueness of conditional expectations)**.** *Let $\tilde{X}_0$ and $\tilde{X}_1$ be two versions of the conditional expectation of X given $\mathcal{G}$. Then*

$$\tilde{X}_0 = \tilde{X}_1 \textit{ almost surely}$$

Proof. For all $G \in \mathcal{G}$ we have

$$\mathbb{E}[\tilde{X}_0 \mathbb{1}_G] = \mathbb{E}[X \mathbb{1}_G] = \mathbb{E}[\tilde{X}_1 \mathbb{1}_G]$$

so that $\mathbb{E}[(\tilde{X}_1 - \tilde{X}_0)\mathbb{1}_G] = 0$. Note $G = \{\tilde{X}_0 < \tilde{X}_1\}$ is in $\mathcal{G}$ since $\tilde{X}_1$ and $\tilde{X}_0$ are both $\mathcal{G}$-measurable by definition. Since $(\tilde{X}_1 - \tilde{X}_0)\mathbb{1}_G \geq 0$, by the pigeon-hole principle we have $(\tilde{X}_1 - \tilde{X}_0)\mathbb{1}_G \geq 0 = 0$ meaning $\tilde{X}_1 - \tilde{X}_0 \geq 0$ almost surely. By symmetry $\tilde{X}_1 - \tilde{X}_0 \leq 0$ almost surely as well. □

Proposition (Existence of conditional expectations)**.** *If X is integrable, the conditional expectation of X given $\mathcal{G}$ exists.*

Notation: The conditional expectation of X given $\mathcal{G}$ is denoted $\mathbb{E}(X|\mathcal{G})$. In the special case where $\mathcal{G} = \sigma(Y)$ we write $\mathbb{E}(X|Y)$ for $\mathbb{E}(X|\sigma(Y))$.
Projection property Let X be integrable and Z bounded and $\mathcal{G}$-measurable. Then

$$\mathbb{E}[XZ] = \mathbb{E}[\mathbb{E}(X|\mathcal{G})Z]$$

The proof follows from the definition of $\mathbb{E}(X|\mathcal{G})$ and standard measure theory.

Proposition (Mean squared error minimisation)**.** *Suppose X is square-integrable and let $\tilde{X} = \mathbb{E}(X|\mathcal{G})$. Then*

$$\mathbb{E}[(X - \tilde{X})^2] \leq \mathbb{E}[(X - Y)^2]$$

for all $\mathcal{G}$-measurable square integrable Y.

Sketch of proof. Let $Z = \tilde{X} - Y$. Then

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

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

Stochastic Financial Models 11

Michael Tehranchi

October 31, 2022

1 Properties of conditional expectations

Example. Suppose $\Omega = \bigcup_n G_n$, where the G_n are disjoint and $\mathbb{P}(G_n) > 0$. Let $\mathcal{G}$ be the sigma-algebra $\{\bigcup_{n\in I} G_n : I \subseteq \mathbb{N}\}$ and X integrable. Then

$$\mathbb{E}(X|\mathcal{G})(\omega) = \mathbb{E}(X|G_n) \text{ iff } \omega \in G_n$$

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* $\mathbb{E}[f(X)|\mathcal{G}] \geq f[\mathbb{E}(X|\mathcal{G})]$

Proof. The first four properties follow directly from the definition and uniqueness of conditional expectation.

For positivity, apply the definition of conditional expectation with $G = \{\mathbb{E}(X|\mathcal{G}) < 0\}$.

For Jensen's inequality in the case where f has a bounded derivative, from the supporting hyperplane theorem we have

$$f(y) \geq f(x) + f'(x)(y - x)$$

for all x, y. Now, let $x = \mathbb{E}(X|\mathcal{G})$ and $y = X$ and compute the conditional expectation of both sides using additivity and positivity. The conclusion follows from pulling out a known factor since $f'(\mathbb{E}(X|\mathcal{G}))$ is $\mathcal{G}$-measurable. □

Definition. The conditional probability of A given $\mathcal{G}$ is

$$\mathbb{P}(A|\mathcal{G}) = \mathbb{E}(\mathbb{1}_A|\mathcal{G})$$

Remark. Returning to our motivating discussion of what should be called a measurable set: Note if A is $\mathcal{G}$-measurable then $\mathbb{P}(A|\mathcal{G}) = \mathbb{1}_A$ which only takes the values $\{0,1\}$. Conversely, suppose $\mathbb{P}(A|\mathcal{G})$ only takes the values $\{0,1\}$. Then there is a $\mathcal{G}$-measurable set B such that $\mathbb{P}(A|\mathcal{G}) = \mathbb{1}_B$. Note that

$$\mathbb{P}(A) = \mathbb{E}[\mathbb{1}_A] = \mathbb{E}[\mathbb{1}_B] = \mathbb{P}(B)$$

and

$$\mathbb{P}(A \cap B) = \mathbb{E}[\mathbb{1}_A \mathbb{1}_B] = \mathbb{E}[\mathbb{1}_B] = \mathbb{P}(B)$$

and hence

$$\mathbb{E}[(\mathbb{1}_A - \mathbb{1}_B)^2] = \mathbb{P}(A) - 2\mathbb{P}(A \cap B) + \mathbb{P}(B) = 0$$

so $\mathbb{1}_A = \mathbb{1}_B$ a.s. So A is *almost* a $\mathcal{G}$-measurable set.

2 Filtrations, adaptedness and martingales

Definition. A *filtration* is a family $(\mathcal{F}_n)_{n\geq 0}$ of sigma-algebras such that $\mathcal{F}_{n-1} \subseteq \mathcal{F}_n$ for all $n \geq 1$.

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

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

Definition. A stochastic process $(X_n)_{n\geq 0}$ is *adapted* to a filtration $(\mathcal{F}_n)_{n\geq 0}$ iff X_n is $\mathcal{F}_n$ measurable for all $n \geq 0$.

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

Definition. The filtration $(\mathcal{F}_n)_{n\geq 0}$ *generated* by a process $(X_n)_{n\geq 0}$ is

$$\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$$

(i.e. the smallest filtration for which the process is adapted)

Definition. A discrete-time process $(X_n)_{n\geq 0}$ is a *martingale* with respect to a filtration $(\mathcal{F}_n)_{n\geq 0}$ iff X_n is integrable for all n and

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

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

$$\mathbb{E}(X_n - X_{n-1}|\mathcal{F}_{n-1}) = 0 \text{ for all } n \geq 1$$

Stochastic Financial Models 12

Michael Tehranchi

November 5, 2022

1 Motivation from finance

We will soon look at multi-period models with risky asset prices $(S_n)_{n\geq 0}$.

Definition. A risk-neutral measure is a measure equivalent to $\mathbb{P}$ such that $\frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}(S_n|\mathcal{F}_{n-1}) = S_{n-1}$ for all $n \geq 1$.

Remark. An equivalent measure $\mathbb{Q}$ is risk-neutral iff the process $(M_n)_{n\geq 0}$ is a d-dimensional martingale under $\mathbb{Q}$ where

$$M_n = (1+r)^{-n} S_n.$$

Theorem (Fundamental theorem of asset pricing)**.** *In a finite horizon multi-period model, there is no arbitrage if and only if there exists a risk-neutral measure.*

2 Examples of martingales

A martingale is defined relative to a filtration. Sometimes we need a way to specify the filtration. Here is useful way:

Definition. The filtration $(\mathcal{F}_n)_{n\geq 0}$ *generated* by a process $(X_n)_{n\geq 0}$ is

$$\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$$

(i.e. the smallest filtration for which the process is adapted)

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
Example.

- Given a filtration $(\mathcal{F}_n)_{n\geq 0}$ and an integrable random variable X.
- Let $Z_n = \mathbb{E}(X|\mathcal{F}_n)$ for $n \geq 0$.

Then $(Z_n)_{n\geq 0}$ is a martingale.

- That Z_n is integrable is from the definition of conditional expectation.
- and $\mathbb{E}(Z_n|\mathcal{F}_{n-1}) = \mathbb{E}[\mathbb{E}(X|\mathcal{F}_n)|\mathcal{F}_{n-1}] = \mathbb{E}(X|\mathcal{F}_{n-1}) = Z_{n-1}$ by the tower property.

3 Martingale transform

Definition. A 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 adepted 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

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

for some $C > 0$, by the triangle inequality and the assumption the boundedness of (H_n) and the integrability of (X_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 $(X_n)_{n\geq 0}$. □

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)

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

Proof. The investor's wealth is $X_n = \theta_n^0 + \theta_n^\top S_n$. The self-financing condition is

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

Solving for θ_n^0 and substituting

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

Hence

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

□

Stochastic Financial Models 13

Michael Tehranchi

November 5, 2022

1 Stopping times

Definition. A *stopping time* for a filtration $(\mathcal{F}_n)_{n\geq 0}$ is a random variable T valued in $\{0, 1, 2, \dots, +\infty\}$ such that

$$\{T \leq n\} \in \mathcal{F}_n \text{ for all } n \geq 0$$

Example.

- Let $(X_n)_{n\geq 0}$ be an 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.)

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.

Proposition. *A random time T is a stopping time if and only if*

$$\{T = n\} \in \mathcal{F}_n \textit{ for all } n \geq 0$$

Proof. Suppose $\{T = n\} \in \mathcal{F}_n$ for all n. Then $\{T \leq n\} = \cup_{k=0}^{n}\{T = k\} \in \mathcal{F}_n$ since $\mathcal{F}_k \subseteq \mathcal{F}_n$ for all $k \leq n$.

Conversely, if T is a stopping time, then $\{T = n\} = \{T \leq n\} \cap \{T \leq n-1\}^c \in \mathcal{F}_n$. $\square$

2 Optional sampling theorem

Definition. Let $(X_n)_{n\geq 0}$ be an adapted process and T a stopping time. The *stopped process* $(X_{n\wedge T})_{n\geq 0}$ is defined by

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

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

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 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. □

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. □

Theorem (Optional sampling theorem)**.** *Let $(X_n)_{n\geq 0}$ be a martingale and T a bounded stopping time (i.e. there exists a constant N such that $T \leq N$ 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. First, if (M_n) is a martingale, then by the tower property

$$\begin{aligned}\mathbb{E}(M_n) &= \mathbb{E}[\mathbb{E}(M_n|\mathcal{F}_{n-1})] \\ &= \mathbb{E}(M_{n-1})\end{aligned}$$

so by induction $\mathbb{E}(M_n) = M_0$ for any non-random time n.

Now, since $M_n = X_{T\wedge n}$ is a martingale and $M_0 = X_0$ and $M_N = X_{T\wedge N} = X_T$, we are done. □

3 Optional stopping theorem

For interesting applications, we need to remember

Theorem (Monotone convergence theorem)**.** *Given a non-negative increasing sequence $(X_n)_n$ of random variables, we have*

$$\lim_n \mathbb{E}(X_n) = \mathbb{E}(\lim_n X_n).$$

Theorem (Dominated convergence theorem)**.** *Let* $(X_n)_n$ *be an almost surely convergent sequence of random variables. If* $\sup_n |X_n|$ *is integrable then we have*

$$\lim_n \mathbb{E}(X_n) = \mathbb{E}(\lim_n X_n).$$

Theorem (Optional stopping theorem)**.** *Let* T *be a stopping time and* $(X_n)_{n\geq 0}$ *be a martingale such that* $\sup_n |X_{n\wedge T}|$ *integrable. Then* $X_T = \lim_n X_{n\wedge T}$ *exists and*

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

Remark. A simple condition for integrability is $|X_{n\wedge T}| \leq C$ for all n, for some constant $C > 0$.

Proof when T is finite almost surely. By optional sampling

$$\mathbb{E}(X_{n\wedge T}) = X_0$$

The conclusion follows from the dominated convergence theorem since $X_{n\wedge T} \to X_T$. □

Example

- 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 $\mathbb{P}(\xi_n = \pm 1) = \frac{1}{2}$.
- 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.

- 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.

Stochastic Financial Models 14

Michael Tehranchi

November 7, 2022

1 Examples of optional stopping

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 independent $\mathbb{P}(\xi_n = \pm 1) = \frac{1}{2}$.
Claim: the process $w^{S_n} z^n$ is a martingale iff $w + w^{-1} = 2z^{-1}$. Indeed, note

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

Example 1. Fix an integer $a > 0$ and let $\tau = \inf\{n \geq 0 : S_n = -a\}$. Our goal is to find the probability generating function $\mathbb{E}(z^\tau)$ for fixed $0 < z < 1$.

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? Well, note that

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

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.

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

Example 2. Fix integers $a, b > 0$ and let $T = \inf\{n \geq 0 : S_n \in \{b, -a\}\}$. Again, our goal is to find the probability generating function $\mathbb{E}(z^T)$ for fixed $0 < z < 1$.

Let $w + w^{-1} = 2z^{-1}$ as before set

$$M_n = w^{S_n + \frac{a-b}{2}} z^n + w^{-S_n + \frac{b-a}{2}} z^n$$

Note $(M_n)_{n\geq 0}$ is a martingale, and

$$M_T = (w^{\frac{a+b}{2}} + w^{-\frac{(a+b)}{2}} z^T$$

so by the OST

$$\mathbb{E}(M_T) = (w^{\frac{a+b}{2}} + w^{-\frac{(a+b)}{2}})\mathbb{E}(z^T) = M_0 = w^{\frac{a-b}{2}} + w^{\frac{b-a}{2}}$$

to yield the formula

$$\mathbb{E}(z^T) = \frac{w^{\frac{a-b}{2}} + w^{\frac{b-a}{2}}}{w^{\frac{a+b}{2}} + w^{-\frac{(a+b)}{2}}} = \frac{w^a + w^b}{1 + w^{a+b}}$$

Notice that the OST is applicable in this case since $M_{n\wedge T}$ is bounded for either choice of w. Furthermore, the choice of w doesn't matter to the final formula since it is symmetric in w and w^{-1}. (Check that we recover the previous formula when $b \to +\infty$.)

2 Submartingales and supermartingales

Definition. A *submartingale* with respect to a filtration $(\mathcal{F}_n)_{n\geq 0}$ is an integrable adapted process $(X_n)_{n\geq 0}$ such that

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

Proposition. *An integrable adapted process $(X_n)_{n\geq 0}$ is a submartingale iff and only if*

$$\mathbb{E}(X_n|\mathcal{F}_m) \leq X_m \textit{ for all } 0 \leq m \leq n$$

Proof. The 'if' direction follows from the given definition. To prove the 'only if' direction: Fix m. The claim is true for $n = m + 1$. Suppose it is true for $n = m + k$ for some $k \geq 1$. Then

$$\begin{aligned}
\mathbb{E}(X_{m+k+1}|\mathcal{F}_m) &= \mathbb{E}[\mathbb{E}(X_{m+k+1}|\mathcal{F}_{m+k})|\mathcal{F}_m] \\
&\geq \mathbb{E}[X_{m+k}|\mathcal{F}_m] \\
&\geq X_m
\end{aligned}$$

using the assumed submartingale property of $(X_n)_n$, the positivity of conditional expectation and the induction hypothesis. The claim follows by induction. □

Definition. A *supermartingale* with respect to a filtration $(\mathcal{F}_n)_{n\geq 0}$ is an integrable adapted process $(X_n)_{n\geq 0}$ such that

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

Remark. A process $(X_n)_n$ is a supermartingale iff $(-X_n)_n$ is a submartingale.

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 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 N such that $T \leq N$ a.s. The conclusion follows from $\mathbb{E}(M_N) \geq M_0 = 0$ since $M_N = X_T - X_S$. □

Stochastic Financial Models 15

Michael Tehranchi

November 10, 2022

1 Controlled Markov processes

Markov process

Definition. A process $(X_n)_{n\geq 0}$ has the *Markov property* iff

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

for all $n \geq 1$, (measurable) sets A, and where $(\mathcal{F}_n)_{n\geq 0}$ is the filtration generated by $(X_n)_{n\geq 0}$.
A *Markov process* is a process with the Markov property.

There are two ways to think about Markov processes:
Transition probabilities. A Markov process can be built from

- Initial condition $X_0 = x$
- The one-step transition probabilities $P(n, x, A) = \mathbb{P}(X_n \in A|X_{n-1} = x)$ for $n \geq 1$ and measurable sets $A \subseteq \mathcal{X}$.

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 ξ, such that

  $$\mathbb{P}(G(n, x, \xi_n) \in A) = P(n, x, A)$$

  for all A.
- 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
$$G(n, x, v) = x + v$$
- 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 $n \geq 1$.

Example from finance. Let $d = 1$.

- Suppose that the risky asset prices $(S_n)_{n\geq 0}$ are such that $S_n = S_{n-1}\xi_n$, where $(\xi_n)_{n\geq 1}$ are IID. Then $(S_n)_{n\geq 0}$ is a Markov process.
- Given the initial wealth $X_0 = x$
- Given the previsible trading strategy $(\theta_n)_{n\geq 1}$
- Let
$$\begin{aligned} X_n &= (1+r)X_{n-1} + \theta_n[S_n - (1+r)S_{n-1}] \\ &= (1+r)X_{n-1} + \theta_n S_{n-1}[\xi_n - (1+r)] \end{aligned}$$
- The investor's wealth is a controlled Markov process with previsible control $U_n = \theta_n S_{n-1}$.

2 Bellman equation

Stochastic optimal control.

Given a controlled Markov process $(X_n^U)_{n\geq 0}$ and a (non-random) time horizon N we wish to

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

over previsible controls $(U_k)_{1\leq k\leq N}$.

The value function

Idea: consider a sequence of sub-problems.

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

where the supremum is over previsible controls $(U_k)_{n+1\leq n\leq N}$.

The function V is called the *value function* for the problem.

Suppose the controlled Markov process evolves as $X_n = G(n, X_{n-1}, U_n, \xi_n)$ for $n \geq 1$.

Definition. The system of equations

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

is called the *Bellman equation* for the problem.

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

Stochastic Financial Models 16

Michael Tehranchi

November 11, 2022

1 Dynamic programming principle

Given

- A sequence $(\xi_n)_{n\geq 1}$ of independent random variables generating a filtration $(\mathcal{F}_n)_{n\geq 0}$
- A function $G(\cdot,\cdot,\cdot,\cdot)$
- An initial condition X_0

for any previsible $(U_n)_{n\geq 1}$ construct the controlled Markov process by

$$X_n^U = \begin{cases} X_0 & \text{if } n = 0 \\ G(n, X_{n-1}^U, U_n, \xi_n) & \text{if } n \geq 1 \end{cases}$$

Now given

- A non-random time horizon $N > 0$
- Suitably integrable functions $f(\cdot,\cdot)$ and $g(\cdot)$

we seek to maximise

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

Theorem (The dynamic programming principle)**.** *Let V solve the Bellman equations:*

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

and suppose for each n and x there is an optimal solution $u^*(n,x)$ to the maximisation problem, so that

$$V(n-1,x) = f(n, u^*(n,x)) + \mathbb{E}[V(n, G(n,x,u^*(n,x),\xi_n))] \ \textit{for all } 1 \leq n \leq N, \ \textit{and } x$$

Fix an initial condition $X_0^* = X_0$ and let

$$U_n^* = u^*(n, X_{n-1}^*),$$
$$X_n^* = G(n, X_{n-1}^*, U_n^*, \xi_n) \text{ for all } 1 \le n \le N$$

so that $X_n^{U^*} = X_n^*$ for all $0 \le n \le N$. Then $(U_n^*)_{1\le n\le N}$ is the optimal control and V is the value function.

For the proof we need a general fact:

Lemma. *Given a sigm-algebra $\mathcal{G}$ and two random variables X and Y such that X is $\mathcal{G}$-measurable and Y is independent of $\mathcal{G}$. For any suitably integrable function $h(\cdot,\cdot)$ we have*

$$\mathbb{E}[h(X,Y)|\mathcal{G}] = \mathbb{E}[h(x,Y)]\big|_{x=X}$$

Remark. An alternative notation for the conclusion of the lemma is

$$\mathbb{E}[h(X,Y)|\mathcal{G}](\omega) = \mathbb{E}[h(X(\omega),Y)] \text{ for almost all } \omega$$

Idea of proof of the lemma. The conclusion is easy to check via the definition of conditional expectation in the case where h factorises as $h(x,y) = \phi(x)\psi(y)$. The general case follows from the monotone class theorem from Probability & Measure. □

Proof of the dynamic programming principle. Fix X_0 and let $(U_n)_{1\le n\le N}$ be a previsible control, and consider the associated controlled process $(X_n^U)_{0\le n\le N}$. Let

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

Claim: $(M_n)_{0\le n\le N}$ is a supermartigale.

Indeed, this process is adapted and integrable and using the lemma we have

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

since V solves the Bellman equation.

Hence, using $V(N,x) = g(x)$ for all x, we have

$$\begin{aligned}
\mathbb{E}\left[\sum_{k=1}^{N} f(k,U_k) + g(X_N^U)|\mathcal{F}_n\right] &= \mathbb{E}[M_N|\mathcal{F}_n] \\
&\le M_n \\
&= V(n, X_n) + \sum_{k=1}^{n} f(k, U_k)
\end{aligned}$$

By the tower property

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

since $\sigma(X_n^U) \subseteq \mathcal{F}_n$.

Finally note that if $U = U^*$, then the process M is a martingale, and there is equality in each inequality above, and hence

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

as claimed. □

2 Optimal investment

Pure investment.

Given a market with interest rate r and d risky assets with prices $(S_n)_{n\geq 0}$

- wealth evolves as $X_n = (1+r)X_{n-1} + \theta_n^\top[S_n - (1+r)S_{n-1}]$
- Given a time horizon N, a natural goal is to

  $$\text{maximise } \mathbb{E}[U(X_N)]$$

  where U is the investor's utility funciton.

Let $d = 1$.

- Suppose $S_n = S_{n-1}\xi_n$, where $(\xi_n)_{n\geq 1}$ are IID.
- Let

  $$X_n = (1+r)X_{n-1} + \eta_n[\xi_n - (1+r)]$$

  where $\eta_n = \theta_n S_{n-1}$.

The Bellman equation is

$$\begin{aligned} V(N,x) &= U(x) \\ V(n-1,x) &= \max_\eta \mathbb{E}[V(n,(1+r)x + \eta(\xi - (1+r))] \end{aligned}$$

- Generally, intractable
- but suppose the utility is CARA: $U(x) = -e^{-\gamma x}$

- Guess: $V(n,x) = U[x(1+r)^{N-n}]A_n$
- Check: Assuming correct

$$\begin{aligned}
&\max_\eta \mathbb{E}[V(n,(1+r)x + \eta(\xi - (1+r))] \\
&= U[x(1+r)^{N-n+1}]A_n \min_\eta \mathbb{E}[e^{-\gamma(1+r)^{N-n}\eta\xi}]e^{\gamma(1+r)^{N-n}\eta(1+r)} \\
&= U[x(1+r)^{N-n+1}]A_n\alpha
\end{aligned}$$

  where $\alpha = \min_t \mathbb{E}[e^{-t\xi}]e^{t(1+r)}$
- $A_n = \alpha^{N-n}$
- Optimal strategy

$$\theta_n^* = \frac{\eta_n^*}{S_{n-1}} = \frac{t^*}{\gamma(1+r)^{N-n}S_{n-1}}$$

  where $t^* = \text{argmin}_t \mathbb{E}[e^{-t\xi}]e^{t(1+r)}$

Stochastic Financial Models 17

Michael Tehranchi

November 14, 2022

1 Another example of optimal investment

This time with consumption.

- Same market model set-up as before.
- Now we allow the investor to consume C_n during the interval $(n-1, n]$
- wealth evolves as $X_n = (1+r)(X_{n-1} - C_n) + \theta_n[S_n - (1+r)S_{n-1}]$
- Given a time horizon N, a natural goal is to

$$\text{maximise } \mathbb{E}\left[\sum_{k=1}^{N} U(C_k) + U(X_N)\right]$$

where U is the investor's utility function

The Bellman equation is

$$\begin{aligned} V(N, x) &= U(x) \\ V(n-1, x) &= \max_{c,\eta} \mathbb{E}\left[U(c) + V(n, (1+r)(x-c) + \eta(\xi - (1+r)))\right] \end{aligned}$$

- Generally, intractable
- but suppose the utility is CRRA: $U(x) = \frac{1}{1-R}x^{1-R}$ for $x > 0$, where $R > 0, R \neq 1$.
- Guess: $V(n, x) = U(x)A_n$
- Check: Correct for $n = N$. Assume correct for $n = k$ for some $k \leq N$. The

$$\begin{aligned} V(k-1, x) & \max_{c,\eta}\{U(c) + \mathbb{E}[V(k, (1+r)(x-c) + \eta(\xi - (1+r))]\} \\ &= x^{1-R} \max_{c,\eta} \Big\{U(c/x) \\ &+ A_k(1-c/x)^{1-R}\mathbb{E}\left[U\left((1+r) + \frac{\eta}{x-c}(\xi - (1+r))\right)\right] \Big\} \\ &= x^{1-R} \max_{c} \{U(c/x) + A_k U(1-c/x)\alpha\} \end{aligned}$$

where $\alpha = (1-R)\max_t \mathbb{E}\left[U\left((1+r) + t(\xi - (1+r))\right)\right]$

- Now optimise over $s = c/x$: differentiate and set equal to zero to get

$$s_k^{-R} = (1 - s_k)^{-R} A_k \alpha \Rightarrow s_k = \frac{1}{1 + (A_k \alpha)^{1/R}}$$

- plug this back in

$$V(k-1, x) = U(x)(1 + (A_k \alpha)^{1/R})^R$$

- So we would be done if

$$A_{k-1} = (1 + (A_k \alpha)^{1/R})^R \text{ for all } k \leq N$$

- Solving this recursion[1] with $A_N = 1$ yields

$$A_n = \left(\frac{1 - \alpha^{(N-n+1)/R}}{1 - \alpha^{1/R}} \right)^R \quad \text{for } 0 \leq n \leq N$$

- Optimal strategy

$$C_n^* = X_{n-1}^* s_n = \frac{X_{n-1}^*}{1 + (A_n \alpha)^{1/R}} = \frac{X_{n-1}^*(1 - \alpha^{1/R})}{1 - \alpha^{(N-n+2)/R}}$$
$$\theta_n^* = \frac{\eta_n^*}{S_{n-1}} = \frac{t^*(X_{n-1}^* - C_n^*)}{S_{n-1}}$$

where $t^* = \text{argmax}_t \mathbb{E}\left[U\left((1+r) + t(\xi - (1+r))\right)\right]$

2 Infinite-horizon problems

- Consider a controlled Markov process

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

where $(U_n)_{n \geq 1}$ is the previsible control.

- Problem:

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

where $0 < \beta < 1$

- The value function is

$$V(x) = \mathbb{E}\left[\sum_{k=1}^{\infty} \beta^{k-1} f(U_k) | X_0 = x\right]$$

[1] Let $c_k = A_k^{1/r}$ so that $c_{k-1} = 1 + \alpha^{1/R} c_k$ with $c_N = 1$. Then note that this implies $c_k = 1 + \alpha^{1/R} + \cdots + \alpha^{(N-k)/R}$

- The Bellman equation is

$$V(x) = \max_u \{f(u) + \beta \mathbb{E}[V(G(x,u,\xi))]\}$$

- When is the solution of the Bellman equation the value function?

Theorem. *Suppose $f(u) \geq 0$ for all u and that V is a non-negative solution to the Bellman equation. Suppose $u^*(x)$ is the maximiser of*

$$f(u) + \beta \mathbb{E}[V(G(x,u,\xi))]$$

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

$$\beta^n \mathbb{E}[V(X_n^*)] \to 0$$

then V is the value function and U^* is the optimal control.

Proof. Given a control $(U_n)_{n\geq 1}$ let

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

Note $(M_n)_{n\geq 0}$ is a supermartingale

$$\begin{aligned} \mathbb{E}[M_n - M_{n-1} | \mathcal{F}_{n-1}] &= \beta^{n-1}\left(f(U_n) + \beta \mathbb{E}[V(X_n)|\mathcal{F}_{n-1}] - V(X_{n-1})\right) \\ &\leq 0 \end{aligned}$$

with equality if $U = U^*$. Hence

$$\begin{aligned} V(x) &= M_0 \\ &\geq \mathbb{E}[M_n] \\ &= \mathbb{E}\left[\sum_{k=1}^{n} \beta^{k-1} f(U_k)\right] + \beta^n \mathbb{E}[V(X_n)] \end{aligned}$$

with equality if $U = U^*$.

Now since $V \geq 0$, we have

$$\begin{aligned} V(x) &\geq \mathbb{E}\left[\sum_{k=1}^{n} \beta^{k-1} f(U_k)\right] \\ &\to \mathbb{E}\left[\sum_{k=1}^{\infty} \beta^{k-1} f(U_k)\right] \end{aligned}$$

for any control, where we have used that $f \geq 0$ and the monotone convergence theorem.

And for $U = U^*$ we have

$$\begin{aligned} V(x) &= \mathbb{E}\left[\sum_{k=1}^{n} \beta^{k-1} f(U_k^*)\right] + \beta^n \mathbb{E}[V(X_n^*)] \\ &\to \mathbb{E}\left[\sum_{k=1}^{\infty} \beta^{k-1} f(U_k^*)\right] \end{aligned}$$

since $\beta^n \mathbb{E}[V(X_n^*)] \to 0$ by assumption. $\square$

Stochastic Financial Models 18

Michael Tehranchi

November 16, 2022

1 Optimal stopping problems

- Consider a Markov process of the form

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

where $(\xi)_{n\geq 1}$ are independent.

- Fix a horizon N and consider the problem:

$$\text{maximise } \mathbb{E}\left[g(X_T)|X_0 = x\right]$$

over stopping times $0 \leq T \leq N$.

- **The Bellman equation** is

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

Theorem. *Let*

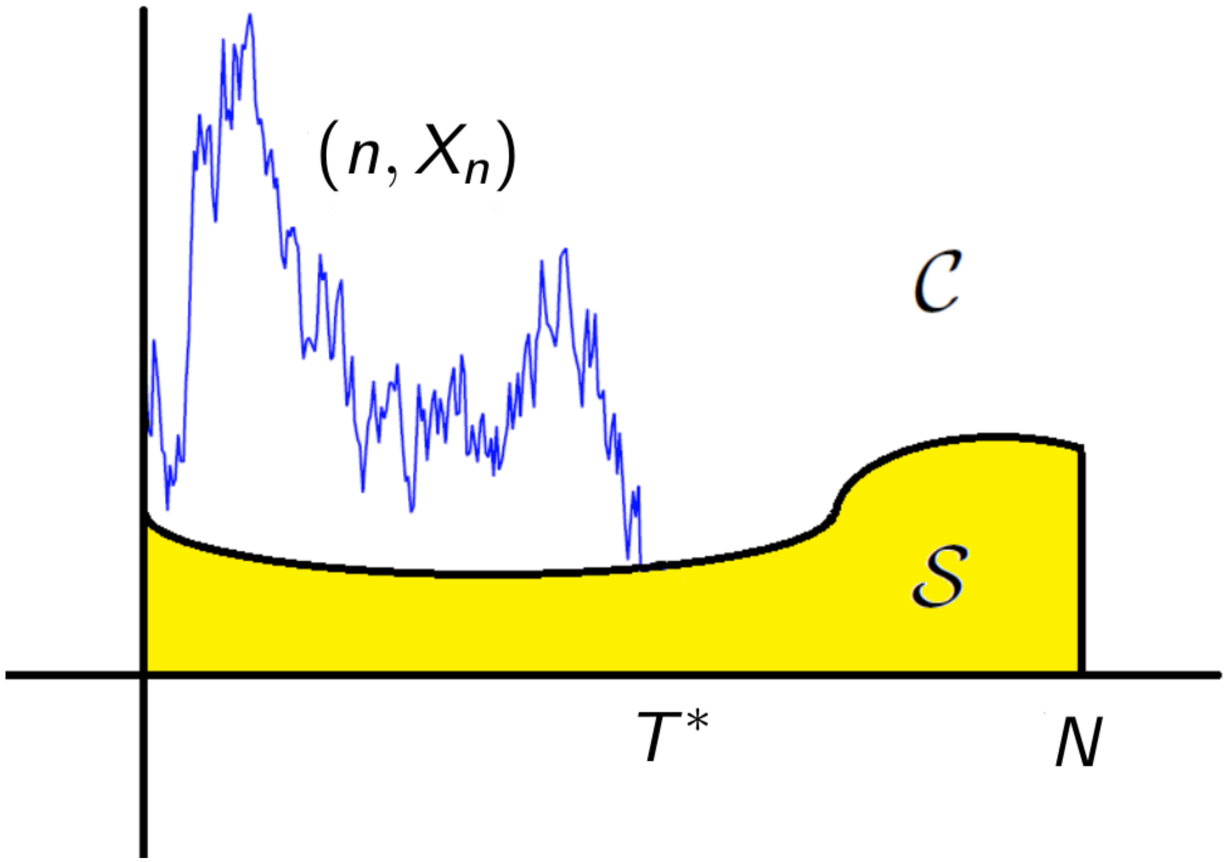
$$T^* = \inf\{n \geq 0 : V(n, X_n) = g(X_n)\}$$

Then T^* is optimal.

Let

$$\begin{aligned} \mathcal{C} &= \{(n,x) : V(n,x) > g(x)\} \\ \mathcal{S} &= \{(n,x) : V(n,x) = g(x)\} \end{aligned}$$

$$T^* = \inf\{n \geq 0 : (n, X_n) \in \mathcal{S}\}$$



Proof. Let

$$M_n = V(n, X_n)$$

First note that $(M_n)_{0 \le n \le N}$ is a supermartingale. Indeed,

$$\begin{aligned}\mathbb{E}[M_n|\mathcal{F}_{n-1}] &= \mathbb{E}[V(n, G(n, X_{n-1}, \xi_n)|X_{n-1}] \\ &\le V(n-1, X_{n-1}) \\ &= M_{n-1}\end{aligned}$$

by the Bellman equation.

Also note

$$M_n \ge g(X_n)$$

for all n.

By the optional sampling theorem

$$\mathbb{E}[g(X_T)] \le \mathbb{E}[M_T] \le M_0 = V(0, x)$$

for any stopping time $T \le N$.

Now, note $(M_{n\wedge T^*})_{0\leq n\leq N}$ is a martingale. Indeed, on the event $\{n \leq T^*\}$ we have $(n-1, X_{n-1}) \in \mathcal{C}$ and hence

$$\mathbb{E}[V(n, G(n, X_{n-1}, \xi_n)|X_{n-1}] = V(n-1, X_{n-1})$$

by the definition of T^*.

Also

$$M_{T^*} = g(X_{T^*}).$$

By the optional sampling theorem

$$\mathbb{E}[g(X_{T^*})] = \mathbb{E}[M_{T^*}] = M_0 = V(0, x)$$

so the upper bound is achieved. □

2 Multi-period arbitrage

The set-up. Consider a market

- with a risk-free asset with interest rate r
- and d risky assets with time n prices $(S_n)_{n\geq 0}$.
- The investor holds the portfolio $\theta_n \in \mathbb{R}^d$ of risky assets during the time interval $(n-1, n]$, where θ_n is $\mathcal{F}_{n-1}$-measurable

The wealth of a self-financing investor evolves as

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

Hence

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

The investor holds

$$\theta_n^0 = X_{n-1} - \theta_n^\top S_{n-1}$$

in the bank during the time interval $(n-1, n]$.

Definition. An *arbitrage* is a previsible process $(\varphi_n)_{1\leq n\leq N}$ such that

$$\sum_{k=1}^{N} (1+r)^{N-k} \varphi_k^\top [S_k - (1+r)S_{k-1}] \geq 0 \text{ almost surely}$$

and

$$\mathbb{P}\left(\sum_{k=1}^{N} (1+r)^{N-k} \varphi_k^\top [S_k - (1+r)S_{k-1}] > 0 \right) > 0$$

If φ is an arbitrage, then an investor would always prefer the investment strategy $\theta + \varphi$ to the strategy θ.

Definition. A risk-neutral measure is a measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ under which the discounted asset price process

$$M_n = (1+r)^{-n} S_n$$

is a martingale, that is,

$$\frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}(S_n|\mathcal{F}_{n-1}) = S_{n-1}$$

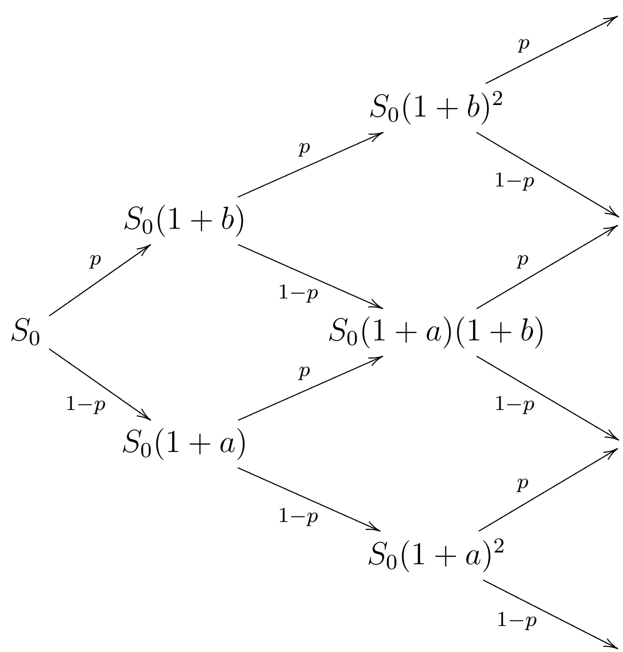
for all $n \geq 1$.

Theorem (Fundamental theorem of asset pricing)**.** *In a finite horizon multi-period model, there is no arbitrage if and only if there exists a risk-neutral measure.*

3 Introduction to the binomial

The Cox–Ross–Rubinstein binomial model

- $d = 1$ and $S_n = S_{n-1}\xi_n$
- $(\xi_n)_{n\geq 1}$ are IID copies of ξ
- $\mathbb{P}(\xi = 1+b) = p = 1 - \mathbb{P}(\xi = 1+a)$
- $S_0 > 0$ and $-1 < a < b$



Theorem. *Consider the N-step binomial model. There exists a risk-neutral measure if and only if $a < r < b$. When it exists it is the unique measure $\mathbb{Q}$ such that $(\xi_n)_{1 \le n \le N}$ are IID under $\mathbb{Q}$ with*

$$\mathbb{Q}(\xi = 1 + b) = q = \frac{r - a}{b - a} = 1 - \mathbb{Q}(\xi = 1 + a).$$

Proof. Suppose such a risk-neutral measure $\mathbb{Q}$ exists. Then by definition

$$\begin{aligned}(1 + r)S_{n-1} =& \mathbb{E}^{\mathbb{Q}}(S_n | \mathcal{F}_{n-1}) \\ =& S_{n-1}(1 + b)\mathbb{Q}(\xi_n = 1 + b | \mathcal{F}_{n-1}) \\ & + S_{n-1}(1 + a)\mathbb{Q}(\xi_n = 1 + a | \mathcal{F}_{n-1})\end{aligned}$$

and hence

$$\mathbb{Q}(\xi_n = 1 + b | \mathcal{F}_{n-1}) = q = 1 - \mathbb{Q}(\xi_n = 1 + a | \mathcal{F}_{n-1}).$$

Note $0 < q < 1$ if and only if $a < r < b$. Also, under this condition, the conditional distribution of ξ_n is independent of n and $\mathcal{F}_{n-1}$, so the $(\xi_n)_{1 \le n \le N}$ are IID. □

Stochastic Financial Models 19

Michael Tehranchi

November 23, 2022

1 Pricing and hedging European claims

Definition. A *European* contingent claim is an asset that pays an $\mathcal{F}_N$-measurable amount Y at a fixed time N.

A European claim is often called *plain vanilla* if its payout of the form $Y = g(S_N)$ for some function g. For instance a call option with payout $Y = (S_N - K)^+$ is a vanilla contingent claim. Otherwise, a claim whose payout depends on the entire path of the underlying asset price is called *exotic*.

Consider the binomial model with $a < r < b$ and a claim with payout $g(S_N)$.

Let
$$V(n,s) = \frac{1}{(1+r)^{N-n}} \mathbb{E}^{\mathbb{Q}}[g(S_N)|S_n = s]$$
for all $0 \le n \le N$ and let
$$\theta_n = \frac{V(n, S_{n-1}(1+b)) - V(n, S_{n-1}(1+a))}{S_{n-1}(b-a)}$$
for $1 \le n \le N$

Theorem. *The wealth process starting from $X_0 = V(0, S_0)$ employing trading strategy $(\theta_n)_{1 \le n \le N}$ is such that*
$$X_n = V(n, S_n) \text{ for all } 0 \le n \le N.$$
In particular, $X_N = g(S_N)$ so the strategy replicates the claim and $V(n, S_n)$ is the unique time-n no-arbitrage price of the claim.

Note that V solves
$$\begin{aligned} V(N,s) &= g(s) \\ (1+r)V(n-1,s) &= qV(n, s(1+b)) + (1-q)V(n, s(1+a)) \end{aligned}$$
for $1 \le n \le N$ by Markov Chains.

Proof. Suppose $X_{n-1} = V(n-1, S_{n-1})$. Then

$$\begin{aligned} X_n =& (1+r)X_{n-1} + \theta_n[S_n - (1+r)S_{n-1}] \\ =& q\mathbb{V}(b) + (1-q)\mathbb{V}(a) \\ &+ \frac{\mathbb{V}(b) - \mathbb{V}(a)}{b-a}[\xi_n - (1+r)] \\ =& \mathbb{V}(b)\frac{\xi_n - (1+a)}{b-a} + \mathbb{V}(a)\frac{1+b-\xi_n}{b-a} \\ =& V(n, S_n) \end{aligned}$$

where $\mathbb{V}(c) = V(n, S_{n-1}(1+c))$. Since $X_0 = V(0, S_0)$ by assumption, the induction is complete. $\square$

Note $M_n = (1+r)^{-n}V(n, S_n)$ is a $\mathbb{Q}$-martingale, as it should by the FTAP!

2 American claims

Definition. Given an adapted process $(Y_n)_{0\le n\le N}$, an *American* contingent claim is a contract that pays its owner Y_n if the owner chooses to exercise the contract at time n.

Example. An American call gives its owner the right, but not the obligation, to buy a certain stock for a fixed strike price K for at *any time* up to the expiry N. The payout if exercised at time n is $(S_n - K)^+$.

The time-n price of an American claim in a complete market with unique risk neutral measure $\mathbb{Q}$ can be calculated as

$$\pi_n = \max_{n\le T\le N} \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{(1+r)^{T-n}}Y_T|\mathcal{F}_n\right]$$

where the maximum is over stopping times T.

By the dynamic programming principle

$$\begin{aligned} \pi_N &= Y_N \\ \pi_{n-1} &= \max\left\{Y_{n-1}, \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}(\pi_n|\mathcal{F}_{n-1})\right\} \end{aligned}$$

An optimal stopping time is

$$T^* = \min\{0 \le n \le N : \pi_n = Y_n\}$$

but it need not be unique.

Proposition. *Suppose the adapted process $(Y_n)_{0\le n\le N}$ which specifies the payout of an American claim of expiry N is such that $(Y_n(1+r)^{-n})_{0\le n\le N}$ is a submartingale in a complete market with unique risk neutral measure $\mathbb{Q}$. Then an optimal exercise time is $T = N$.*

Proof. We will show that $\pi_n = \frac{1}{(1+r)^{N-n}}\mathbb{E}^{\mathbb{Q}}[Y_N|\mathcal{F}_n]$ for all $0 \leq n \leq N$. Therefore, the no-arbitrage price of the American claim is always equal to the price of a European claim with payout Y_N; that is, it optimal to wait until the expiry to exercise.

To show this, note $\pi_N = Y_N$, so the claim is true when $n = N$. Now supposing $\pi_n = \frac{1}{(1+r)^{N-n}}\mathbb{E}^{\mathbb{Q}}[Y_N|\mathcal{F}_n]$ for some $n \leq N$, note that

$$\begin{aligned}\frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}[\pi_n|\mathcal{F}_{n-1}] &= \frac{1}{(1+r)^{N-n+1}}\mathbb{E}^{\mathbb{Q}}[Y_N|\mathcal{F}_{n-1}] \\ &\geq Y_{n-1}\end{aligned}$$

by the tower property and the assumed submartingale property. Hence

$$\begin{aligned}\pi_{n-1} &= \max\left\{Y_{n-1}, \frac{1}{1+r}\mathbb{E}^{\mathbb{Q}}[\pi_n|\mathcal{F}_{n-1}]\right\} \\ &= \frac{1}{(1+r)^{N-n+1}}\mathbb{E}^{\mathbb{Q}}[Y_N|\mathcal{F}_{n-1}]\end{aligned}$$

and the backward induction is complete. □

Stochastic Financial Models 20

Michael Tehranchi

November 23, 2022

1 Put options

Definition. A European *put* is the right but not the obligation to sell certain asset at a fixed strike price K at a fixed expiry date N.

By an argument analogous to the case of a European call, the payout is

$$Y = (K - S_N)^+$$

Theorem (Put-call parity)**.** *Consider an arbitrage-free market with risk-free asset, a stock and a European call of strike K and expiry N.*

The payout of a European put of strike K and expiry N is attainable, and its unique time n no-arbitrage price is

$$P_n = K(1+r)^{-(N-n)} - S_n + C_n$$

where C_n is the time n price of the call.

Proof. Note

$$(K - S_N)^+ = K - S_N + (S_N - K)^+$$

identically. The replicating portfolio is to buy one call, to sell one share of the stock, and to hold the amount $K(1+r)^{-(N-n+1)}$ in the risk-free asset during the interval $(n-1, n]$. □

2 Continuous-time finance

From discrete to continuous. Motivation

- Let $S_n = S_0 \xi_1 \cdots \xi_n$ be the stock price in the binomial model
- Then $\log S_n = \log S_0 + X_1 + \ldots + X_n$ is a random walk with $(X_n)_n$ are independent copies of X with $\mathbb{P}(X = \beta) = p = 1 - \mathbb{P}(X = \alpha)$ where $\beta = \log(1+b)$ and $\alpha = \log(1+a)$.
- Now time step n corresponds to time $t = n\delta$ where δ is very small.

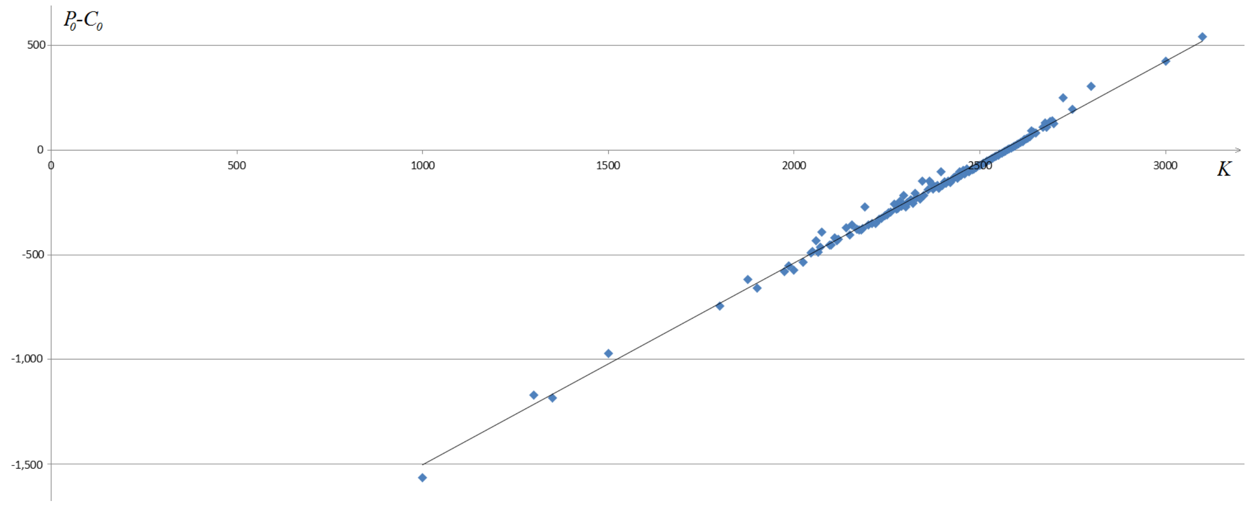


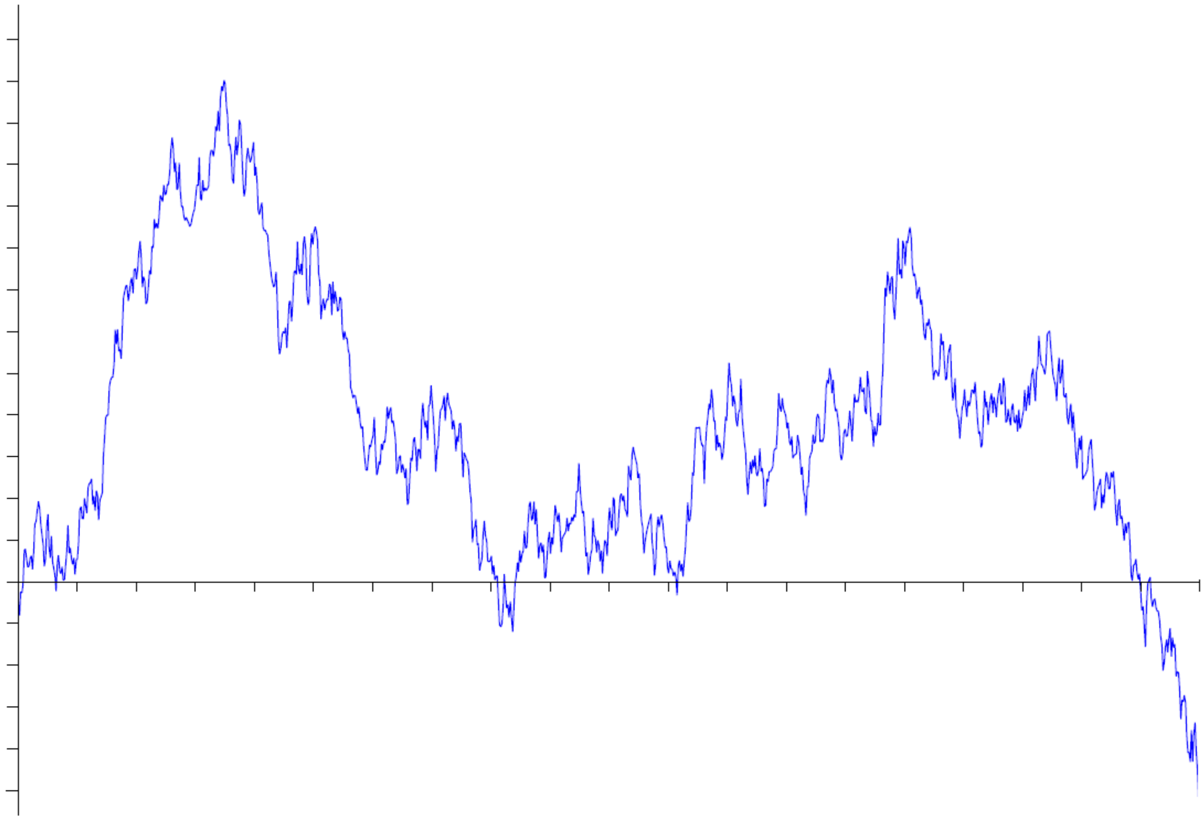
Figure 1: European put prices minus call prices versus strike

- Let $\hat{S}_t = S_{t/\delta}$
- Then
$$\log \hat{S}_t = \log S_0 + \mu t + \sigma W_t$$
where
  - $\mu = \mathbb{E}(X)/\delta$
  - $\sigma^2 = \mathrm{Var}(X)/\delta$
  - $W_t - W_s$ is independent of W_u for $0 \le u \le s < t$ and where $u/\delta, s/\delta, t/\delta$ are integers
  - and by the limit theorem
$$\frac{W_t - W_s}{\sqrt{t-s}} \approx N(0,1)$$
as $\delta \downarrow 0$ (and hence $m, n \to \infty$, where $n = t/\delta$ and $m = t/\delta$)

3 Introduction to Brownian motion

Definition. A *Brownian motion* $(W_t)_{t \ge 0}$ is a stochastic process such that

- $t \mapsto W_t$ is continuous
- $W_0 = 0$
- $W_t - W_s$ is independent of $(W_u)_{0 \le u \le s}$
- $W_t - W_s \sim N(0, t-s)$



Stochastic Financial Models 21

Michael Tehranchi

November 23, 2022

1 Properties of Brownian motion

Theorem (Wiener 1923). *Brownian motion exists.*

Remark. A Brownian motion is called a Wiener process in the US.

Theorem. *Brownian motion is a martingale in its filtration $\mathcal{F}_t = \sigma(W_s : 0 \le s \le t)$.*

Proof. Brownian motion is integrable, adapted and

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

for $0 \le s \le t$ by the independence of $W_t - W_s$ and $\mathcal{F}_s$. □

Theorem. *Brownian motion is a Markov process.*

Proof. Since W_s is $\mathcal{F}_s$ measurable and $W_t - W_s$ is independent of $\mathcal{F}_s$ for $0 \le s \le t$, we have

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

□

Definition. A process $(X_t)_{t\ge0}$ is *Gaussian* iff the random variables $X_{t_1}, \ldots, X_{t_n}$ are jointly normal for all $0 \le t_1 \le \ldots \le t_n$, i.e. the random variable $\sum_{i=1}^n a_i X_{t_i}$ is normally distributed for all constants $a_1, \ldots, a_n$.

Theorem. *The following are equivalent*

1. *$(W_t)_{t\ge0}$ is a Brownian motion*
2. *$(W_t)_{t\ge0}$ is a Gaussian process such that*
   - *$t \mapsto W_t$ is continuous*

- $\mathbb{E}[W_t] = 0$ *for all* $t \geq 0$
- $\mathbb{E}[W_s W_t] = s$ *for all* $0 \leq s \leq t$

Proof. Suppose $(W_t)_{t\geq 0}$ is a Brownian motion. Fix $0 = t_0 \leq t_1 \leq \ldots \leq t_n$ and $a_1, \ldots, a_n$. Note

$$\sum_{i=1}^{n} a_i W_{t_i} = \sum_{i=1}^{n} b_i (W_{t_i} - W_{t_{i-1}})$$

where $b_k = \sum_{i=k}^{n} a_i$. Since $W_{t_1} - W_{t_0}, \ldots, W_{t_n} - W_{t_{n-1}}$ are independent normals, and the linear combination of independent normals is normal, we have that $(W_t)_{t\geq 0}$ is Gaussian with $\mathbb{E}[W_t] = \mathbb{E}[W_0] = 0$ and

$$\begin{aligned} \mathbb{E}[W_s W_t] &= \mathbb{E}[W_s^2] + \mathbb{E}[W_s(W_t - W_s)] \\ &= \operatorname{Var}(W_s) + \mathbb{E}(W_s)\mathbb{E}(W_t - W_s) \\ &= 0. \end{aligned}$$

for $0 \leq s \leq t$, since W_s and $W_t - W_s$ are independent.

Conversely, suppose $(W_t)_{t\geq 0}$ is a continuous Gaussian process such that $\mathbb{E}[W_t] = 0$ and $\mathbb{E}[W_s W_t] = s$ for all $0 \leq s \leq t$. Then for $0 \leq u \leq s \leq t$ we have

$$\begin{aligned} \operatorname{Cov}(W_u, W_t - W_s) &= \mathbb{E}[W_u W_t] - \mathbb{E}[W_u W_s] \\ &= u - u = 0 \end{aligned}$$

By normality, the increment $W_t - W_s$ is independent of W_u. By Gaussianity, the increment is independent of $(W_u)_{0\leq u\leq s}$.

Theorem. *Let $(W_t)_{t\geq 0}$ be a Brownian motion. Then each of the following processes are also Brownian motions.*

1. $\tilde{W}_t = cW_{t/c^2}$, *for any constant* $c \neq 0$.
2. $\tilde{W}_t = W_{t+T} - W_T$ *for any constant* $T \geq 0$.
3. $\tilde{W}_0 = 0$ *and* $\tilde{W}_t = tW_{1/t}$ *for* $t > 0$.

Proof. Check that each process is a continuous mean-zero Gaussian process with the correct covariance. [For 3, we technically need the Brownian law of large number $\frac{W_s}{s} \to 0$ as $s \to \infty$ to prove continuity of $\tilde{W}$ at $t = 0$.] □

2 Reflection principle

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

Proof. Consider the case $a > 0$. (The case $a < 0$ is similar.) We must show

$$\sup_{t \geq 0} W_t > a \text{ almost surely}$$

By Brownian scaling, for any $c > 0$ and $0 < a < b$, we have

$$\begin{aligned}
\mathbb{P}(a < \sup_{t \geq 0} W_t < b) &= \mathbb{P}(a < \sup_{t \geq 0} cW_{t/c^2} < b) \\
&= \mathbb{P}(a/c < \sup_{s \geq 0} W_s < b/c) \quad \text{letting } t/c^2 = \\
&\to 0
\end{aligned}$$

by sending $c \uparrow \infty$. Since $Z = \sup_{t \geq 0} W_t \geq W_0 = 0$, we have shown that $Z \in \{0, +\infty\}$ almost surely.

Let $\hat{Z} = \sup_{t \geq 1}(W_t - W_1)$. Note Z and $\hat{Z}$ have the same distribution, so $\hat{Z} \in \{0, +\infty\}$ almost surely.

Note that $\{\hat{Z} = \infty\} = \{Z = \infty\}$ since $\sup_{0 \leq t \leq 1} W_t$ is finite by the continuity of Brownian motion. Hence

$$\begin{aligned}
p = \mathbb{P}(Z = 0) &= \mathbb{P}(Z = 0, \hat{Z} = 0) \\
&\leq \mathbb{P}(W_1 \leq 0, \hat{Z} = 0) \\
&= \frac{1}{2}\mathbb{P}(\hat{Z} = 0) = \frac{1}{2}p
\end{aligned}$$

so $p = 0$. Hence $\sup_{t \geq 0} W_t = \infty$ almost surely. □

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

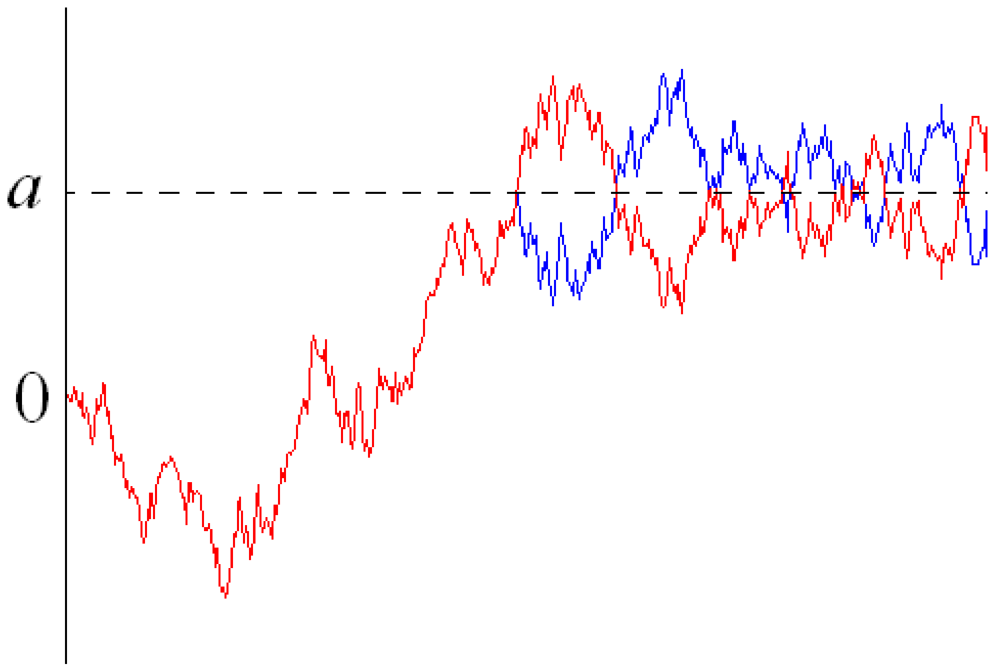
Proof. Omitted. The idea is Brownian motion is a *strong* Markov process. □

Applying this with the finite stopping time T_a together with the symmetry of Brownian motion, we have

Theorem (Reflection principle)**.** *Let $(W_t)_{t \geq 0}$ be a Brownian motion and let*

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

Then $(\tilde{W}_t)_{t \geq 0}$ is a Brownian motion.



Reflection principle: Key formula

$$\boxed{\mathbb{P}(\max_{0\le s\le t} W_s \ge a, W_t \le b) = \mathbb{P}(W_t \ge 2a-b) \text{ for } a \ge 0, b \le a}$$

Proof. We have

$$\begin{aligned}\mathbb{P}(\max_{0\le s\le t} W_s \ge a, W_t \le b) &= \mathbb{P}(\tilde{W}_t \ge 2a-b)\\ &= \mathbb{P}(W_t \ge 2a-b)\end{aligned}$$

□

Stochastic Financial Models 22

Michael Tehranchi

November 25, 2022

1 Applications of the reflection principle

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

$$\begin{aligned} F_{M_t,W_t}(a,b) &= \mathbb{P}(M_t \leq a, W_t \leq b) \\ &= \mathbb{P}(W_t \leq b) - \mathbb{P}(M_t > a, W_t \leq b) \\ &= \mathbb{P}(W_t \leq b) - \mathbb{P}(W_t \geq 2a - b) \text{ for } a \geq 0, b \leq a \\ &= F_{W_t}(b) + F_{W_t}(2a-b) - 1 \\ &= \Phi\left(\frac{b}{\sqrt{t}}\right) + \Phi\left(\frac{2a-b}{\sqrt{t}}\right) - 1 \end{aligned}$$

where $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$ is the standard normal distribution function.

The joint density of (M_t, W_t) is

$$\begin{aligned} f_{M_t,W_t}(a,b) &= \frac{\partial^2}{\partial a \partial b} F_{M_t,W_t}(a,b) \\ &= -2 f'_{W_t}(2a-b) \text{ for } a \geq 0, b \leq a \\ &= \frac{2(2a-b)}{t^{3/2}} \varphi\left(\frac{2a-b}{\sqrt{t}}\right) \end{aligned}$$

where $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ is the standard normal density

Proposition. *For suitable g and $a \geq 0$ we have*

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

Proof.

$$\begin{aligned} \mathbb{E}[g(W_t) \mathbb{1}_{\{M_t \geq a\}}] &= \mathbb{E}[g(W_t) \mathbb{1}_{\{M_t \geq a, W_t \geq a\}}] + +\mathbb{E}[g(W_t) \mathbb{1}_{\{M_t \geq a, W_t \leq a\}}] \\ &= \mathbb{E}[g(W_t) \mathbb{1}_{\{W_t \geq a\}}] + \mathbb{E}[g(2a - W_t) \mathbb{1}_{\{2a - W_t \leq a\}}] \end{aligned}$$

□

Another consequence for $a \geq 0$

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

Hence, For $a \geq 0$, we have

$$\begin{aligned} \mathbb{P}(T_a \leq t) &= \mathbb{P}(M_t \geq a) \\ &= 2\left[1 - \Phi\left(\frac{a}{\sqrt{t}}\right)\right] \\ &= 2\Phi\left(-\frac{a}{\sqrt{t}}\right) \end{aligned}$$

using the symmetry of the normal distribution.

Hence T_a has density

$$f_{T_a}(t) = \frac{a}{t^{3/2}}\varphi\left(\frac{a}{\sqrt{t}}\right) \text{ for } t > 0$$

2 Cameron–Martin theorem

Motivation. Example sheet 1

- Let $Z \sim N(0,1)$.
- $\mathbb{E}[g(a+Z)] = \mathbb{E}[e^{aZ - a^2/2}g(Z)]$ for any $a \in \mathbb{R}$ and suitable g.
- Proof: Change of variables formula for integration.

Generalisation

- Let $Z \sim N_n(0, I)$ multi-variate normal.
- $\mathbb{E}[g(a+Z)] = \mathbb{E}[e^{a^\top Z - \|a\|^2/2}g(Z)]$ for any $a \in \mathbb{R}^n$ and suitable g.
- Essentially the same proof.

Theorem (Cameron–Martin theorem)**.** *Let $(W_t)_{t \geq 0}$ be a Brownian motion. For fixed $t \geq 0$ and $c \in \mathbb{R}$ we have*

$$\mathbb{E}[g((W_s + cs)_{0 \leq s \leq t})] = \mathbb{E}[e^{cW_t - c^2 t/2}g((W_s)_{0 \leq s \leq t})]$$

for suitable functions g from the space of continuous functions on $[0,t]$ to the real line.

Sketch of proof. By measure theory, it is enough to consider functions g of the form

$$g(w) = G(w(t_1), \dots, w(t_n))$$

for a function G on $\mathbb{R}^n$, where $0 = t_0 < t_1 < \dots < t_n = t$.

$$\begin{aligned}
\mathbb{E}[g((W_s + cs)_{0 \le s \le t}) &= \mathbb{E}[G(W_{t_1} + ct_1, \dots, W_{t_n} + ct_n)] \\
&= \mathbb{E}[G((\sum_{i=1}^{k} \sqrt{t_i - t_{i-1}}(Z_i + a_i))_{k=1}^{n})]] \\
&= \mathbb{E}[e^{a^\top Z - \|a\|^2/2} G((\sum_{i=1}^{k} \sqrt{t_i - t_{i-1}} Z_i)_{k=1}^{n})]] \\
&= \mathbb{E}[e^{cW_t - c^2 t/2} g((W_s)_{0 \le s \le t})]
\end{aligned}$$

where $Z_i = \frac{W_{t_i} - W_{t_{i-1}}}{\sqrt{t_i - t_{i-1}}}$ are iid $N(0,1)$ for $1 \le i \le n$ and $a_i = c\sqrt{t_i - t_{i-1}}$ so that

$$a^\top Z = \sum_{i=1}^{n} a_i Z_i = W_t$$

and

$$\|a\|^2 = \sum_{i=1}^{n} a_i^2 = c^2 t$$

$\square$

Stochastic Financial Models 23

Michael Tehranchi

November 28, 2022

1 An application of Cameron–Martin

Proposition. *Let $(W_t)_{t\geq 0}$ be a Brownian motion. For $a \geq 0$ we have*

$$\begin{aligned}\mathbb{P}(\max_{0\leq s\leq t}(W_s + cs) \leq a) &= \mathbb{P}(W_t \leq a - ct) - e^{2ca}\mathbb{P}(W_t \geq a + ct) \\ &= \Phi\left(\frac{a - ct}{\sqrt{t}}\right) - e^{2ca}\Phi\left(\frac{-a - ct}{\sqrt{t}}\right)\end{aligned}$$

Proof.

$$\begin{aligned}\mathbb{P}(\max_{0\leq s\leq t}(W_s + cs) \leq a) =&\mathbb{E}[\mathbb{1}_{\{\max_{0\leq s\leq t}(W_s+cs)\leq a\}}] \\ =&\mathbb{E}[e^{cW_t - c^2t/2}\mathbb{1}_{\{\max_{0\leq s\leq t} W_s \leq a\}}] \\ =&\mathbb{E}[e^{cW_t - c^2t/2}\mathbb{1}_{\{W_t\leq a\}}] \\ &- \mathbb{E}[e^{c(2a-W_t)-c^2t/2}\mathbb{1}_{\{W_t\geq a\}}] \\ =&\mathbb{E}[\mathbb{1}_{\{W_t+ct\leq a\}}] - e^{2ac}\mathbb{E}[\mathbb{1}_{\{W_t-ct\geq a\}}]\end{aligned}$$

□

To discuss risk-neutral measures, we need

Theorem (Cameron–Martin reformulation)**.** *Let $(W_t)_{t\geq 0}$ be a Brownian motion under a given measure $\mathbb{P}$. Fix $T > 0$ and $c \in \mathbb{R}$, and define an equivalent measure $\mathbb{Q}$ by*

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

Then the process $(W_t - ct)_{0\leq t\leq T}$ is a Brownian motion under $\mathbb{Q}$.

Proof. Fix a function g on $C[0, T]$. Then

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}[g((W_t - ct)_{0\leq t\leq T})] &= \mathbb{E}^{\mathbb{P}}[e^{cW_T - c^2T/2}g((W_t - ct)_{0\leq t\leq T})] \\ &= \mathbb{E}^{\mathbb{P}}[g((W_t)_{0\leq t\leq T})]\end{aligned}$$

by the first formulation of Cameron–Martin. So the process $(W_t - ct)_{0\leq t\leq T}$ has the same law under $\mathbb{Q}$ as the process $(W_t)_{0\leq t\leq T}$ has under $\mathbb{P}$. □

2 Heat equation

Proposition. *Fix a suitable g and let*

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

where $Z \sim N(0,1)$. Then u solves the heat equation

$$\partial_\tau u = \frac{1}{2}\partial_{xx}u$$

with boundary condition $u(0,x) = f(x)$.

Proof when g is well-behaved by example sheet 1,

$$\begin{aligned}\partial_\tau u &= \frac{1}{2\sqrt{\tau}}\mathbb{E}[Zg'(x+\sqrt{\tau}Z)] \\ &= \frac{1}{2}\mathbb{E}[g''(x+\sqrt{t}Z)] \\ &= \frac{1}{2}\partial_{xx}u\end{aligned}$$

If g is less well-behaved, then write

$$u(\tau,x) = \int f(y)p(\tau;x,y)dy$$

where

$$p(\tau;x,y) = \frac{1}{\sqrt{2\pi\tau}}\exp\left(-\frac{(y-x)^2}{2\tau}\right)$$

is the transition density of the Brownian motion (also called the *heat kernel* or *Green's function*) and use the fact that $p(\cdot;\cdot,y)$ satisfies the heat equation.

Since p is very well-behaved, interchange of derivatives and integrals is allowed by the dominated convergence theorem, provided that f has exponential growth. $\square$

3 Black–Scholes model

- A risk-free asset with constant (instantaneously compounded) interest rate r.
- A risky stock with time t price $(S_t)_{t\geq 0}$ where

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

  and $(W_t)_{t\geq 0}$ is a Brownian motion.

A risk neutral measure in this context is an equivalent measure $\mathbb{Q}$ under which the discounted stock price $(e^{-rt}S_t)_{t\geq 0}$ process is a martingale.

Theorem (Risk-neutrality in Black–Scholes)**.** *Over any horizon $T \geq 0$, there is a risk-neutral measure $\mathbb{Q}$ with density*

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

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

Proof. By Cameron–Martin, the process $\hat{W}_t = W_t - ct$ is a Brownian motion under $\mathbb{Q}$. Notice that

$$\begin{aligned} e^{-rt}S_t &= S_0 e^{(\mu-r)t+\sigma W_t} \\ &= S_0 e^{(\mu-r+c\sigma)t+\sigma\hat{W}_t} \\ &= S_0 e^{-\sigma^2 t/2+\sigma\hat{W}_t} \end{aligned}$$

is a martingale under $\mathbb{Q}$ by example sheet 4. □

Black–Scholes pricing

Definition. Consider a European contingent claim with time T payout Y. Within the Black–Scholes model, the time t price is

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

where $\mathbb{Q}$ is the risk-neutral measure.

Note $(e^{-rt}\pi_t)_{0\leq t\leq T}$ is a $\mathbb{Q}$-martingale.

For a vanilla European contingent claim with payout $Y = g(S_T)$ the price is

$$\begin{aligned} \pi_t &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}(g(S_T)|\mathcal{F}_t) \\ &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[g(S_t e^{(r-\sigma^2/2)(T-t)+\sigma(\hat{W}_T-\hat{W}_t)}|\mathcal{F}_t] \\ &= V(t, S_t) \end{aligned}$$

where

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

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

The Black–Scholes pricing function V for a European vanilla claim solves the *Black–Scholes PDE*

$$\boxed{\partial_t V + rs\partial_s V + \frac{1}{2}\sigma^2 s^2 \partial_{ss} V = rV}$$

with boundary condition $V(T,s) = g(s)$. One way to see this is to change variables and related V to a heat equation. See example sheet 4.

Stochastic Financial Models 24

Michael Tehranchi

December 4, 2022

1 Black–Scholes PDE

The Black–Scholes pricing function V for a European vanilla claim solves the *Black–Scholes PDE*

$$\boxed{\partial_t V + rs\partial_s V + \frac{1}{2}\sigma^2 s^2 \partial_{ss} V = rV}$$

with boundary condition $V(T, s) = g(s)$. One way to see this is to differentiate the integral formula for $V(t, s)$. Here is another approach:

- Due to the martingale property we have for $\delta > 0$ that

  $$e^{r\delta} V(t, s) = \mathbb{E}^{\mathbb{Q}}(V(t + \delta, s\xi))$$

  where $\xi = e^{(r-\sigma^2/2)\delta+\sigma\sqrt{\delta}Z}$ and $Z \sim N(0, 1)$.

- Taylor expand the expectation about (t, s)

  $$V(t + \delta, s\xi) = V(t, \delta) + \frac{\partial V}{\partial t}\delta + \frac{\partial V}{\partial s} s(\xi - 1) + \frac{1}{2}\frac{\partial^2 V}{\partial s^2} s^2(\xi - 1)^2 + \dots$$

- Now

  $$\mathbb{E}(\xi) = e^{r\delta} \approx 1 + r\delta$$

  and

  $$\mathbb{E}[(\xi - 1)^2] = e^{(2r+\sigma^2)\delta} - 2e^{r\delta} + 1 \approx \sigma^2\delta$$

- Now divide by δ and take the limit.

2 Black–Scholes greeks

- No-arbitrage price of European claim with time-T payout Y is

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

- If $Y = g(S_T)$ is the payout of a vanilla claim, then

$$\pi_t = V(t, S_t)$$

where V solves the Black–Scholes PDE

In the binomial model, in order to replicate the claim, at time t on the event $\{S_{t-\delta} = s\}$ you must hold

$$\frac{V(t, s(1+b)) - V(t, s(1+a))}{s(b-a)} \approx \frac{\partial V}{\partial s}(t, s)$$

shares of the underlying asset between times $t - \delta$ and t. In the Black–Scholes model, the quantity $\frac{\partial V}{\partial s}$ is called the *delta* of the claim.

The quantity $\frac{\partial^2 V}{\partial s^2}$ measures the sensitivity of delta with respect to movements of the stock price and is known as the *gamma* of the claim.

Proposition. *If the payout function g is increasing, then the delta is always non-negative. If g is convex, then the gamma is always non-negative.*

Proof. From the formula

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

it is clear (using the argument from example sheet 1) that $V(t, \cdot)$ is increasing when g is increasing, and that $V(t, \cdot)$ is convex when g is convex. □

3 Black–Scholes formula

The Black–Scholes price of a European call

$$\begin{aligned}\pi_t &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(S_T - K)^+|\mathcal{F}_t] \\ &= \boxed{S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)}\end{aligned}$$

where

$$d_1 = -\frac{\log(K/S_t)}{\sigma\sqrt{T-t}} + \left(\frac{r}{\sigma} + \frac{\sigma}{2}\right)\sqrt{T-t}$$

and

$$d_2 = -\frac{\log(K/S_t)}{\sigma\sqrt{T-t}} + \left(\frac{r}{\sigma} - \frac{\sigma}{2}\right)\sqrt{T-t}$$

Derivation: Let $\delta = T - t$ and $\xi = e^{(r-\sigma^2/2)\delta + \sigma\sqrt{\delta}Z}$ where $Z \sim N(0,1)$.

$$\begin{aligned} V(t,s) &= e^{-r\delta}\mathbb{E}[(s\xi - K)^+] \\ &= e^{-r\delta}\mathbb{E}[(s\xi - K)\mathbb{1}_{\{\xi > K/s\}}] \\ &= s\mathbb{E}(e^{-r\delta}\xi\mathbb{1}_{\{\xi > K/s\}}) - -e^{-r\delta}K\mathbb{P}(\xi > K/s) \end{aligned}$$

Note that $\mathbb{P}(\xi > K/s) = 1 - \Phi(-d_2) = \Phi(d_2)$. By the change of variables formula for normal random variables (see example sheet 1), the law of ξ under $\hat{\mathbb{P}}$ is the same as the law of $e^{\sigma^2\delta}\xi$ under $\mathbb{P}$, where $\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} = e^{-r\delta}\xi$. Hence

$$\mathbb{E}(e^{-r\delta}\xi\mathbb{1}_{\{\xi > K/s\}}) = \mathbb{P}(\xi > Ke^{-\sigma^2\delta}/s) = 1 - \Phi(-d_2 - \sigma\sqrt{\delta}) = \Phi(d_1)$$

The delta of the European call option is then (check)

$$\frac{\partial V}{\partial s} = \Phi(d_1)$$

4 Black–Scholes prices of barrier-type claims

Consider a market with a stock with price $(S_t)_{t\geq 0}$.

- Given a European contingent claim with payout Y and expiry T
- and given a level B
- A *down-and-in* version of the claim has payout $Y\mathbb{1}_{\{\min_{0\leq t\leq T} S_t \leq B\}}$
- *down-and-out* has payout $Y\mathbb{1}_{\{\min_{0\leq t\leq T} S_t > B\}}$
- *up-and-in* has payout $Y\mathbb{1}_{\{\max_{0\leq t\leq T} S_t \geq B\}}$
- *up-and-out* has payout $Y\mathbb{1}_{\{\max_{0\leq t\leq T} S_t < B\}}$

Example. An up-and-in call option with strike K and barrier B, gives the owner of the option the right, but not the obligation, to buy the stock at time T for price K, provided that the price of the stock exceeds B at some time between time 0 and time T.

Proposition. *Within the Black–Scholes model, the initial price of an up-and-out claim with payout*

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

is the same as that of a vanilla option with payout

$$g(S_T)\mathbb{1}_{\{S_T \leq B\}} - (B/S_0)^{2r/\sigma^2 - 1} g(B^2 S_T/S_0^2)\mathbb{1}_{\{S_T \leq S_0^2/B\}}$$

Proof. Since $\{\max_{0\le t\le T} S_t < B\} \subseteq \{S_T < B\}$, we have

$$g(S_T)\mathbb{1}_{\{\max_{0\le t\le T} S_t<B\}} = g(S_T)\mathbb{1}_{\{S_T<B\}} - g(S_T)\mathbb{1}_{\{\max_{0\le t\le T} S_t\ge B, S_T<B\}}$$

Looking at the second term on the right, letting $b = \log(B/S_0)/\sigma$ and $c = r/\sigma - \sigma/2$.

$$g(S_T)\mathbb{1}_{\{\max_{0\le t\le T} S_t\ge B, S_T\le B\}} = g(S_0 e^{\sigma(W_T+cT)})\mathbb{1}_{\{\max_{0\le t\le T}(W_t+ct)\ge b, W_T+cT\le b\}}$$

where $(W_t)_{0\le t\le T}$ is a Brownian motion under the risk-neutral measure.

By the Cameron–Martin theorem, the expected value is

$$\mathbb{E}[e^{cW_T-c^2T/2} g(S_0e^{\sigma W_T})\mathbb{1}_{\{\max_{0\le t\le T} W_t\ge b, W_T\le b\}}]$$

by the reflection principle

$$= \mathbb{E}[e^{c(2b-W_T)-c^2T/2} g(S_0e^{\sigma(2b-W_T)})\mathbb{1}_{\{W_T\ge b\}}]$$

by symmetry of W_T

$$= e^{2bc}\mathbb{E}[e^{cW_T-c^2T/2} g(S_0e^{2b\sigma}e^{\sigma W_T})\mathbb{1}_{\{W_T\le -b\}}]$$

by Cameron–Martin again

$$= e^{2bc}\mathbb{E}[g(S_0e^{2b\sigma}e^{\sigma(W_T+cT)})\mathbb{1}_{\{W_T+cT\le -b\}}]$$

Rewriting

$$e^{2bc}\mathbb{E}[g(e^{2b\sigma}S_T)\mathbb{1}_{\{S_T\le S_0e^{-b\sigma}\}}] = (B/S_0)^{2r/\sigma^2-1}\mathbb{E}[g(B^2S_T/S_0^2)\mathbb{1}_{\{S_T\le S_0^2/B\}}]$$

□