3508 Financial Mathematics Notes

Based on the 2015 spring lectures by Dr J Walton

OUTDATED

03/10/14

M3508: Mathematics of Finance

Dr Jamie Walton

Morgan Stanley and
University College London

Autumn 2014

walton@math.ucl.ac.uk

https://www.facebook.com/DrJamieWalton

Thurs 4pm HW deadline

1 Introduction

1.1 What is financial mathematics?

Mathematical finance as a discipline covers the broad area of all mathematics applicable to finance and banking. For example:

- Stock prediction
- Portfolio theory (Capital Asset Pricing Model)
- Utility theory (of investors)
- Game theory (for trading, eg. 4G licence auction, Nash equilibria)
- Valuation theory (for valuing companies/assets; Corporate Finance)
- Real option theory (for valuing decision-making)
- VaR - Value at Risk (for risk-management)
- Derivative pricing

This course will look at the most mathematically precise, wide-ranging and, in my opinion, interesting of these topics: derivative pricing (not least because that is my job).

Derivative pricing dates back to 1900 (Bachelier) though in reality took-off with the Nobel-prize winning theories of Black-Scholes-Merton and the advent of derivative trading on exchanges in 1973. Derivatives are now the largest assets traded globally and drive the world's capital economies. The mathematics underlying the pricing of derivatives is the fastest growing and probably the largest area of mathematical research currently although the bulk of research is proprietary and unpublished.

1.2 What is a derivative?

First, we need to define the main asset classes of finance. These are:

- **Equities**: Stocks/shares, stock indices (FTSE, S&P Nasdaq...)

  loans US govt debt UK equivalent
- **Interest Rates**: Bonds, swaps, treasuries, gilts...

- **Foreign Exchange (FX)**: EURUSD, GBPUSD, USDJPY...
- **Commodities**: Oil, gas, precious metals, porkbellies, weather...
- **Credit**: CDS, CDO...

credit default swap

Definition 1. *A* derivative *is any product whose value is dependent on one or more underlying assets.*

Two common types of derivatives are forwards and options:

Definition 2. *A* forward *allows you to buy or sell an underlying asset at a future date at a level determined today. The level where you buy or sell is often called the* strike. *

Forwards
- OTC market

Forwards are usually traded for zero value so that all cash-flows, which can be either positive or negative, occur on the expiration date.

The value of a long forward at time T is given by:

buy

$$Fwd(T) = S(T) - K$$

where S is the underlying asset and K is the strike price agreed at the start. This is called the *payoff* function.

Futures
- exchanges

Futures have the same payoff as a forward but are usually traded on exchanges and have a different profile as margin needs to be posted over the life of the option to cover the change in value of the position. This leads to a difference in accounting for a forward and a future.

So the value of a forward can go positive or negative. What if we only want to be exposed to positive payoffs?

Definition 3. *An* option *is the right to buy/sell the underlying asset at a fixed strike on a given future date.*

An option to buy the underlying is called a call *option and the option to sell the underlying is called a* put *option.*

As the option is a right and not an obligation, it is only exercised if it has positive value (we say it is *in-the-money*).

The payoff of call option struck at K is:

$$Call(T) = max(S(T) - K, 0)$$

There are two common types of options:

* \$110 = spot price i.e. current price $S(0)$

K = \$120 = 1 yr fwd $F(1)$ forward price / strike price

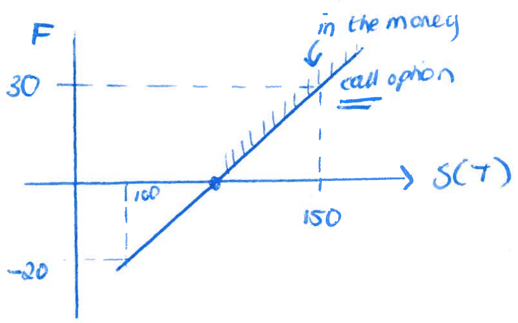
Definition 4. *A* European option *is only exercised on the expiration date. An* American option *can be exercised at any point during the life of the option.*

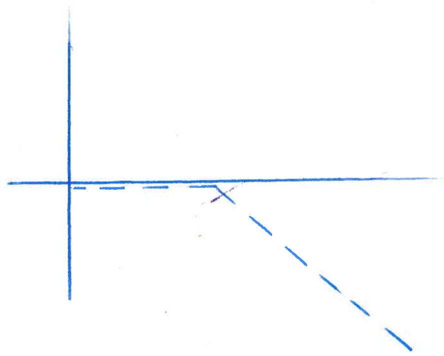
Since an option has positive value at a future date, a fee called a *premium* has to be paid to buy options. Calculating the price of an option will be the primary concern of this course.

1.3 What mathematical techniques will be useful?

make sure you revise / learn

- Linear algebra
- Probability theory: expectation, PDFs, normal distributions, central limit theorem, variance/covariance
- Probability measures: filtrations, change of measures
- Conditional expectation and martingales
- Brownian motion
- Stochastic calculus, Ito's lemma
- PDEs, Feyman-Kac





short an option to buy

Buy underlying = call
Sell underlying = put

Long = buy
Short = sell

Suggested Reading

Introductory Level

J. Hull, *Options, Futures and Other Derivatives*, Prentice Hall, various editions

M. Baxter and A. Rennie, *Financial Calculus*, CUP, 1996

Stanley R. Pliska, *Introduction to Mathematical Finance – Discrete Time Models*, Blackwell, 1997

Sheldon M. Ross, *An Elementary Introduction to Mathematical Finance: Options and Other Topics*, CUP, 2003

Further Reading

N.H. Bingham and R. Kiesel, *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives*, Springer, 2002

Mark Joshi, *The Concepts and Practice of Mathematical Finance*, CUP, 2003

Paul Wilmott, Sam Howison and Jeff Dewynne, *The Mathematics of Financial Derivatives: A Student Introduction*, CUP, 1997

Anecdotal

Peter L. Bernstein, *Capital Ideas: Improbable Origins of Modern Wall Street*, The Free Press, 1992

Peter L. Bernstein, *Against the Gods: The Remarkable Story of Risk*, Wiley, 1996

Michael Lewis, *Liar's Poker: Playing the Money Markets*, Coronet, 1999

Roger Lowenstein, *When Genius Failed: The Rise and Fall of Long Term Capital Management*, Fourth Estate, 2001

Frank Partnoy, *F.I.A.S.C.O.: Guns, Booze and Bloodlust – The Truth About High Finance*, Profile Books, 1998

2 Introduction to Foreign Exchange

Foreign exchange is the largest market in the world with a reported \$1.5 trillion traded every day. The vast majority is traded interbank and the options market is primarily over-the-counter (OTC). All asset classes have some FX dependency, for example, if you are a UK investor in US equities you have exposure to the US dollar (USD) exchange rate against the GB pound (GBP).

2.1 Notation

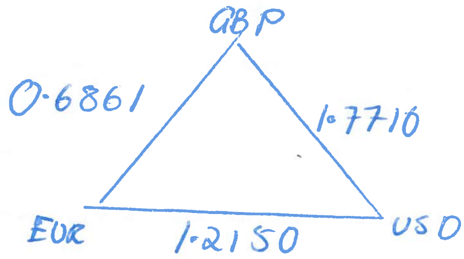
The underlying asset in FX is the exchange rate between 2 currencies which are denoted by 3-letter abbreviations:

- GBP: GB pound (also known as Sterling)
- USD: US dollar
- JPY: Japanese yen
- EUR: Euro
- CHF: Swiss Franc (Swissie)
- AUD: Australian dollar (Aussie)
- NZD: New Zealand dollar (Kiwi, or the Bird)
- CAD: Canadian Dollar (Loonie)

So the Sterling-dollar exchange rate is written as GBPUSD = 1.7710 which means 1 GBP is worth 1.7710 USD.

Buying GBPUSD means buying GBP and selling USD.
A call option on GBPUSD is a call on GBP/put on USD.

In reality all transactions are an exchange of assets. When you buy UK shares such as Vodafone you are exchanging VOD for GBP, ie. buying VODGBP.



2.2 Triangular arbitrage

There is a natural relationship between exchange rates:

Buying GBPUSD @ 1.7710
Selling EURUSD @ 1.2150

in the same amount of USD (same USD notional) is equivalent to:

selling EURGBP

So what is the exchange rate for EURGBP?

It had better be 1.2150/1.7710 = 0.6861 otherwise I can make free money! This is known as *triangular arbitrage* and, normally, transaction costs such as commission and bid/ask spread prevent this from occurring.

Arbitrage is the primary mechanism stabilising the markets and we will talk more about different arbitrage opportunities next week.

Can have quadruple arbitrage

2.3 The carry trade

Suppose I am a JPY based investor earning 0% interest on my savings. I may prefer to put all my cash in GBP and earn 5% interest. This is the simplest example of a class of "arbitrage trades" known as *carry trades*.

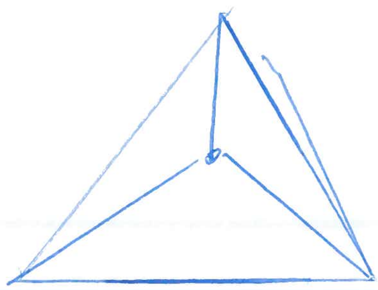
Note: I have put the word arbitrage in quotes because this trade does not lead to guaranteed returns, the clear risk here is that you are exposed to exchange rate moves.

2.4 Siegel's paradox

Lets suppose GBPUSD falls from 2.00 to 1.5, ie 25%. For a USD based investor, USDGBP has risen from 0.5 to 0.6667, ie a gain of 33%.

Does this asymmetry matter?

We shall talk about this later in the course and see some of the consequences.



2.5 Covered interest rate arbitrage

FX forwards are normally traded for zero cost. How can we calculate the *strike* value for an FX forward?

Example 5. *Let us calculate the value of a 1 year USDJPY forward. We know the values:*

- *Spot USDJPY* = 120.00
- *USD interest rate* = 4%
- *JPY interest rate* = 1%

What else do we need to know?

- Expected return on USDJPY?
- Volatility of USDJPY?
- More generally, expected return distribution of USDJPY?

NO!

The price of a forward can be determined purely by *arbitrage*.
It is independent of where spot is in one year's time.
It is independent of the expected distribution of returns on spot.

2.5.1 Graphical explanation of covered interest rate arbitrage

$$Time\ t = 0: \quad USD \xrightarrow{Spot} JPY$$

$$4\% = r_{USD} \downarrow \qquad\qquad \downarrow r_{JPY} = 1\%$$

$$Time\ t = 1y: \quad USD \xrightarrow{Fwd} JPY$$

To deliver a fwd in one year's time, either,

- Invest USD and earn USD interest r_{USD}. Convert to JPY at fwd value in 1 year's time, or
- Convert to JPY today and earn JPY interest r_{JPY} for 1 year.

These must be equivalent or there is an arbitrage.

2.5.2 Arbitrage equation

any amount

Suppose we have a notional amount of USD denoted N_{USD}. Then

- $Spot * N_{USD} * (1 + r_{JPY} * T)$ time
- $N_{USD} * (1 + r_{USD} * T) * Forward$

By arbitrage, these are equivalent so

$$Fwd = Spot * \frac{(1 + r_{JPY}T)}{(1 + r_{USD}T)}$$

This form of arbitrage is called *covered interest rate arbitrage* and causes the relationship between spot and forward to be determined by covered interest rate parity.

$$Spot * N_{USD} * (1 + r_{JPY} * T) = N_{USD} * (1 + r_{USD} * T) * Fwd$$

$$Spot * \frac{(1 + r_{JPY} T)}{(1 + r_{USD} T)} = Fwd$$

09/10/14

3 Forward pricing

Using Foreign Exchange as our example we have showed how the price of forwards can be determined using arbitrage (in that example, covered interest rate arbitrage). The value of an FX forward can be determined from the benefit of owning JPY (ie. the JPY interest) versus the loss from not owning USD (the USD interest). From this we can define a general principle for forward pricing as follows.

Proposition 1. *The value of a forward is derived from the value of the underlying asset together with the gains/losses associated with holding/not holding the asset until expiry.*

Fwd=S+(gains from holding currency)-(losses from not holding asset)

3.1 Equity futures

We will demonstrate this principle using equity futures. First, we need to explain the concept of a dividend.

Definition 2. *Owning a share gives the shareholder the right to a* dividend payment *at certain intervals during the year. This payment is usually made in cash.*

A dividend payment rewards long-term investors in a stock and enables them to earn money from holding stocks without having to close their position. Large stable companies try to give regular divident payments to stockholders. Growing companies often don't give dividends but expect the stockholders to benefit from the growth of the stock. Microsoft famously gave no dividends in the 90's although in recent years they have started paying dividends.

Example 3. *Suppose UCL shares are trading at* 10*GBP and interest rates are at* 5%*. UCL pays a dividend of* 0.50 *GBP in* 6 *months time. What is the value of the* 1 *year future on UCL?*

$$UCL(1y) = UCL(0) * (1 + 5\%) - 0.5 * (1 + 5\%/2) = 9.9875$$

Holding shares- get dividends
Why hold shares? Long term source of finance
Utility firms want long term shareholders ($ their pensions)

Interest rate compounded monthly ≠ yearly interest

4 Interest rates

4.1 Compound interest

Suppose you invest a sum P at an (annual) interest rate of r, compounded m times a year. This means that simple interest of r/m is charged m times per year, with the sums being compounded. So if you invest principal P, at the end of one year you will have

$$P(1+r/m)^m.$$

$P(1+r/m)^{mt}$ ← for multiple years

The total interest you receive is $P(1+r/m)^m - P$, which is equivalent to an *effective interest rate* of

$$(1+r/m)^m - 1.$$

Example 4. *A credit card company charges* 12.9%, *compounded monthly. This is equivalent to an annual rate of*

$$\left(1+\frac{0.129}{12}\right)^{12} - 1 = 1.01075^{12} - 1 = 0.1369\cdots.$$

Thus the effective annual rate is almost 13.7%.

As we compound more frequently, the effective rate increases. As $m \to \infty$, we get

$$(1+r/m)^m \to e^r.$$

If we invest $V(0) = P$ at rate r *compounded continuously*, this means that after t years we have

$$V(t) = Pe^{rt}.$$

(We do not need t to be an integer.) So, taking $t = 1$, the *effective rate* is

$$\frac{\text{total interest}}{\text{initial investment}} = \frac{V(1)-V(0)}{V(0)} = \frac{Pe^{rt}-P}{P} = e^r - 1.$$

4.2 Present value

Suppose that we are free to borrow and lend money at interest rate r, compounded m times a year. What is the value to us today of a payment $V(t)$ that we will receive at time t? We must *discount* the sum $V(t)$ to allow for the passage of time: it will be worth less at time $t > 0$ than it is now.

A sum $V(0)$ invested today will be worth

$$V(0)(1 + r/m)^{mt}$$

at time t. This equals $V(t)$ when

$$V(0) = (1 + r/m)^{-mt}V(t),$$

which is called the *present value* or *discounted value* at time 0 of the payment $V(t)$ at time t. Similarly, the present value of a payment $V(t)$ at time t, with interest rate r compounded continously, is

$$V(t)e^{-rt}.$$

The present value of a sequence of payments made over time is obtained by adding up the discounted values of the separate payments.

The value e^{-rt} is known as the *discount factor*.

4.3 Bonds

A *bond* is a financial instrument that pays a regular sum, known as the *coupon* for a specified period of time, and an additional final payment, known as the *principal*. The final payment is made at the same time as the last coupon payment. A *zero-coupon* bond has only a final payment.

The *yield* of a bond is its rate of return. Typically, the yield of a bond with given face value and coupon will vary with its time to maturity. We can draw a graph of bond yield against maturity: this is know as a *yield curve*.

Contingent claims e.g. option contingent on option being in the money

5 Representing assets and claims

5.1 Contingent claims

What are the basic ingredients we need to construct an option?

- A set Ω of "states of natures" or *contingencies*: these represent the possible outcomes or "possible worlds".
- A schedule of payments or claim $X\colon \Omega \to \mathbb{R}$. In other words, a function that tells us the financial consequences of each possible outcome $\omega \in \Omega$.

Example 5. *Call option with strike price K and expiration date T. Let us take $\Omega = \mathbb{R}^+$, where ω corresponds to the event "$S(T) = \omega$". Then*

$$\begin{aligned} X(\omega) &= (\omega - K)^+ \\ &= \max\{\omega - K, 0\} \end{aligned}$$

Example 6. *Consider a stock that has value 1 today, and may go either up or down tomorrow. Let us have $\Omega = \{\text{Heads}, \text{Tails}\}$, with*

$$\begin{aligned} X(\text{H}) &= 1.01 \\ X(\text{T}) &= 0.99. \end{aligned}$$

(We could write $\omega_1 = \text{H}$ and $\omega_2 = \text{T}$.)

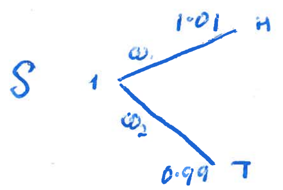
What about if we just hold onto a sum in cash? Let B be the payoff from doing this, so

$$\begin{aligned} B(\text{H}) &= 1 \\ B(\text{T}) &= 1; \end{aligned}$$

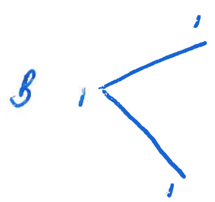
whatever we do, we end up with the same amount.

Suppose we buy M units of the stock. Let G be our overall position tomorrow, so G is the value of M units of stock, minus the amount M that we have spent. (In other words, G is the amount we have gained.) We have

$$\begin{aligned} G(H) &= M(1.01 - 1) = 0.01M \\ G(T) &= M(0.99 - 1) = -0.01M. \end{aligned}$$



one or other will happen



We can write this example in vector notation: let

$$\mathbf{X} = (X(H), X(T)) = (1.01, 0.99)$$

and

$$\mathbf{B} = (B(H), B(T)) = (1, 1).$$

Then

$$\begin{aligned}
\mathbf{G} &= (G(H), G(T)) \\
&= M\mathbf{X} - M\mathbf{B} \\
&= (0.01M, -0.01M).
\end{aligned}$$

We will either gain or lose. Should we expect each to happen with probability 1/2?

Example 7. *Suppose we have two assets X and Y that behave independently but identically. So*

$$\begin{tikzcd}[row sep=small]
& \mathit{X(1)=1.01} & & & \mathit{Y(1)=1.01} \\
\mathit{X(0)=1} \arrow[ur] \arrow[dr] & & & \mathit{Y(0)=1} \arrow[ur] \arrow[dr] & \\
& \mathit{X(1)=0.99} & & & \mathit{Y(1)=0.99}
\end{tikzcd}$$

We define $\Omega = (HH, HT, TH, TT)$. In vector notation, we have

$$\mathbf{X} = (1.01, 1.01, 0.99.0.99), \qquad \mathbf{Y} = (1.01, 0.99, 1.01, 0.99).$$

Suppose we invest M by diversifying $M/2$ on X and $M/2$ on Y. We end up with gain

$$\begin{aligned}
\mathbf{G} &= \frac{M}{2}(\mathbf{X} + \mathbf{Y}) - M\mathbf{B} \\
&= M(0.01, 0, 0, -0.01).
\end{aligned}$$

Should we expect each outcome to be equally likely? It looks as though diversification has reduced the risk.

5.2 The portfolio space

The basic setup we have is as follows.

- A set $\Omega = (\omega_1, \ldots, \omega_m\}$ of basic outcomes or states of nature. Ω is often known as the *sample space.*
- Probabilities on the states: $\mathbb{P}(\omega_i) \geq 0$, $\sum_{i=1}^m \mathbb{P}(\omega_i) = 1$.
- Contingent claims (or, in probabilistic language, *random variables*) are functions $X \colon \Omega \to \mathbb{R}$. We say that X has *value* $X(\omega_i)$ on contingency ω_i.

We will often use the vector space notation: X can be represented by the vector

$$\mathbf{X} = (X(\omega_1), \ldots, X(\omega_m)) \in \mathbb{R}^m.$$

The *riskless asset* is

$$\mathbf{1} = (1, \ldots, 1),$$

as the value of the riskless asset does not depend on ω (although it may depend on time).

Often we will have more than one asset in play, so we can produce *portfolios of assets.* For instance, suppose we have assets with names $1, \ldots, N$. These have prices $S_1(0), \ldots, S_N(0)$ now, and prices $S_1(T, \omega), \ldots, S_N(T, \omega)$ at time T if we are in state ω. Using the vector notation, we get prices

$$\mathbf{S}_1(T), \ldots, \mathbf{S}_N(T)$$

at time T, where these N vectors are in $\mathbb{R}^m$.

The vectors $\mathbf{S}_1(T), \ldots, \mathbf{S}_N(T)$ do not need to be independent. We define the *portfolio space* to be

$$\operatorname{lin}\{\mathbf{S}_1(T), \ldots, \mathbf{S}_N(T)\}.$$

Why does this make sense? Suppose we have H_1 units of S_1, H_2 units of S_2 and so on. We allow H_i to be either positive or negative: if $H_i > 0$ we say that we are *long* in S_i, while if $H_i < 0$ we say that we are *short* in S_i. We can write our portfolio of assets as

$$\mathbf{H} = (H_1, \ldots, H_N).$$

[This is a vector in $\mathbb{R}^N$, while the prices $\mathbf{S}_1(T), \ldots, \mathbf{S}_N(T)$ are in $\mathbb{R}^m$.]

What is the value $V_T(\mathbf{H})$ at time T of the resulting portfolio? For each state ω, we have

$$V_T(\mathbf{H}, \omega) = \sum_{i=1}^{N} H_i S_i(T, \omega).$$

In vector notation,

$$\mathbf{V}_T(\mathbf{H}) = \sum_{i=1}^{N} H_i \mathbf{S}_i(T).$$

The set of vectors that we can achieve in this way is given by the portfolio space.

Question: What is the value of the portfolio $\mathbf{H}$ at time 0? Should it be $\sum_{i=1}^{N} H_i S_i(0)$?

T (expiration occurs at time T

$$\begin{tikzcd}[row sep=small] & S_H \quad \max(S_H - K, 0) = S_H - K \\ S(0) \arrow[ur] \arrow[dr] & \\ & S_L \quad \max(S_L - K, 0) = 0 \text{ since } K > S_L \end{tikzcd}$$

5.3 Replicating an option: an important example

Let us return to the example of a call option X with strike price K and underlying stock S. Suppose that the stock is worth $S(0)$ today, and either goes up to S_H or down to S_L tomorrow.

$$\begin{tikzcd}[row sep=small] & S_H = S(T, \text{High}) \\ S(0) \arrow[ur] \arrow[dr] & \\ & S_L = S(T, \text{Low}) \end{tikzcd}$$

We assume that $S_L < K < S_H$. Recall that the value of the option X tomorrow is $(S_L - K)^+ = 0$ if the stock goes down and $(S_H - K)^+ > 0$ if the stock goes up. In vector notation, we have

$$\mathbf{X} = (0, S_H - K).$$

Now consider a portfolio composed of the stock S and the riskless asset. Does the call option $\mathbf{X}$ lie in the portfolio space?

In vector notation, we have

$$\mathbf{S} = (S_L, S_H), \qquad \mathbf{1} = (1, 1),$$

and the question is whether $\mathbf{X} \in \text{lin}\{\mathbf{1}, \mathbf{S}\}$. In other words, we want* to solve the equation

$$u\mathbf{1} + v\mathbf{S} = \mathbf{X},$$

where we write $\mathbf{H} = (u, v)$ for our portfolio. This is equivalent to the simultaneous equations

$$\begin{aligned} u + vS_L &= 0 \\ u + vS_H &= S_H - K, \end{aligned}$$

which are solved by

$$v = \frac{S_H - K}{S_H - S_L}$$

* i.e. can we recreate this option within this portfolio space

and

$$u = -vS_L = -S_L\frac{S_H - K}{S_H - S_L}.$$

It follows that *we can replicate the option by a combination of the stock and the riskless asset.* This is very significant.

Example 8. *Consider the following asset.*

$$\begin{tikzcd}[row sep=small] & S_H = 2 \\ S(0) = 1 \arrow[ur] \arrow[dr] & \\ & S_L = 1/2 \end{tikzcd}$$

A call option with strike price $K = 1/2$ will have value 1.5 if the stock goes up and 0 if the stock goes down. How do we replicate this? We have

$$\mathbf{S} = (1/2, 2), \qquad \mathbf{1} = (1, 1),$$

and

$$\mathbf{X} = (0, 3/2) = -\frac{1}{2}\mathbf{1} + \mathbf{S}.$$

So the call can be replicated at time 1 by a portfolio consisting of $-1/2$ unit of the riskless asset and one unit of the stock.

6 Arbitrage

6.1 An arbitrage argument

Suppose we have assets S_1 and S_2, and that S_1 is always worth the same as S_2 at time T. In other words, we have

$$S_1(T, \omega) = S_2(T, \omega)$$

for every state $\omega \in \Omega$. Shouldn't S_1 and S_2 have the same price today?

Let us suppose that S_1 and S_2 have different prices, say

$$p(S_1) < p(S_2),$$

where we write $p(S)$ for the price today (at time $t = 0$) of asset S. We can use this situation to make money in the following way: at time 0 we buy a unit of S_1 and sell (or go short) a unit of S_2. We receive the price difference $D = p(S_2) - p(S_1)$ in cash, which we place in a bank. We now do nothing until time T. At time T, we sell our unit of S_1, and use it to purchase one unit of S_2, leaving us with no holding in either asset. Since we have $S_1(T, \omega) = S_2(T, \omega)$ for every state ω, the sale of S_1 pays for the purchase of S_2. We are left with the money in the bank, which will by now have increased to $(1 + r)D$, where rD is the quantity of interest we have earned.

So if $p(S_1) < p(S_2)$ (or similarly if $p(S_1) > p(S_2)$) we are able to make a profit without any risk at all: this process is known as *arbitrage*.

We can express the same argument in another way, by considering portfolios. Suppose that S_1 and S_2 are worth the same in all states at time T, but that $p(S_1) < p(S_2)$. Let $\mathbf{H}$ be the portfolio consisting of 1 unit of S_1, -1 units of S_2 and $D = p(S_2) - p(S_1)$ units of cash. The price of $\mathbf{H}$ at time 0 is

$$p(\mathbf{H}) = p(S_1) - p(S_2) + D = 0,$$

while we have

$$V_T(\mathbf{H}, \omega) = S_1(T, \omega) - S_2(T, \omega) + (1 + r)D = (1 + r)D > 0,$$

for every $\omega \in \Omega$. So we have a portfolio with cost 0 at time 0, but positive value in every state at time T. This is clearly a good way to make money!

6.2 Two types of arbitrage

In the example above, we saw a portfolio which guarantees a profit regardless of the state at time T.

Definition 9. *We say that a portfolio* $\mathbf{H}$ *is a* sure-thing arbitrage *if*

$$V_0(\mathbf{H}) = 0$$

and, for every $\omega \in \Omega$,

$$V_T(\mathbf{H}, \omega) > 0.$$

As we have seen above, a sure-thing arbitrage offers us a way to make a guaranteed with no risk. It may also happen that we have a possibility of profit without risk, as follows.

Definition 10. *We say that a portfolio* $\mathbf{H}$ *is an* arbitrage opportunity *if*

$$\begin{aligned} &V_0(\mathbf{H}) = 0, \\ &V_T(\mathbf{H}) \geq 0 \qquad \text{for every } \omega \in \Omega \end{aligned}$$

and

$$V_T(\mathbf{H}, \omega) > 0 \qquad \text{for some } \omega \in \Omega.$$

It is clear that if there is no arbitrage opportunity (for a given set of assets) then there is no sure-thing arbitrage. Both assumptions imply that price is a linear functional.

Lemma 11. *Suppose we have assets with claims* $\mathbf{X}, \mathbf{X}_1, \ldots, \mathbf{X}_k$. *If* $\mathbf{X} = \sum_{i=1}^k \lambda_i \mathbf{X}_i$ *and there is no sure-thing arbitrage then the current price satisfies* $p(\mathbf{X}) = \sum_{i=0}^k \lambda_i p(\mathbf{X}_i)$.

Proof. Consider the portfolio $\mathbf{H}$ consisting of $p(\mathbf{X}) - \sum_{i=1}^k \lambda_i p(\mathbf{X}_i)$ units of the riskless asset, -1 units of X, and λ_i units of X_i for each i. If $p(\mathbf{X}) > \sum_{i=1}^k \lambda_i p(\mathbf{X}_i)$ then $\mathbf{H}$ is a sure-thing arbitrage. If $p(\mathbf{X}) < \sum_{i=1}^k \lambda_i p(\mathbf{X}_i)$ then $-\mathbf{H}$ is a sure-thing arbitrage. □

6.3 Valuing a call option

Let us return to the two-state model of a call option X with strike price K and underlying stock S. The stock is worth $S(0)$ today, and either goes up to S_H or down to S_L tomorrow.

$$\begin{tikzcd}
 & S_H = S(T, \text{High}) \\
S(0) \arrow[ur] \arrow[dr] & \\
 & S_L = S(T, \text{Low})
\end{tikzcd}$$

Assuming $S_L < K < S_H$, we found that we could replicate the call option by

$$\mathbf{X} = u\mathbf{1} + v\mathbf{S},$$

where

$$u = -S_L \frac{S_H - K}{S_H - S_L}, \qquad v = \frac{S_H - K}{S_H - S_L}.$$

If the interest rate is r we get

$$p(\mathbf{1}) = \frac{1}{1+r}, \qquad p(S) = S(0),$$

and so

$$\begin{aligned} p(X) &= \frac{u}{1+r} + vS(0) \\ &= \frac{S_H - K}{S_H - S_L}\left(S(0) - \frac{S_L}{1+r}\right) \end{aligned} \tag{1}$$

It is helpful to write this in terms of present value. We adopt the following convention.

Definition 12. *The present value of a payment R at a fixed point in the future is denoted R^*.*

In this case, $S_H^* = S_H/(1+r)$, and so on. Using this notation, we rewrite (1) as

$$p(X) = \frac{S_H^* - K^*}{S_H^* - S_L^*}\left(S(0) - S_L^*\right) = \frac{S(0) - S_L^*}{S_H^* - S_L^*}\left(S_H^* - K^*\right).$$

6.4 Put-Call Parity

If we assume that there is no arbitrage opportunity, we can also derive a useful (and quite general) relationship between prices of put and call options.

Theorem 13 (Put-Call Parity). *Let C be a European call option and P a European put option on a stock S, and suppose that both options have the same strike price K and expiry date T. If there is no sure-thing arbitrage then*

$$P(0) - C(0) = K^* - S(0).$$

Proof. We define two portfolios as follows. Let $\mathbf{H}_1$ consist of one unit of stock and one put option, and let $\mathbf{H}_2$ consist of K^* of the riskless asset and one call option. Recall that, at time T, the call option is worth

$$C(T) = \max\{S(T) - K, 0\}$$

and the put option is worth

$$P(T) = \max\{K - S(T), 0\}.$$

At time T, the value of $\mathbf{H}_1$ is

$$S(T) + P(T) = S(T) + \max\{K - S(T), 0\} = \max\{S(T), K\},$$

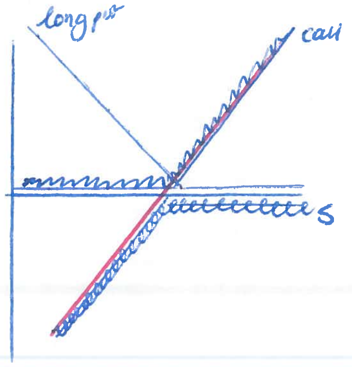
and the value of $\mathbf{H}_2$ is

$$K + C(T) = K + \max\{S(T) - K, 0\} = \max\{S(T), K\}$$

Thus $\mathbf{H}_1$ and $\mathbf{H}_2$ have the same value at time T (in every state). The assumption of no sure-thing arbitrage implies that at time 0 we have $p(\mathbf{H}_1) = p(\mathbf{H}_2)$, and so $S(0) + P(0) = K^* + C(0)$. □

Put-Call Parity only applies to European options. With American options, we may choose to exercise the option before its expiry, so that we would no longer be holding the option at time T.

Remember shortcut to calculate put if you know call



— denotes long call, short put

$H(+1C, -1P) \Leftrightarrow F(K)$

must be careful of present value

K^* present value

16/10/14

7 Basic ideas from probability

7.1 Expected value

Recall that for a finite sample space $\Omega = \{\omega_1, \ldots, \omega_m\}$, we say that $\mathbb{P} : \Omega \to [0, 1]$ is a probability on Ω if $\sum_{i=1}^m \mathbb{P}(\omega_i) = 1$. The probability of $A \subset \Omega$ is then $\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega)$.

A *random variable* or *contingent claim* is a function $X : \Omega \to \mathbb{R}$. The *expected value* of X is then

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[X] &= \sum_{i=1}^m \mathbb{P}(\omega_i) X(\omega_i) \\ &= \mathbf{P} \cdot \mathbf{X}, \end{aligned}$$

where $\mathbf{P} = (\mathbb{P}(\omega_1), \ldots, \mathbb{P}(\omega_m))$.

We have the following familiar examples.

Example 1 (Coin tossing). *Suppose that $\Omega = \{H, T\}$ and $\mathbb{P}(H) = \mathbb{P}(T) = 1/2$ (so we have a* fair coin*). If we make a bet that pays 2 on heads and 0 on tails then the payoff is*

$$\mathbf{X} = (X(H), X(T)) = (2, 0)$$

and

$$\mathbb{E}[X] = \frac{1}{2} 2 + \frac{1}{2} 0 = 1.$$

Example 2 (Roll of a die). *We have $\Omega = \{\omega_1, \ldots, \omega_6\}$, and (assuming a fair die) $\mathbb{P}(\omega_i) = 1/6$ for every i. Suppose the payoff is defined by $X(\omega_i) = i$ for every i. Then*

$$\begin{aligned} \mathbb{E}[X] &= \frac{1}{6}(1 + 2 + \cdots + 6) \\ &= 3.5 \end{aligned}$$

A more interesting example arises when we consider the two-state model of a call option.

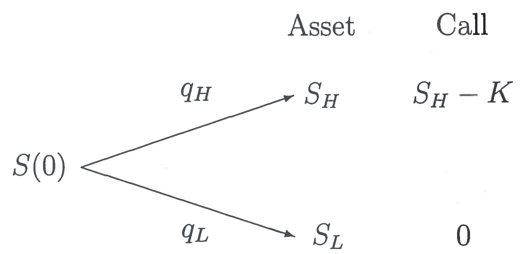
Why talk about probability when we don't need it to value options?

Fake probability - synthetic probability, i.e. risk neutral probability

Example 3 (Valuation of a call). *Recall our two-state model of a call option. A stock is worth $S(0)$ today, and either goes up to S_H or down to S_L at time T. We consider a call option X with strike price K, where $S_L < K < S_H$, and so claim $(S(T)-K)^+$ at time T. We used replication of the call to obtain a no-arbitrage valuation of*

$$\frac{S(0) - S_L^*}{S_H^* - S_L^*}(S_H^* - K^*).$$

Suppose that $S_L^* < S(0) < S_H^*$. (Why is this a reasonable assumption?) We can describe the situation as follows.



We can define 'high' and 'low' probabilities by present value

$$q_H = \frac{S(0) - S_L^*}{S_H^* - S_L^*}$$

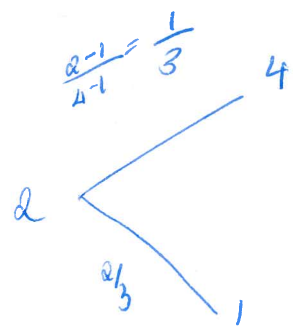
and

$$q_L = 1 - q_H = \frac{S_H^* - S(0)}{S_H^* - S_L^*}.$$

This gives a probability measure $\mathbb{Q}$ with

$$\mathbb{E}_{\mathbb{Q}}[X^*] = q_H(S_H^* - K^*) + q_L \cdot 0,$$ since call $(S(T)-K)^+$

which is equal to our valuation of the call option.



The probability measure $\mathbb{Q}$ in Example 3 is called the *synthetic probability*, and has some important properties. The expected value of the stock S at time T is

$$\mathbb{E}_{\mathbb{Q}}[S(T)] = q_L S_L + q_H S_H.$$

Since q_L and q_H are written in terms of present value, it makes sense to do the same for $S(T)$. We get

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}[S(T)^*] &= q_L S_L^* + q_H S_H^* \\ &= \frac{S_H^* - S(0)}{S_H^* - S_L^*} S_L^* + \frac{S(0) - S_L*}{S_H^* - S_L^*} S_H^* \\ &= S(0). \end{aligned}$$

This is important enough to put in a box:

$$\boxed{\mathbb{E}_{\mathbb{Q}}[S(T)^*] = S(0)}$$

In other words, after discounting to today's value, the expected price at time T with respect to the synthetic probability is the same as the price today.

Now consider a portfolio H composed from u units of the riskless asset and v units of stock, so

$$\mathbf{H} = u\mathbf{1} + v\mathbf{S}(T).$$

The expected value of H with respect to the synthetic probability is given by

$$\begin{aligned} \mathbb{E}_Q[V_T(H)] &= u + v\mathbb{E}_{\mathbb{Q}}[S(T)] \\ &= u + v(1+r)S(0), \end{aligned}$$

and so

$$\boxed{\mathbb{E}_{\mathbb{Q}}[V_T(H)^*] = \tfrac{u}{1+r} + vS(0).}$$

In other words, discounting to present value, the expected value of $\mathbf{H}$ with respect to the synthetic probability measure at time $t = T$ is the same as the price of the portfolio at time $t = 0$.

future value of portfolio is same as it is today

7.2 Risk-neutral probabilities

An investor will usually take account of the risk attached to any financial investment. Some investors are *risk-averse*, and will seek to avoid risk, or else expect to be paid for taking on a risk – in other words, if the risk is greater then they will expect a larger expected return. Similarly, investors may desire to gamble, in which case they may be willing to accept a lower expected return in exchange for a higher risk. For instance, in Example 1, a risk-averse player will not pay as much as 1 for a game with expected return 1, while a more speculative player might be willing to pay more than 1.

A player who is willing to enter the game in Example 1 for a fee equal to the expected payoff is said to be *risk-neutral*. Such a player is willing to exchange a payment of 1 with certainty for a payment that has expected value 1.

In Example 3 above, *the market behaves as a risk-neutral agent with respect to the synthetic probability.* In other words, the market prices the call option X as though it were a risk-neutral agent with the synthetic probability. For this reason, the synthetic probability is also known as the *risk-neutral probability*.

Note that the synthetic probability measure $\mathbb{Q}$ is a function of market prices, and does not depend on our beliefs about the future. Usually, when talking about probabilities of future events we have to justify our statements with evidence of what is more or less likely to happen in the future. This is not the case with the synthetic probability: if the synthetic probability exists, it is simply a description of market prices.

7.3 Probability distributions

If X is a random variable defined on $\Omega = \{\omega_1, \dots, \omega_m\}$, the sets

$$(X = x) = \{\omega \in \Omega : X(\omega) = x\},$$

for $x \in \mathbb{R}$, partition ω. In order to calculate the expectation $\mathbb{E}_{\mathbb{P}}[X]$ it is enough to know the probability of the event $(X = x)$, which is equal to

$$\mathbb{P}(X = x) = \sum_{\omega \in (X=x)} \mathbb{P}(\omega).$$

The expectation is then

$$\boxed{\mathbb{E}_{\mathbb{P}}[X] = \textstyle\sum_x x\mathbb{P}(X = x).}$$

When calculating $\mathbb{E}_{\mathbb{P}}(X)$, we don't need to look at Ω: it is enough to know just the probability that X takes a specific value.

When X can take infinitely many values, this approach needs some modification (we may have $\mathbb{P}(X = x) = 0$ for every x). But it is enough to know the probability that X lies in any given range.

Definition 4. *The* distribution function F_X *of* X *is defined by* cumulative distribution function

$$F_X(x) = \mathbb{P}(X \leq x).$$

We can recover $\mathbb{P}(X = x)$ from the distribution function, since $\mathbb{P}(X = x) = F_X(x) - \lim_{t\to 0+} F_X(x - t)$.

If F_X is sufficiently nice, then we can write it as an integral.

Definition 5. *We say that* $f : \mathbb{R} \to [0, \infty)$ *is a* probability density function *for* X *if*

$$F_X(x) = \int_{-\infty}^{x} f(t)d(t).$$

for all $x \in \mathbb{R}$.

In this case, the expectation of X is

$$\boxed{\mathbb{E}[X] = \textstyle\int_{-\infty}^{\infty} x f(x)dx,}$$

as long as the integral is defined. More generally, for any reasonable function $g : \mathbb{R} \longrightarrow \mathbb{R}$,

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx.$$

For instance $\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f(x) dx$.

[All of this can be set up more formally in the context of *Measure Theory*, when the expectation can be defined as $\int X d\mathbb{P}$. We could also give a definition of a "reasonable" function – for the moment, it is enough to say that all functions in this course (and all functions that you are likely to meet) are reasonable.]

Example 6 (Standard normal distribution). *A random variable is said to be* standard normal*, or to be $N(0, 1)$, if it has density function*

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Of course, to check that this is *a density function, we need to check that*

$$\int_{-\infty}^{\infty} \phi(x) dx = 1.$$

Example 7 (Warning about expectation). *Consider the function*

$$f(x) = \frac{1}{\pi} \frac{1}{1 + x^2}.$$

Since $\int_{-\infty}^{\infty} f(x) dx = 1$ (exercise!), this is a distribution function (known as the Cauchy distribution*). But*

$$\int_{-\infty}^{\infty} x f(x) dx$$

is not defined, so a random variable with the Cauchy distribution doesn't have an expectation.

8 Risk-neutral measures and arbitrage

8.1 Introduction

Consider a one-period model of a market with $\Omega = \{\omega_1, \ldots, \omega_m\}$ and assets $S_1, \ldots, S_N$. The current price of the ith asset is $S_i(0)$, while the vector of prices at time $t = 1$ is $\mathbf{S}_i = (S_i(\omega_1, 1), \ldots, S_i(\omega_m, 1))$.

Definition 8. *A probability measure Q on Ω is risk-neutral if*

$$Q(\omega_i) > 0 \qquad \forall i$$

← always check this
c/w can hide an arbitrage opportunity

and

expected value under Q of S_n (present value) is same as value today →

$$\mathbb{E}_Q[S_n(1)^*] = S_n(0) \qquad \forall n.$$

As we shall see, there is a close relationship between risk-neutral measures and the absence of arbitrage. An important role in our discussion will be played by complete markets.

Definition 9. *We say that a market with $\Omega = \{\omega_1, \ldots, \omega_m\}$ and assets $S_1, \ldots, S_N$ is* complete *if*

$$\text{lin}\{\mathbf{S}_1, \ldots, \mathbf{S}_N\} = \mathbb{R}^m.$$

Thus a complete market is one in which every possible contingent claim can be replicated by a suitable portfolio. Note that we could have $N > m$, in which case the replicating portfolio will not be unique.

It will be useful to define the following contingent claims.

Definition 10. *We write e_i for the asset with contingent claim*

$$\mathbf{e}_i = (0, \ldots, 0, \overset{i\text{th asset}}{1}, 0, \ldots, 0),$$

where $\mathbf{e}_i(\omega_i) = 1$ and $\mathbf{e}_i(\omega_j) = 0$ for $j \neq i$. The assets e_i are known as Arrow-Debreu securities. Also a basis set.

incomplete leads to arbitrage

$e_1 \ (1, 0, 0 \ldots)$
$e_2 \ (0, 1, 0 \ldots)$
$e_3 \ (0, 0, 1, 0 \ldots)$

8.2 Complete markets with no arbitrage

Suppose that our model is complete and has no arbitrage opportunities. It follows that the time $t = 0$ price of each $\mathbf{e}_i$ is positive (why?), i.e.

$$p(\mathbf{e}_i) > 0.$$

Now suppose the interest rate is r, so

$$p(\mathbf{1}) = \frac{1}{1+r}.$$

Let us define

$$q_i = (1+r)p(\mathbf{e}_i) > 0.$$

We can replicate the riskless asset by Arrow-Debreu securities:

$$\mathbf{1} = \sum_{i=1}^{m} \mathbf{e}_i.$$

So, by our assumption of no-arbitrage, we have

$$p(\mathbf{1}) = \sum_{i=1}^{m} p(\mathbf{e}_i).$$

Thus

$$\sum_{i=1}^{m} q_i = 1$$

and

$$q_i > 0 \qquad \forall i.$$

In other words, $\mathbb{Q} = (q_1, \ldots, q_m)$ gives a probability measure on Ω. We shall prove that $\mathbb{Q}$ is a risk-neutral measure.

Suppose that $\mathbf{X} = (x_1, \ldots, x_m)$ is any contingent claim. Then the portfolio with x_i units of the Arrow-Debreu security e_i for each i has price

$$\begin{aligned} p(X) &= p\left(\sum_{i=1}^{m} x_i \mathbf{e}_i\right) \\ &= \sum_{i=1}^{m} x_i p(\mathbf{e}_i) \\ &= \sum_{i=1}^{m} q_i \frac{x_i}{1+r}, \end{aligned}$$

need to show value of portfolio = present value since $q_i = (1+r)p(\mathbf{e}_i)$. It follows that

$$p(X) = \mathbb{E}_{\mathbb{Q}}[X^*].$$

In particular, for $1 \leq i \leq N$, we have

$$S_n(0) = \mathbb{E}_{\mathbb{Q}}[S_n(1)^*].$$

So $\mathbb{Q}$ is a risk-neutral measure. Thus a market without arbitrage opportunities prices assets as though it is a risk-neutral agent using the probability measure $\mathbb{Q}$.

The risk-neutral probability measure was defined by the *state-prices*

$$q_i = p((1+r)\mathbf{e}_i).$$

In this case (the complete market), there is a unique risk-neutral probability: since the Arrow-Debreu securities can be replicated, we have

$$p(\mathbf{e}_i) = \mathbb{E}_q[e_i(1)^*] = \frac{1}{1+r}\mathbb{E}_q[e_i(1)] = \frac{q_i}{1+r}.$$

Since (by no-arbitrage) $p(\mathbf{e}_i)$ is determined by the time 0 prices of the X_i, it follows that the q_i are also determined by these prices.

Note that this argument doesn't work when the Arrow-Debreu securities are not replicable.

8.3 Markets with a risk-neutral measure

We now drop the assumption that there is no arbitrage opportunity, and instead assume that there is a risk-neutral measure. We will prove some interesting consequences, beginning with the fact that any contingent claim has a unique price. *Note that we do* not *assume that the market is complete.*

so all banks have same price

Theorem 11 (Law of One Price). *In a market with a risk-neutral measure $\mathbb{Q}$, all portfolios that replicate a contingent claim $\mathbf{X}$ have the same time 0 price.*

Proof. Suppose that $\mathbf{X}$ is replicated by portfolios $\mathbf{H} = (H_1, \ldots, H_N)$ and $\mathbf{H}' = (H'_1, \ldots, H'_N)$. Then

$$\mathbb{E}_{\mathbb{Q}}[X(1)^*] = \mathbb{E}_{\mathbb{Q}}\Big[\sum_{i=1}^{N} H_i S_i(1)^*\Big]$$

risk neutral measure $\mathbb{Q}$

$$= \sum_{i=1}^{N} H_i \mathbb{E}_{\mathbb{Q}}[S_i(1)^*]$$

$$= \sum_{i=1}^{N} H_i S_i(0).$$

So

$$\mathbb{E}_{\mathbb{Q}}[X(1)^*] = V_0(H)$$

and, similarly,

$$\mathbb{E}_{\mathbb{Q}}[X(1)^*] = V_0(H').$$ = Value at time 0 of H'

Thus both replicating portfolios have the same time 0 value. □

The Law of One Price implies that there is no sure-thing arbitrage. But we can prove something stronger.

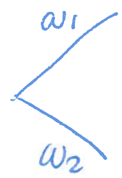
Theorem 12. *In a market with a risk-neutral measure $\mathbb{Q}$, there is no arbitrage opportunity.*

Proof. Suppose that the portfolio $\mathbf{H} = (H_1, \ldots, H_N)$ is an arbitrage opportunity, so for some j we have

$$V_0(\mathbf{H}) = 0 \quad \text{Value today = nothing}$$
$$V_1(\mathbf{H}, \omega_i) \geq 0 \quad \forall i$$
$$V_1(\mathbf{H}, \omega_j) > 0.$$

Why buy option instead of replicating?

You may think it is worth more - believe a price is going up



Then

$$\mathbb{E}_{\mathbb{Q}}[V_1(\mathbf{H})] = \sum_{i=1}^{N} \mathbb{Q}(\omega_i) V_1(\mathbf{H}, \omega_i) \geq Q(\omega_j) V_1(\mathbf{H}, \omega_j) > 0.$$

But, as $\mathbb{Q}$ is risk-neutral,

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}[V_1(H)] &= \mathbb{E}_{\mathbb{Q}}\Big[\sum_{i=1}^{N} H_i S_i(1)\Big] \\ &= \sum_{i=1}^{n} H_i \mathbb{E}_{\mathbb{Q}}[S_i(1)] \\ &= (1+r) \sum_{i=1}^{N} H_i S_i(0) \\ &= (1+r) V_0(\mathbf{H}) \\ &= 0, \end{aligned}$$

which is a contradiction. $\square$

In the argument above we used the fact that $Q(\omega_j) > 0$ to show that an arbitrage opportunity would have (strictly) positive expected value at time 1 but 0 value at time 0. The argument would not work if we were allowed to have states j with probability 0 (as we could 'hide' an arbitrage opportunity using those states).

In a complete market with no arbitrage opportunities, we know that there is a unique risk-neutral probability measure. When the market is not complete there may be multiple risk-neutral probability measures. However, the measures must agree on certain claims.

Theorem 13. *If $\mathbb{Q}$ and $\mathbb{Q}'$ are risk-neutral probability measures and X is replicable then*

$$\mathbb{E}_{\mathbb{Q}}[X(1)^*] = \mathbb{E}_{\mathbb{Q}'}[X(1)^*].$$

Proof. If X is replicated by $\mathbf{H}$, i.e.

$$\mathbf{X} = \sum_{i=1}^{N} H_i \mathbf{S}_i,$$

then

$$
\begin{aligned}
\mathbb{E}_{\mathbb{Q}}[X(1)^*] &= \sum_{i=1}^{N} H_i \mathbb{E}_{\mathbb{Q}}[S_i(1)^*] \\
&= \sum_{i=1}^{N} H_i S_i(0) \\
&= V_0(\mathbf{H}).
\end{aligned}
$$

Similarly, we have

$$\mathbb{E}_{\mathbb{Q}'}[X(1)^*] = V_0(\mathbf{H}),$$

and the theorem follows immediately. □

Notice that if we know that X is replicable, and there is a risk-neutral measure $\mathbb{Q}$, then

$$\mathbb{E}_{Q}[X(1)^*] = p(X).$$

In other words, if we have a risk-neutral measure then we can value every replicable asset by taking expectation. We don't need to work with a replicating portfolio.

We have drawn a number of conclusions from the existence of a risk-neutral measure. So far we only know that a risk-neutral measure exists when the market is complete. However, risk-neutral measures can also exist in incomplete markets.

Our next target is to prove the following result.

Theorem 14 (No-Arbitrage Theorem). *In any one-period model with a finite Ω, the following are equivalent:*

1. *There is a risk neutral measure.*
2. *There are no arbitrage opportunities.*

$\Updownarrow$ if and only if

On every past paper

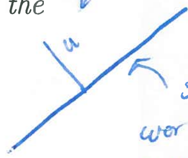
9 The No-Arbitrage Theorem

9.1 Hyperplanes and convex sets

Recall that if $\Omega = \{\omega_1, \ldots, \omega_k\}$, then our claims (or random variables) can be written as vectors in $\mathbb{R}^k$, or equivalently $\mathbb{R}^\Omega$. We want to prove a result about risk-neutral measures and the absence of arbitrage opportunities, but first we will need a geometric fact.

u is perp. to hyperplane

Definition 15. *The* hyperplane $H_{\mathbf{u},p}$ *in* $\mathbb{R}^k$ *is the set of solutions to the equation*



separates world in to 2 bits

$$\langle \mathbf{x}, \mathbf{u} \rangle = p,$$

where $u \in \mathbb{R}^k$ *and* $p \in \mathbb{R}$.

The hyperplane $H_{\mathbf{u},p}$ separates $\mathbb{R}^k$ into two half-spaces, given by

$$\{\mathbf{x} \in \mathbb{R}^k : \langle \mathbf{x}, \mathbf{u} \rangle > p\}$$

and

$$\{\mathbf{x} \in \mathbb{R}^k : \langle \mathbf{x}, \mathbf{u} \rangle < p\}.$$

Note that $\mathbf{u}$ is perpendicular to the hyperplane, since if $\mathbf{x}_1, \mathbf{x}_2 \in H_{\mathbf{u},p}$ then

$$\langle \mathbf{u}, \mathbf{x}_1 - \mathbf{x}_2 \rangle = \langle \mathbf{u}, \mathbf{x}_1 \rangle - \langle \mathbf{u}, \mathbf{x}_2 \rangle = p - p = 0.$$

We will need to recall two more definitions.

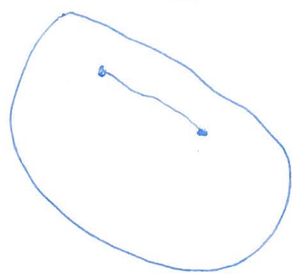
Definition 16. *Let W be a subspace of* $\mathbb{R}^k$. *The subspace* orthogonal *to W is*

$$W^\perp = \{\mathbf{x} \in \mathbb{R}^k : \langle \mathbf{x}, \mathbf{y} \rangle = 0 \quad \forall \mathbf{y} \in W\}.$$

Definition 17. *A subset C of* $\mathbb{R}^k$ *is* convex *if for every* $\mathbf{x}, \mathbf{y} \in C$ *and* $0 \leq \lambda \leq 1$, *we have*

$$\lambda \mathbf{x} + (1 - \lambda)\mathbf{y} \in C.$$

every x, y line lies in C



9.2 The Separating Hyperplane Theorem

Our main tool will be the following.

Theorem 18 (Separating Hyperplane Theorem). *Let W be a subspace of $\mathbb{R}^k$ and let C be a bounded, closed convex set. If*

$$C \cap W = \emptyset$$

then there is a vector $\mathbf{u} \in W^{\perp}$ such that

$$\langle \mathbf{u}, \mathbf{y} \rangle > 0 \quad \forall \mathbf{y} \in C.$$

Why is this called the Separating Hyperplane Theorem? Note that in the proof above we have

$$\langle \mathbf{u}, \mathbf{w} \rangle = 0 \quad \forall \mathbf{w} \in W$$

and

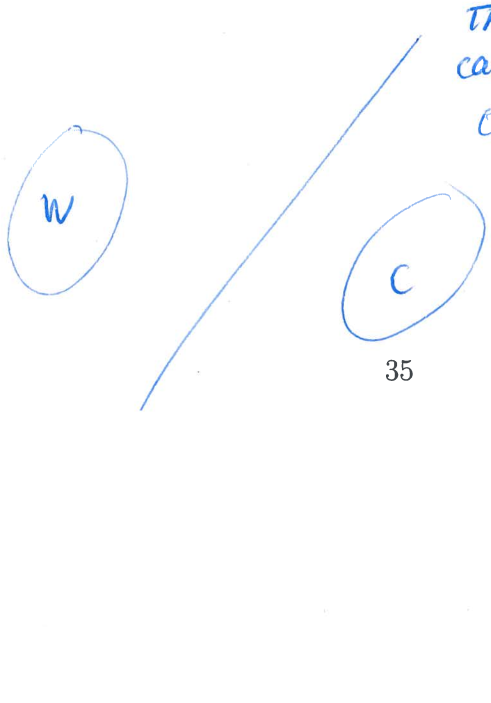
$$\langle \mathbf{u}, \mathbf{x} \rangle \geq d^2 \quad \forall \mathbf{x} \in C.$$

It follows that the hyperplane $H_{\mathbf{u},d^2/2}$ divides $\mathbb{R}^k$ into two half-spaces, each of which contains one of W and X.

No proof, just statement

Used to be examined but not any more

Theorem says:
can draw a line
that separates them



9.3 Gain vectors

Definition 19. *The* gain vector $\mathbf{G}^*(\mathbf{H})$ *of a portfolio* $\mathbf{H}$ *is*

$$\mathbf{G}(\mathbf{H})^* = \sum H_i \Delta \mathbf{S}_i^*,$$

where $\Delta \mathbf{S}_i^*$ *is the vector with coordinates*

$$\Delta S_i(\omega_j)^* = S_i(1, \omega_j)^* - S_i(0).$$

Note that $\Delta \mathbf{S}_i$ is the change between time 0 and time 1 of the (discounted) price of the ith asset. We have

$$\begin{aligned} \mathbf{G}(\mathbf{H}, \omega_j)^* &= \sum H_i \left(S_i(1, \omega_j)^* - S_i(0) \right) \\ &= \sum H_i S_i(1, \omega_j)^* - \sum H_i S_i(0) \\ &= V_1(\mathbf{H}, \omega_j)^* - V_0(\mathbf{H}). \end{aligned}$$

So the gain vector $\mathbf{G}(\mathbf{H})^*$ is the claim corresponding to the gains (or losses) of portfolio $\mathbf{H}$ in the various states.

Example 20. *Suppose we have* N *assets* $S_1, \ldots, S_N$. *If* $\mathbf{H}$ *consists of 1 unit of asset* 1 *then*

$$\mathbf{G}(\mathbf{H})^* = (\Delta S_1(\omega_1)^*, \ldots, \Delta S_1(\omega_N)^*).$$

Example 21. *If* $\mathbf{H}$ *is a portfolio consisting only of the riskless asset then* $\mathbf{G}(\mathbf{H})^* = \mathbf{0}$. *Likewise, if* $\mathbf{H}$ *and* $\mathbf{H}'$ *differ only by the riskless asset then* $\mathbf{G}(\mathbf{H})^* = \mathbf{G}(\mathbf{H}')^*$.

By allowing $\mathbf{H}$ to vary over the collection of all possible portfolios, we get a corresponding collection of gain vectors.

Lemma 22. *Suppose we have* N *assets* $S_1, \ldots S_N$. *The set*

$$W = \{\mathbf{G}(\mathbf{H})^* : \mathbf{H} \in \mathbb{R}^{N+1}\}$$

is a subspace of $\mathbb{R}^\Omega$.

Proof. This follows immediately from the fact that the mapping $\mathbf{H} \to \mathbf{G}(\mathbf{H})^*$ is linear. $\square$

$H = (u\mathbb{1} + rS(0)) = (H_1 S_1(0) + H_2 S_2(0))$

$= (u\mathbb{1} + rS(1)) = (H_1 S_1(1) + H_2 S_2(1))$

So $G = (H_1 (S_1(1) - S_1(0)))$ $(H_2 (S_2(1) - S_2(0)))$

ΔS_i

Note that the gain vector $\mathbf{G}(\mathbf{H})^*$, which is a vector in $\mathbb{R}^\Omega$, can also be thought of as the vector representation of a random variable, which we write as $G(\mathbf{H})^*$

We can rewrite the definition of an arbitrage opportunity in terms of the gain vector.

Lemma 23. *There is an opportunity for arbitrage if and only if there is a portfolio* $\mathbf{H}$ *with*

$$G(\mathbf{H},\omega)^* \geq 0 \quad \forall \omega$$
$$G(\mathbf{H},\omega)^* > 0 \quad \textit{for some } \omega$$

Proof. If $\mathbf{H}$ is an arbitrage opportunity then $\mathbf{G}(\mathbf{H})^*$ satisfies the conditions. On the other hand, if $\mathbf{G}(\mathbf{H})^*$ satisfies the conditions, then adding $-V_0(\mathbf{H})$ cash to the portfolio creates an arbitrage opportunity. □

We are finally ready to prove the No-Arbitrage Theorem.

examinable

Theorem 24 (No-Arbitrage Theorem). *In any one-period model with a finite* Ω*, the following are equivalent:*

1. *There is a risk neutral measure.*
2. *There are no arbitrage opportunities.*

Proof. Suppose that $\Omega = \{\omega_1, \ldots, \omega_k\}$, we have assets $S_1, \ldots, S_N$, and there is no arbitrage opportunity. We show that there is a risk-neutral measure.

Let
$$W = \{\mathbf{G}(\mathbf{H})^* : \mathbf{H} \in \mathbb{R}^{N+1}\}$$
be the subspace of $\mathbb{R}^k$ given by gain vectors of portfolios. Let X^+ be the subset of $\mathbb{R}^k$ given by

positive subset X

$$X^+ = \{\mathbf{x} \in \mathbb{R}^k : x_i \geq 0 \ \forall i, \mathbf{x} \neq \mathbf{0}\},$$

and let P be the subset of $\mathbb{R}^k$ defined by

$$P = \{\mathbf{x} \in \mathbb{R}^k : x_1 + \cdots + x_k = 1, x_i \geq 0 \ \forall i\}.$$ components add up to 1

[Note that P can be thought of as the set of all probability measures on Ω.]

We have
$$W \cap X^+ = \emptyset,$$

since any vector in $W \cap X^+$ is an arbitrage opportunity. [Any such vector would be the gain vector of some portfolio, with every coordinate positive and some coordinate strictly positive.] Since $P \subset X^+$, we have

$$W \cap P = \emptyset.$$

Since P is closed and convex, it follows from the Separating Hyperplane Theorem that there is a vector $\mathbf{u}$ such that

$$\mathbf{u} \in W^\perp$$

and

$$\langle \mathbf{u}, \mathbf{x} \rangle > 0 \ \forall \mathbf{x} \in P.$$

As $\mathbf{e}_i \in P$ for every i, we have $u_i = \langle \mathbf{u}, \mathbf{e}_i \rangle > 0$ for every i. So, defining

$$q_i = \frac{u_i}{u_1 + \ldots + u_k},$$

we see that

$$\sum_{i=1}^{k} q_i = 1$$

and

$$q_i > 0 \ \forall i.$$

We claim that $\mathbb{Q} = (q_1, \ldots, q_k)$ is a risk-neutral probability measure. All that remains is to check that

$$\mathbb{E}_{\mathbb{Q}}[S_i(1)^*] = S_i(0) \ \forall i,$$

or equivalently

$$\mathbb{E}_{\mathbb{Q}}[\Delta S_i^*] = 0 \ \forall i.$$

But this is immediate, as if $\mathbf{H}$ is the portfolio consisting of one unit of S_i, then

$$\mathbb{E}_{\mathbb{Q}}[\Delta S_i^*] = \mathbb{E}_{\mathbb{Q}}[G(\mathbf{H})^*] = \langle \mathbb{Q}, \mathbf{G}(\mathbf{H})^* \rangle = 0,$$

since $\mathbb{Q} = \mathbf{u}/(u_1 + \cdots + u_k)$ and $\mathbf{u} \in W^\perp$. We conclude that $\mathbb{Q}$ is a risk-neutral measure.

For the converse, suppose that $\mathbb{Q}$ is a risk-neutral measure. [Actually, we have proved the converse already, but here it is again in the language of gain vectors.]

If $\mathbf{H}$ is an arbitrage opportunity, then there is j such that

$$G(\mathbf{H},\omega)^* \geq 0 \ \forall \omega$$
$$G(\mathbf{H},\omega_j)^* > 0.$$

So, summing over states ω_i,

$$\mathbb{E}_{\mathbb{Q}}[G(\mathbf{H})^*] = \sum_i q_i G(\mathbf{H},\omega_i) \geq q_j G(\mathbf{H},\omega_j) > 0.$$

But, summing over assets,

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}[G(\mathbf{H})^*] &= \mathbb{E}_{\mathbb{Q}}\big[\sum_n H_n \Delta S_n^*\big] \\ &= \sum_n H_n \mathbb{E}_{\mathbb{Q}}[\Delta S_n^*] \\ &= 0, \end{aligned}$$

which gives a contradiction. □

In general, a given market may have infinitely many risk-neutral measures. In fact, any vector belonging to $W^{\perp} \cap P$ in the proof above is risk-neutral.

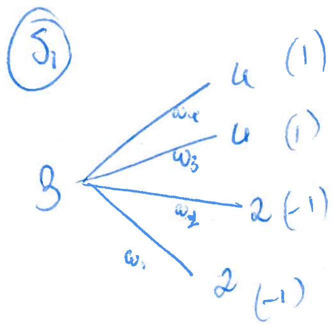
Example 25. *Consider the following market with $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, assets S_1 and S_2, and interest rate $r = 0$.* 2 assets/shares

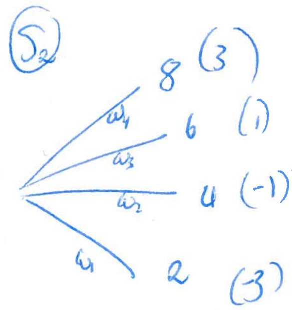
n	$S_n(0)$	$S_n(1,\omega_1)$	$S_n(1,\omega_2)$	$S_n(1,\omega_3)$	$S_n(1,\omega_4)$	
1	*3*	*2*	*2*	*4*	*4*	S_1
2	*5*	*2*	*4*	*6*	*8*	S_2

What are the risk-neutral measures for this model? We have the following values for ΔS_n.

$S(1)-S(0)$ $S(\omega_1)-S(0)$

n	$\Delta S_n(1,\omega_1)$	$\Delta S_n(1,\omega_2)$	$\Delta S_n(1,\omega_3)$	$\Delta S_n(1,\omega_4)$
1	*-1*	*-1*	*1*	*1*
2	*-3*	*-1*	*1*	*3*





So we must satisfy the equations

$$-q_1 - q_2 + q_3 + q_4 = 0$$
$$-3q_1 - q_2 + q_3 + 3q_4 = 0.$$

since sum q prob = 1

Solutions to this are of form $(\lambda, \mu, \mu, \lambda)$. Since we must also have $q_1 + q_2 + q_3 + q_4 = 1$ and $q_i > 0$ for every i, we get the solution set

end up with 0 probability - not allowed for risk neutral measure

$q_1 = q_4 = \lambda$

$q_2 = q_3 = \mu$

$$\{(p/2, (1-p)/2, (1-p)/2, p/2) : 0 < p < 1\}.$$

So risk-neutral probability measures exist for this model, and we can conclude that there are no arbitrage opportunities.

*

Note that this model determines a unique risk-neutral value for an asset with claim (x_1, x_2, x_3, x_4) if and only if $x_1 + x_4 = x_2 + x_3$.

Example 26. *Consider the following model, and assume the interest rate is* $r = 1/9$.

n	$S_n(0)$	$S_n(1,\omega_1)^*$	$S_n(1,\omega_2)^*$
1	5	6	4
2	1	1	1

Clearly $\mathbb{Q} = (1/2, 1/2)$ is a risk-neutral measure. Since the two assets are linearly independent, and there are two states, the market is complete. It follows that $\mathbb{Q}$ is the unique risk-neutral probabiltiy measure.

Now consider an asset X with claim

$X(\omega_1)$	$X(\omega_2)$
7	2

How can we value X? We have

$$\mathbb{E}_{\mathbb{Q}}[X(1)^*] = \left(\frac{1}{2}\cdot 7 + \frac{1}{2}\cdot 2\right)/(1+r) = \frac{9}{2}\cdot\frac{9}{10} = \frac{81}{20} = 4.05.$$

Since $\mathbb{Q}$ is risk-neutral, we have $X(0)^* = 4.05$.

We could also value X by replication: looking at each ω_i in turn, we have

$$6H_1 + H_2 = 7/(1+r)$$
$$4H_1 + H_2 = 2/(1+r),$$

* we get $2q_1 + 2q_4 = 1$ $\quad p = 2q_1$

$q = \frac{1-p}{2}$

and solving this gives $H_1 = 2.25$, $H_2 = -7.2$. *We can construct a portfolio replicating* X *by borrowing* 7.2 *(i.e. selling the riskless asset) and buying* 2.25 *units of asset 1, which cost* $2.25 \cdot 5 = 11.25$. *The total cost to us is* $11.25 - 7.2 = 4.05$.

At time 1, in state ω_1, *our* 2.25 *units of asset 1 are worth* $2.25 \cdot 6 \cdot (10/9) = 15$, *while our bank debt has become* $7.2 \cdot (10/9) = 8$. *The portfolio therefore has value* $15 - 8 = 7$ *(which discounts to present value of* $7 \cdot (9/10) = 6.3$*). A similar calculation can be performed for* ω_2.

Don't need to bother replicating - just use synthetic probabilities

28/10/14

10 The binomial model

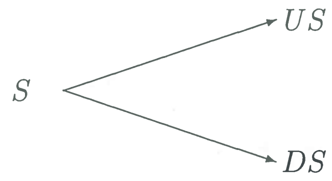
10.1 Single period

An important model for changes in the price of an asset is given by a binomial scaling process. If the current price is S, then after one time period the price will be either US or DS, where we shall always assume

$$D < U.$$

So the change in price at each step is proportional to the current price. (Is this a reasonable assumption?)

We have the following situation.

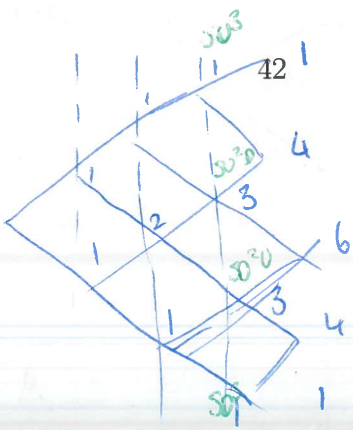


The risk-neutral probabilities are given by

$$q_U = \frac{S - (SD)^*}{(SU)^* - (SD)^*} = \frac{1 - D^*}{U^* - D^*}$$
$$q_D = \frac{(SU)^* - S}{(SU)^* - (SD)^*} = \frac{U^* - 1}{U^* - D^*}.$$

Note that the probabilities depend only on the factors U and D: they do not depend on the current price.

Suppose that the interest rate is $r > 0$ per time unit, compounded continuously. We will work with continuously compounded interest as we shall be interested taking ever shorter time intervals, and this fits better with continuous compounding. [Interest compounded at fixed intervals is the same as continuously compounded interest (at a different rate), if we are only interested in the price at the end of each time interval.]



multiple time steps

numbers indicate no. of ways to get there from before $q_H = \frac{S - S_L^*}{S_H^* - S_L^*} = \frac{S - SD^*}{SH^* - SD^*} = \frac{1-D^*}{H^*-D^*}$

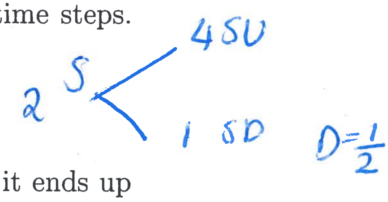
in below example

Then $U = e^r U^*$ and $D = e^r D^*$, so

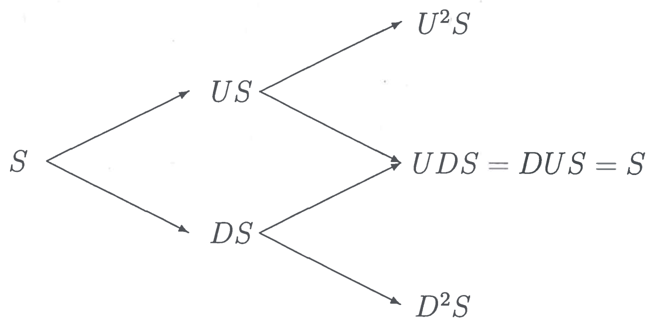
$$q_U = \frac{e^r - D}{U - D}$$
$$q_D = \frac{U - e^r}{U - D}.$$

We will be considering assets that evolve over a number of time steps. Let us assume for the moment that

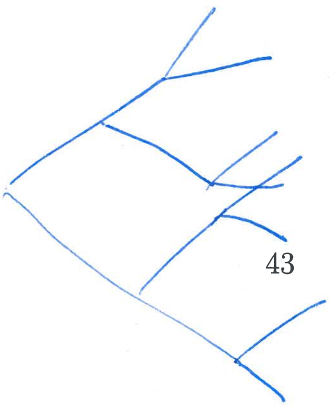
$$\boxed{UD = 1,}$$



so that if the stock goes up in one period and down in the next, it ends up at its original price. For instance, if $DU = 1$, we have the following situation after two steps.



How should we choose U and D so as to get a reasonable model of asset price?



This is a non-recombining tree!

* (probably not examinable)
but don't trust him

10.2 Choosing U and D

Let's consider a small time-step of length Δt. If the interest rate is r, then discounting over time Δt involves a factor $e^{r\Delta t}$, so

$$\mathbb{E}_{\mathbb{Q}}(S(\Delta t)) = e^{r\Delta t}S(0).$$

Since $\mathbb{E}_{\mathbb{Q}}(S(\Delta t)) = qUS + (1-q)DS$, where $q = q_U$, we have ← prob q ↑ + prob q ↓

$$qU + (1-q)D = e^{r\Delta t}. \tag{1}$$

Since we are also assuming

$$UD = 1,$$

we can write $U = e^{\lambda}$ and $D = e^{-\lambda}$, for some $\lambda > 0$. What should this λ be?

Here is a rough justification for one choice of λ. The variance of the $\pm\lambda$ term is proportional to λ^2. Now suppose that we want the price at time 1 to be of the form Se^X, where X is a random variable with variance σ^2. Since we are taking steps of length Δt, we have taken $1/\Delta t$ steps; if we assume that the steps are independent, then X will be the sum of $1/\Delta t$ terms with variance λ^2, and so X will have variance $\lambda^2/\Delta t$. It follows that $\lambda = \sigma\sqrt{\Delta t}$, and so

$$U = e^{\sigma\sqrt{\Delta t}}, \qquad D = e^{-\sigma\sqrt{\Delta t}}.$$

Substituting these values into (1), we get

$$qe^{\sigma\sqrt{\Delta t}} + (1-q)e^{-\sigma\sqrt{\Delta t}} = e^{r\Delta t},$$

which implies

$$q = \frac{e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}}.$$

Since

$$e^x \approx 1 + x$$

if you include higher order terms - all cancel out.

for small x, we get

$$q \approx \frac{r\Delta t + \sigma\sqrt{\Delta t}}{2\sigma\sqrt{\Delta t}} = \frac{1}{2} + \frac{r}{2\sigma}\sqrt{\Delta t},$$

which tends to $1/2$ as $\Delta t \to 0$.

Taking these approximations needs a little justification (are we sure that the errors in our approximation are smaller than the values we end up with?),

don't include higher order terms - always end up with risk neut prob $= \frac{1}{2}$
This is WRONG

$\mathbb{E}_Q(S(\Delta T)) = e^{r\Delta T}S$

so let's do that last calculation more rigorously. We fix r and σ, and consider what happens as $\Delta t \to 0$. For $|x| < 1$, we have the approximation

$$\exp(x) = 1 + x + x^2/2 + O(x^3).$$

If Δt is small, then

$$e^{r\Delta t} = 1 + r\Delta t + O\left((\Delta t)^2\right)$$

and

$$e^{-\sigma\sqrt{\Delta t}} = 1 - \sigma\sqrt{\Delta t} + \sigma^2\Delta t/2 + O\left((\Delta t)^{3/2}\right),$$

so

$$e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}} = \sigma\sqrt{\Delta t} + (r - \sigma^2/2)\Delta t + O\left((\Delta t)^{3/2}\right).$$

(Note that, as Δt is small, $(\Delta t)^2$ is much smaller than $(\Delta t)^{3/2}$.) We also have

$$e^{\sigma\sqrt{\Delta t}} = 1 + \sigma\sqrt{\Delta t} + \frac{\sigma^2}{2}\Delta t + O\left((\Delta t)^{3/2}\right)$$

and

$$e^{-\sigma\sqrt{\Delta t}} = 1 - \sigma\sqrt{\Delta t} + \frac{\sigma^2}{2}\Delta t + O\left((\Delta t)^{3/2}\right)$$

and so

$$e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}} = 2\sigma\sqrt{\Delta t} + O\left((\Delta t)^{3/2}\right).$$

So

$$\begin{aligned}
q &= \frac{(r - \sigma^2/2)\Delta t + \sigma\sqrt{\Delta t} + O\left((\Delta t)^{3/2}\right)}{2\sigma\sqrt{\Delta t} + O((\Delta t)^{3/2})} \\
&= \frac{(r - \sigma^2/2)\sqrt{\Delta t} + \sigma + O\left((\Delta t)\right)}{2\sigma + O\left(\Delta t\right)} \\
&= (1 + O(\Delta t))\,\frac{(r - \sigma^2/2)\sqrt{\Delta t} + \sigma}{2\sigma} \\
&= \frac{1}{2} + \frac{1}{2\sigma}(r - \frac{\sigma^2}{2})\sqrt{\Delta t} + O(\Delta t).
\end{aligned}$$

It turns out that we were right to be more careful, as our value for q is slightly different! This is a subtle point - the term is the same that you get in stochastic calculus

Let's check that the variance estimate is OK with this q: after one step, we have a price

$$Se^{\pm\sigma\sqrt{\Delta t}}.$$

The exponent $X_0 = \pm\sigma\sqrt{\Delta t}$ has mean

$$q(\sigma\sqrt{\Delta t}) + (1-q)(-\sigma\sqrt{\Delta t}) = (2q-1)\sigma\sqrt{\Delta t} = O(\Delta t).$$

So the variance is

$$\mathrm{var}[X_0] = \mathbb{E}[X_0^2] - \mathbb{E}[X_0]^2 = \sigma^2\Delta t + O\left((\Delta t)^2\right),$$

which means that at time 1 the exponent X in Se^X has variance

$$(1/\Delta t)\cdot \mathrm{var}(X_0) = \sigma^2 + O(\Delta t)$$

which tends to σ^2 as $\Delta t \to 0$.

Why do we get a σ^2 in the formula for q? Calculating the mean of X_0 more carefully gives

$$\mathbb{E}X_0 = (2q-1)\sigma\sqrt{\Delta t} = (r - \sigma^2/2)\Delta t + O\left((\Delta t)^{3/2}\right),$$

and so we have a drift of $(r-\sigma^2/2)\Delta t$ per time step. At time 1, we have

$$\mathbb{E}X = r - \sigma^2/2 + O(\sqrt{\Delta t})$$

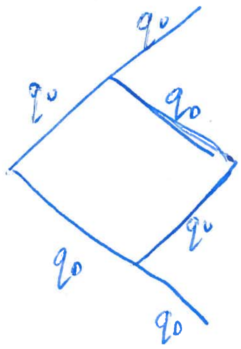
and

$$\mathrm{var}X = \sigma^2 + O(\Delta t).$$

If we choose huge value per U top branch blows up

Make sure to choose a viable model.

Making sure variance grows at right rate



ω_1 $N(\omega_1)=2$ $(q_U)^N = q_U^2$

ω_2 $N(\omega_2) = 1$ $q_U q_D$
ω_3 $N(\omega_3) = 1$

ω_3 $N(\omega_4) = 0$ q_D^2

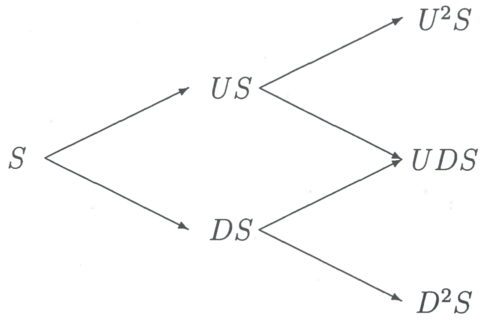
10.3 Multiperiod binomial model

Consider a T-period model. An asset with initial price S_0 will have subsequent prices $S_1, \ldots, S_T$, where each S_i is a random variable. We can represent the possible sequences of asset prices by binary sequences

$$\omega = (\omega_1, \ldots, \omega_T),$$

where a digit 0 corresponds to D and a digit 1 corresponds to U.

For instance, if $T = 2$ there are four paths that the asset price can follow.



Let us write

$$N(\omega) = \sum_{i=1}^{T} \omega_i$$

for the number of increases in price. Since there are $N(\omega)$ increases and $T - N(\omega)$ decreases, the probability of the path ω is

$$\mathbb{Q}(\omega) = q_U^{N(\omega)} q_D^{T-N(\omega)}.$$

$\mathbb{Q}$ is the risk-neutral measure for the T-period process (more about this later). Since there are $\binom{T}{n}$ paths with exactly n increases, we have

$$\mathbb{Q}(S_T = S_0 U^n D^{T-n}) = \binom{T}{n} q_U^n q_D^{T-n}.$$

Now, for $0 \le t \le T$, let us write

$$N_t(\omega) = \sum_{i=1}^{t} \omega_i,$$

so $N_t(\omega)$ is the number of increases in the first t time periods. We get

$$S_t(\omega) = S_0 U^{N_t(\omega)} D^{t-N_t(\omega)},$$

and, if $UD = 1$, this becomes

$$\begin{aligned} S_t(\omega) &= S_0 U^{N_t(\omega)} U^{N_t(\omega)-t} \\ &= S_0 U^{2N_t(\omega)-t}. \end{aligned}$$

10.4 Cox-Ross-Rubinstein valuation

Our aim is to value an option $(S_T - K)^+$ in a multiperiod binomial model with T periods of length Δt. We shall let $T \to \infty$ and $\Delta t \to 0$ in such a way that $T\Delta t$ remains constant and vary U and D so that the variance of $\log S_T$ converges to σ^2.

Let us note first that, for any state ω,

$$S(T,\omega) = S(0)U^{N_T(\omega)}D^{T-N_T(\omega)}.$$

It follows that

$$S(T,\omega) \geq K$$

if and only if

$$U^{N_T(\omega)}D^{T-N_T(\omega)} \geq K/S(0),$$

which is equivalent to

$$N_T(\omega)\log U + (T - N_T(\omega))\log D \geq \log(K/S(0)),$$

or in other words

$$N_T(\omega) \geq \frac{\log(K/S(0)) - T\log D}{\log(U/D)}.$$

Let

$$\hat{n} = \frac{\log(K/S(0)) - T\log D}{\log(U/D)}.$$

(Note that $\hat{n}$ need not be an integer.)

The fair price at time 0 of a call with strike price K and expiry time $T\Delta t$ is

$$V_{\text{call}} = e^{-rT\Delta t}\mathbb{E}\left[(S_T - K)^+\right],$$

where r is the (continuously compounded) interest rate. Using our model for S_T, we get

$$\begin{aligned} V_{\text{call}} &= e^{-rT\Delta t}\sum_{n\geq\hat{n}}(S_0U^nD^{T-n} - K)\cdot\mathbb{Q}(N_T = n) \\ &= e^{-rT\Delta t}\sum_{n\geq\hat{n}}(S_0U^nD^{T-n} - K)\binom{T}{n}q_U^n q_D^{T-n}. \end{aligned}$$

Now if (q_U, q_D) is risk-neutral, we have

$$q_U U + q_D D = e^{r\Delta t},$$

and so defining

$$\hat{q}_U = e^{-r\Delta t} q_U U$$

and

$$\hat{q}_D = e^{-r\Delta t} q_D D$$

gives a new probability measure $(\hat{q}_U, \hat{q}_D)$. We can then rewrite our call valuation as

$$\begin{aligned} V_{\text{call}} &= S_0 e^{-rT\Delta t} \sum_{n \geq \hat{n}} (q_U U)^n (q_D D)^{T-n} \binom{T}{n} - K e^{-rT\Delta t} \sum_{n \geq \hat{n}} \binom{T}{n} q_U^n q_D^{T-n} \\ &= S_0 \sum_{n \geq \hat{n}} \binom{T}{n} \hat{q}_U^n \hat{q}_D^{T-n} - K^* \sum_{n \geq \hat{n}} \binom{T}{n} q_U^n q_D^{T-n}. \end{aligned}$$

GIRSANOV

So

probability

$$\boxed{V_{\text{call}} = S_0 \widehat{\mathbb{Q}}[N_T \geq \hat{n}] - K^* \mathbb{Q}[N_T \geq \hat{n}],}$$

where we write $\mathbb{E}_{\mathbb{Q}}$ for expectation with respect to the measure on N_T coming from the one-step probabilities (q_U, q_D), and $\mathbb{E}_{\widehat{Q}}$ for the measure coming from the one-step probabilities $(\hat{q}_U, \hat{q}_D)$.

We shall determine the limiting value of this expression for V_{call} shortly, but first we will need to recall the Central Limit Theorem.

10.5 Independence and the Central Limit Theorem

If Ω is finite, we say that random variables $X_1, \ldots, X_n$ are *independent* if, for every $x_1, \ldots, x_n$,

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i).$$

If Ω is infinite, we may not be able to use this definition, so we work instead with the following.

Definition 1. *Random variables $X_1, \ldots, X_n$ are* independent *if, for every $x_1, \ldots, x_n$, we have*

$$\mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq x_i).$$

An infinite sequence of random variables $X_1, X_2, \ldots$ is said to be independent if every finite subset of the random variables is independent.

In other words, we work with the distribution functions. (Exercise: Check that the two definitions are equivalent when Ω is finite.)

Independent random variables are particularly nicely behaved. For instance, the following is true.

Theorem 2. *If $X_1, \ldots, X_n$ are independent then*

$$\mathbb{E}[X_1 \cdots X_n] = \prod_{i=1}^{n} \mathbb{E}[X_i].$$

When the random variables have the same distribution, even stronger statements hold.

Definition 3. *A sequence of random variables $X_1, X_2, \ldots$ is said to be* iid *(or* independent and identically distributed*) if the sequence is independent and all the X_i have the same distribution function $F(x)$.*

If $(X_i)_{i=1}^{\infty}$ is an iid sequence, where the random variables all have mean μ and variance σ^2, then we can consider the sum

$$S_n = \sum_{i=1}^{n} X_i.$$

We have

$$\mathbb{E}[S_n] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n\mu$$

and, by independence,

$$\operatorname{var}(S_n) = \sum_{i=1}^{n} \operatorname{var}(S_i) = n\sigma^2.$$

It follows that the random variable

$$Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

has mean 0 and variance 1.

Theorem 4 (Central Limit Theorem). *Let $(X_i)_{i=1}^{\infty}$ be an iid sequence of random variables with mean μ and variance σ^2, and define*

$$Y_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}}.$$

Then, for any real number y,

$$\mathbb{P}(Y_n \leq y) \to \Phi(y)$$

as $n \to \infty$.

Here we have written

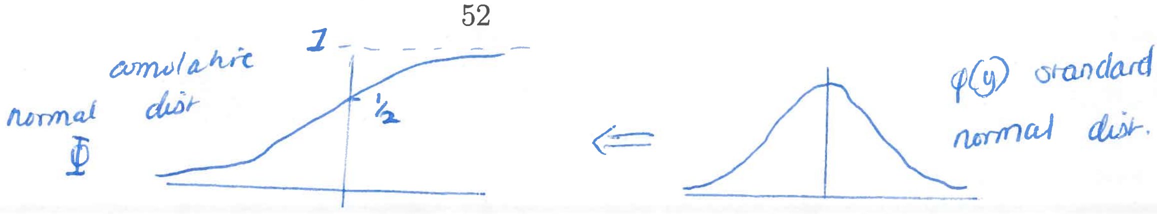
$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} dt$$

for the probability that a standard normal random variable is less than or equal to y.

The Central Limit Theorem is one of the most important theorems in probability theory, and it has many extensions and generalizations. (We won't prove the theorem here.)

A very simple corollary to the Central Limit Theorem is the special case of Bernoulli random veriables. A random variable is *Bernoulli* if it takes only values 0 and 1. If X is Bernoulli with

$$\mathbb{P}(X = 1) = p, \quad \mathbb{P}(X = 0) = 1 - p$$



then

$$\mathbb{E}[X] = p$$

and

$$\mathrm{var}(X) = \mathbb{E}[(X-p)^2] = p(1-p)^2 + (1-p)p^2 = p(1-p).$$

If we add up n Bernoulli random variables X_i with mean p, we get

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i = t\right) = \binom{n}{t} p^t (1-p)^{n-t}.$$

This is a *binomial distribution*, and has mean np and variance $np(1-p)$.

The following theorem follows directly from the Central Limit Theorem, but was known much earlier.

Theorem 5 (De Moivre–Laplace). *Let $X_1, X_2, \ldots$ be iid Bernoulli random variables with mean p. Then for any real number y,*

$$\mathbb{P}\left(\frac{\sum_{i=1}^{n} X_i - np}{\sqrt{np(1-p)}} \le y\right) \to \Phi(y) \qquad (2)$$

as $n \to \infty$.

It is sometimes useful to know how quickly the probability converges to $\Phi(y)$ in the De Moivre–Laplace Theorem. In fact, it can be shown that

$$\left|\mathbb{P}\left(\frac{\sum_{i=1}^{n} X_i - \mu}{\sigma} \le y\right) - \Phi(y)\right| \le \frac{1}{\sigma}, \qquad (3)$$

where we have written $\mu = np$ and $\sigma = \sqrt{np(1-p)}$.

σ standard dev.

10.6 Cox-Ross-Rubinstein continued

Recall that we had reached the valuation

$$V_{\text{call}} = S_0\widehat{\mathbb{Q}}[N_T \geq \hat{n}] - K^*\mathbb{Q}[N_T \geq \hat{n}], \tag{4}$$

where $\mathbb{Q}$ is the measure on the T-period binomial model corresponding to the one-step risk-neutral measure (q_U, q_D), and $\widehat{\mathbb{Q}}$ corresponds to the one-step measure defined by

$$\hat{q}_U = e^{-r\Delta t} q_U U, \quad \hat{q}_D = e^{-r\Delta t} q_D D,$$

and

$$\hat{n} = \frac{\log(K/S(0)) - T\log D}{\log(U/D)}.$$

Let us fix the expiry date at time $t = 1$, and subdivide the interval from time $t = 0$ to $t = 1$ into T equal steps of size $\Delta t = 1/T$, so that

$$T\Delta t = 1.$$

As before, we choose U and D to be

$$U = e^{\sigma\sqrt{\Delta t}}, \quad D = e^{-\sigma\sqrt{\Delta t}},$$

and, since the risk-neutral probability $q = q_U$ satisfies

$$qU + (1-q)D = e^{r\Delta t},$$

we get

$$q = \frac{e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}}.$$

Recall that

$$N_T(\omega) = \sum_{i=1}^{T} \omega_i,$$

so

$$\mathbb{E}_{\mathbb{Q}}[N_T] = T\mathbb{E}_Q[\omega_1] = Tq$$

and

$$\mathbb{E}_{\widehat{\mathbb{Q}}}[N_T] = T\mathbb{E}_{\widehat{Q}}[\omega_1] = T\hat{q},$$

where we have written $\hat{q} = \hat{q}_u$.

Similarly, we can calculate the variance. By independence, we get

$$\mathrm{var}_{\mathbb{Q}}[N_T] = T\mathrm{var}_{\mathbb{Q}}\omega_1 = Tq(1-q)$$

and

$$\mathrm{var}_{\widehat{\mathbb{Q}}}[N_t] = T\mathrm{var}_{\widehat{\mathbb{Q}}}[\omega_1] = T\hat{q}(1-\hat{q}).$$

We can now evaluate our expression for the price at time 0 of a call option with strike price K.

Theorem 6. *With the assumptions above, and writing*

$$d_{\pm} = \frac{\log(S(0)/K) + (r \pm \sigma^2/2)}{\sigma},$$

we have

$$\boxed{\lim_{\Delta t \to 0} V_{\text{call}} = S(0)\Phi(d_+) - K^*\Phi(d_-).}$$

BLACK SCHOLES FORMULA (probably in exam)

Proof. Let us start with the first term of (4).

$$\begin{aligned}
\mathbb{Q}(N_T \geq \hat{n}) &= \mathbb{Q}\left(N_T \geq \frac{\log(K/S(0)) - T\log D}{\log(U/D)}\right) \\
&= \mathbb{Q}\left(N_T \geq \frac{\log(K/S(0)) + T\sigma\sqrt{\Delta t}}{2\sigma\sqrt{\Delta t}}\right) \\
&= \mathbb{Q}\left(N_T - Tq \geq \frac{\log(K/S(0)) + T(1-2q)\sigma\sqrt{\Delta t}}{2\sigma\sqrt{\Delta t}}\right) \\
&= \mathbb{Q}\left(\frac{N_T - Tq}{\sqrt{Tq(1-q)}} \geq \frac{\log(K/S(0)) + T(1-2q)\sigma\sqrt{\Delta t}}{2\sigma\sqrt{T\Delta t\, q(1-q)}}\right)
\end{aligned}$$

Let

$$d_T = \frac{\log(K/S(0)) + T(1-2q)\sigma\sqrt{\Delta t}}{2\sigma\sqrt{T\Delta t\, q(1-q)}}.$$

V_{call} $\underbrace{S, r, \sigma, K, t}_{\text{the variables we need.}}$

Assuming stock prices normally distributed (which they aren't)

Since $T\Delta t = 1$, we get

$$\begin{aligned} d_T &= \frac{\log(K/S(0))}{2\sigma\sqrt{q(1-q)}} + \frac{(1-2q)\sigma\sqrt{T}}{2\sigma\sqrt{q(1-q)}} \\ &= \frac{\log(K/S(0))}{2\sigma\sqrt{q(1-q)}} + \frac{(1-2q)\sigma}{2\sigma\sqrt{\Delta t\, q(1-q)}} \end{aligned}$$

Now

$$\begin{aligned} q &= \frac{e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}} \\ &= \frac{(1 + r\Delta t) - (1 - \sigma\sqrt{\Delta t} + \sigma^2\Delta t/2) + O\big((\Delta t)^{3/2}\big)}{(1 + \sigma\sqrt{\Delta t} + \sigma^2\Delta t/2) - (1 - \sigma\sqrt{\Delta t} + \sigma^2\Delta t/2) + O\big((\Delta t)^{3/2}\big)} \\ &= \frac{r\Delta t + \sigma\sqrt{\Delta t} - \sigma^2\Delta t/2 + O\big((\Delta t)^{3/2}\big)}{2\sigma\sqrt{\Delta t} + O\big((\Delta t)^{3/2}\big)} \\ &= \frac{\sigma + (r - \sigma^2/2)\sqrt{\Delta t} + O(\Delta t)}{2\sigma + O(\Delta t)}, \end{aligned}$$

which tends to $1/2$ as $\Delta t \to 0$.

Also,

$$\begin{aligned} 1 - 2q &= 1 - 2\left(\frac{\sigma + (r - \sigma^2/2)\sqrt{\Delta t} + O(\Delta t)}{2\sigma + O(\Delta t)}\right) \\ &= \frac{-(2r - \sigma^2)\sqrt{\Delta t} + O(\Delta t)}{2\sigma + O(\Delta t)}. \end{aligned}$$

So

$$\begin{aligned} \frac{1-2q}{\sqrt{\Delta t}} &= \frac{-(2r - \sigma^2) + O(\sqrt{\Delta t})}{2\sigma + O(\Delta t)} \\ &\to -\frac{1}{\sigma}\left(r - \frac{\sigma^2}{2}\right) \end{aligned}$$

as $\Delta t \to 0$. So

$$d_T \to \frac{\log(K/S(0)) - (r - \sigma^2/2)}{\sigma} = -d_-$$

as $T \to \infty$ (and so $\Delta t \to 0$).

We have reached

$$\begin{aligned}\mathbb{Q}(N_T \geq \hat{n}) &= \mathbb{Q}\left(\frac{N_T - Nq}{\sqrt{Tq(1-q)}} \geq d_T\right) \\ &= 1 - \mathbb{Q}\left(\frac{N_T - Nq}{\sqrt{Tq(1-q)}} < d_T\right),\end{aligned}$$

where $d_T \to -d_-$ as $T \to \infty$. Since N_T is binomial with mean Tq and variance $\sqrt{Tq(1-q)}$, we would like to use De Moivre-Laplace or the Central Limit Theorem to conclude that

$$\mathbb{Q}\left(\frac{N_T - Nq}{\sqrt{Tq(1-q)}} < d_T\right) \to \Phi(-d_-).$$

The one problem is that q isn't *quite* constant. But this doesn't stop us: we just apply De Moivre-Laplace with the error bound (note that the error is $1/\sqrt{Tq(1-q)} \to 0$ as $T \to \infty$) to reach the same conclusion. Then since, for $y \in \mathbb{R}$,

$$\Phi(y) = \frac{1}{2\pi}\int_{-\infty}^{y} e^{-x^2/2}dx = \frac{1}{2\pi}\int_{-y}^{\infty} e^{-x^2/2}dx = 1 - \Phi(-y),$$

we get

$$\mathbb{Q}(N_T \geq \hat{n}) \to 1 - \Phi(-d_-) = \Phi(d_-)$$

as $T \to \infty$.

Arguing similarly for the other term, we get

$$\tilde{q} = qe^{\sigma\sqrt{\Delta t}}e^{-r\Delta t} \to \frac{1}{2}$$

as $\Delta t \to 0$, while

$$\widehat{\mathbb{Q}}(N_T \geq \hat{n}) = \widehat{\mathbb{Q}}\left(\frac{N_T - T\hat{q}}{\sqrt{T\hat{q}(1-\hat{q})}} \geq \hat{d}_T\right),$$

where

$$d_T = \frac{\log(K/S(0))}{2\sigma\sqrt{\hat{q}(1-\hat{q})}} + \frac{(1-2\hat{q})\sigma}{2\sigma\sqrt{\Delta t\,\hat{q}(1-\hat{q})}}$$

The first term behaves as before. For the second term, note that

$$\hat{q} = qe^{\sigma\sqrt{\Delta t}}e^{-r\Delta t} = \frac{e^{\sigma\sqrt{\Delta t}} - e^{-r\Delta t}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}}.$$

Expanding as before, we get

$$\hat{q} = \frac{\sigma + (r + \sigma^2/2)\sqrt{\Delta t} + O(\Delta t)}{2\sigma + O(\Delta t)},$$

and so

$$\frac{1 - 2\hat{q}}{\sqrt{\Delta t}} = \frac{-2r - \sigma^2 + O(\sqrt{\Delta t})}{2\sigma + O(\Delta t)} \to -\frac{1}{\sigma}\left(r + \frac{\sigma^2}{2}\right).$$

So

$$\hat{d}_T \to \frac{\log(K/S(0)) - (r + \sigma^2/2)}{\sigma} = -d_+$$

as $T \to \infty$. The argument is completed as before: we get

$$\widehat{\mathbb{Q}}(N_T \geq \hat{n}) \to 1 - \Phi(-d_+) = \Phi(d_+)$$

as $T \to \infty$. $\square$

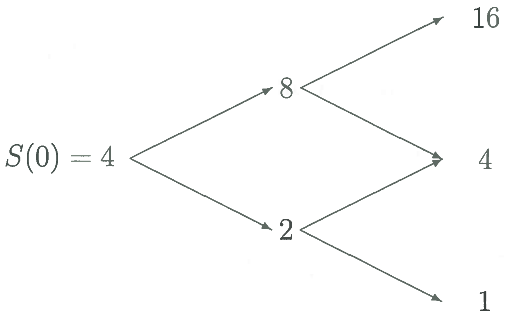
v messy
not v interesting $\Bigg\}$ never been examined

30/10/14 (1)

11 Two-period replication (and beyond)

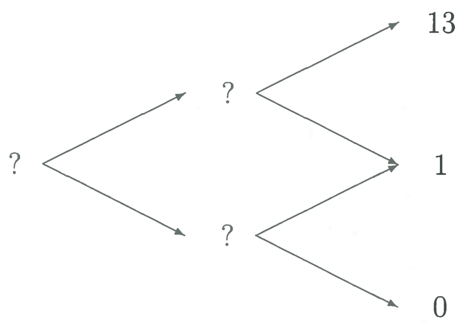
11.1 Two-period example

Consider a two-period binomial model with $U = 2$, $D = 1/2$, $r = 0$ and $S(0) = 4$:



Let us consider an option with strike price 3 that expires at time 2, so the payoff is $X = (S(2) - 3)^+$. How can we price the option at time 0? If we assume that *at each step, the market has no arbitrage opportunity*, we can use replication to work backwards through the tree.

At the moment, we have the following information about X:



We start at time 1, and examine each of the two possible cases. If $S(1) = 8$, then we have the following one-period model:

$$\begin{tikzcd}[column sep=large] & S(2) = 16 \qquad X(2) = 13 \\ S(1) = 8 \arrow[ur] \arrow[dr] & \\ & S(2) = 4 \qquad X(2) = 1 \end{tikzcd}$$

Consider a portfolio (x, y), consisting of x units of stock and y units of the riskless asset. This replicates the option if

$$\begin{aligned} 16x + y &= 13 \\ 4x + y &= 1, \end{aligned}$$

and so $x = 1$, $y = -3$. The replicating portfolio has cost at time 1 equal to $8x + y = 5$.

On the other hand, if $S(1) = 2$, then we have a different one-period model:

$$\begin{tikzcd}[column sep=large] & S(2) = 4 \qquad X(2) = 1 \\ S(1) = 2 \arrow[ur] \arrow[dr] & \\ & S(2) = 1 \qquad X(2) = 0 \end{tikzcd}$$

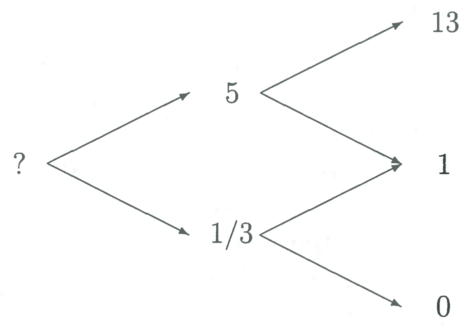
A replicating portfolio (x, y) has

$$\begin{aligned} 4x + y &= 1 \\ x + y &= 0, \end{aligned}$$

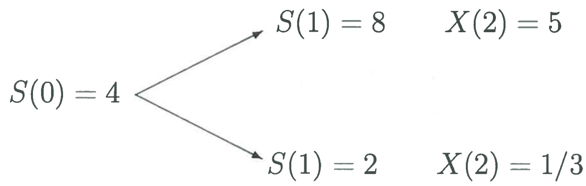
and so $x = 1/3$, $y = -1/3$. The portfolio has cost at time 1 equal to $2x + y = 1/3$.

Thus the cost at time 1 of a portfolio that replicates the claim over the period from $t = 1$ to $t = 2$ is 5 if $S(1) = 8$ and $1/3$ if $S(1) = 2$. Since we are assuming the absence of arbitrage in each period, these must be the prices for the option at time $t = 1$.

We have the following information about X:



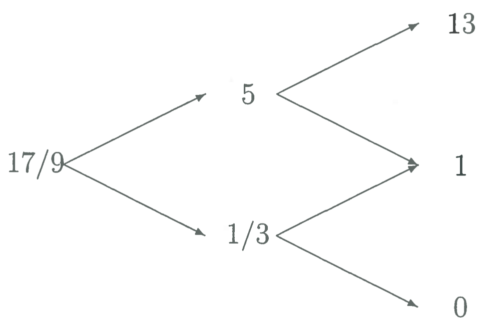
But we can now examine the period from $t = 0$ to $t = 1$. Regarding the price of X at time 1 as a contingent claim, we have the following one-period model:



A replicating portfolio (x, y) for X over this period has

$$\begin{aligned} 8x + y &= 5 \\ 2x + y &= 1/3, \end{aligned}$$

and so $x = 7/9$, $y = -11/9$. The portfolio has cost at time 0 equal to $4x + y = 4(7/9) - (11/9) = 17/9$. So we have the following prices for X:



11.1.1 Self-financing portfolios

How does our two-period example work in terms of portfolios? An initial fund of 17/9 is used to create a portfolio $(x_1, y_1) = (7/9, -11/9)$. At time 1, the value of the portfolio is equal to the the cost of a portfolio (x_2, y_2) replicating X over the second period. The required portfolio (x_2, y_2) depends on whether $S(1) = 8$ or $S(1) = 2$, but so does the value at time 1 of (x_1, y_1): in either case the value of (x_1, y_1) is equal to the cost of (x_2, y_2).

We can describe this dynamical replication with a time-dependent notation H_t, where in the example above we have:

$$H_1 = (7/9, -11/9)$$

and

$$H_2 = \begin{cases} (1, -3) & \text{if } S(1) = 8 \\ (1/3, -1/3) & \text{if } S(1) = 2. \end{cases}$$

Here H_t describes a portfolio constructed at time $t - 1$ for the purpose of replicating a claim at time t. We have

$$\begin{aligned} V_0(H_1) &= 17/9 \\ V_1(H_1) &= V_1(H_2) \\ V_2(H_2) &= X. \end{aligned}$$

The fact that $V_1(H_1) = V_1(H_2)$ says the the portfolio is *self-financing*: in other words, no money is added or removed after the initial investment.

11.1.2 Risk-neutrality

At each stage in the example above we have a one-period risk-neutral probability. For instance at time $t = 1$, we have a risk-neutral probability $\mathbb{Q}_{S(1)}$, which is one of the two possible risk-neutral probabilities depending on whether $S(1) = 8$ or $S(1) = 2$.

If $S(1) = 8$, we have

$$\mathbb{Q}_{S(1)=8} = \left(\frac{8-4}{16-4}, \frac{16-8}{16-4}\right) = \left(\frac{1}{3}, \frac{2}{3}\right).$$

If $S(1) = 2$, we have

$$\mathbb{Q}_{S(1)=2} = \left(\frac{2-1}{4-1}, \frac{4-2}{4-1}\right) = \left(\frac{1}{3}, \frac{2}{3}\right).$$

In either case, we have

$$X(1) = \mathbb{E}_{\mathbb{Q}_{S(1)}}[X(2)^*].$$

Similarly, at time 0, we have a risk-neutral probability

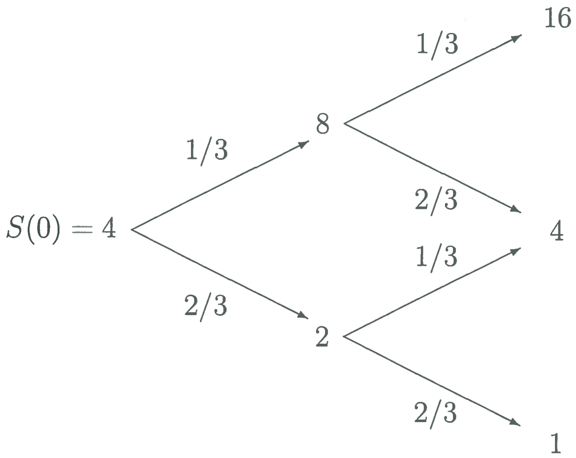
$$\mathbb{Q}_{S(0)} = \left(\frac{4-2}{8-2}, \frac{8-4}{8-2}\right) = \left(\frac{1}{3}, \frac{2}{3}\right),$$

and

$$X(0) = \mathbb{E}_{\mathbb{Q}_{S(1)}}[X(1)^*].$$

Note that the risk-neutral probabilities are the same at each step. But we should have expected this, as they depend only on the factors U and D.

Thus we have the following picture:



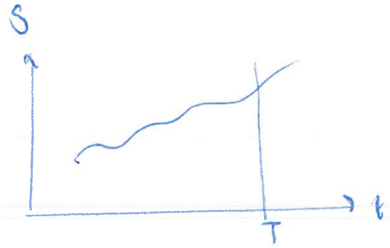
So we can determine the probability of each path $\omega = (\omega_1, \omega_2)$, where $\omega_t \in \{0, 1\}$.

Note that $\mathbb{Q}_{S(1)}$ is the conditional probability under $\mathbb{Q}$ of the prices $S(2)$. For example

$$\mathbb{Q}_{S(1)=8}(S(2) = 16) = \mathbb{Q}(S(2) = 16 | S(1) = 8)$$
$$\mathbb{Q}_{S(1)=4}(S(2) = 1) = \mathbb{Q}(S(2) = 1 | S(1) = 4).$$

We also have

$$X(0) = \mathbb{E}_{\mathbb{Q}_{S(0)}}[X(1)] = \mathbb{E}_{\mathbb{Q}_{S(0)}}\mathbb{E}_{Q_{S(1)}}[X(2)].$$



As asset moves more in money need to hold more shares

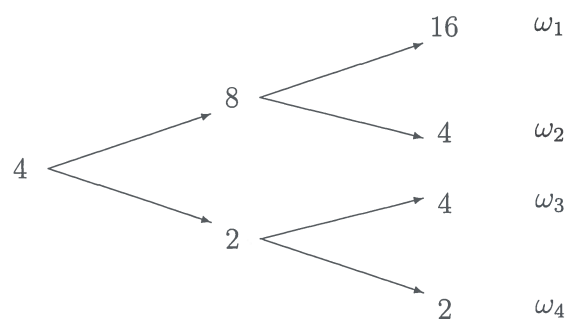
Monotonic relationship

11.1.3 Sample paths

The simplest description of the stock price over time involves writing down the *possible histories*, which we label $\omega_1, \ldots, \omega_4$:

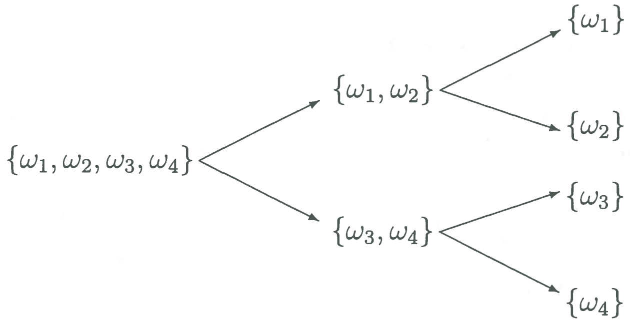
	$S(0)$	S(1)	S(2)
ω_1	4	8	16
ω_2	4	8	4
ω_3	4	2	4
ω_4	4	2	1

These histories describe four possible *states of nature.* We can also think of them as four *sample paths* through the binary tree:



How should we label the nodes in the tree? The prices do not give good labels for the nodes, as several paths may end up with the same price. For instance, ω_2 and ω_3 both end up with price 4. A better approach is to label each node with the set of paths that passes through it. For instance the node with price 8 corresponds to the label $\{\omega_1, \omega_2\}$.

We have the following labels in our example:



Note that, for each t, the nodes at time t give a partition of $\Omega = \{\omega_1, \ldots, \omega_4\}$.

Definition 1. *Let P and P' be partitions of Ω. We say that P' is a* refinement *of P if every set in P' is contained in some set of P. We write $P \leq P'$.*

For instance, in the case of our binomial tree, let us write P_i for the partition at time i. We have

$$\begin{aligned} P_0 &= \big\{\{\omega_1, \omega_2, \omega_3, \omega_4\}\big\} \\ P_1 &= \big\{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\big\} \\ P_2 &= \big\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\big\} \end{aligned}$$

and clearly

$$P_0 \leq P_1 \leq P_2.$$

What's more, if we pick any set $A \in P_i$, then it is partitioned by the sets

$$\{B : B \in P_{i+1}, B \subset A\}.$$

This is obviously true for any tree: the label we give to a node consists of the paths going through it, and the paths going through a given node each go through exactly one of its children. The label A corresponds to a node on the ith layer of the tree, while $\{B : B \in P_{i+1}, B \subset A\}$ correspond to its children.

The partitions P_i will be important. We want to analyze a multiperiod model in terms of its constituent one-period models. The different sample paths correspond to different sequences of prices that are revealed at each time. If we wait until the end of the final time period, we know all the prices and so we know which state of nature (=sample path) we are in. But if we are at an earlier time, we only know some of the initial prices, and so we may not know which state we are in.

For instance, at time 0, we know only $S(0)$, and so we could be in any state. At time 1, we know $S(0)$ and $S(1)$, and so we know whether the asset has gone up or down. If the asset has gone up, we know that we are in $\{\omega_1, \omega_2\}$, although we do not know whether we are in state ω_1 or ω_2; if the asset has gone down, we know that we are in $\{\omega_3, \omega_4\}$, although again we do not know which of the two possibilities we are in. Another way of saying this is that we know the first two nodes on our sample path (i.e. we know $S(0)$ and $S(1)$) but do not know the third. So we have reached some node on the tree: we know we are on one of the sample paths running through that node, but we do not know which. In other words, our current state of knowledge is represented by the label corresponding to the node we have reached.

Finally, note that $S(i)$ is constant on each of the sets in P_i. For instance $S(1)$ has the constant value 8 pm $\{\omega_1, \omega_2\}$, which is in P_1. Again, this is obvious: the paths passing through a node on the ith level of the tree correspond to paths with a particular sequence of prices up to time i. Clearly, these all have the same price at time i.

Filtration "⊆" tree

Partition "=" time step

11.2 Filtrations

Let us put this on a slightly more formal footing.

Definition 2. *(Simple version) A* filtration *on a set* $\Omega = \{\omega_1, \dots, \omega_K\}$ *is a sequence* $(P_t)_{t=0}^T$ *of partitions of* Ω *such that*

$$P_0 = \{\Omega\},$$
$$P_T = \{\{\omega_1\}, \dots, \{\omega_K\}\}$$

and

$$P_1 \leq \dots \leq P_T.$$

A filtration can easily be written in the form of a tree. Indeed, suppose that $(P_t)_{t=0}^T$ is a filtration. We say that $B_{t+1} \in P_{t+1}$ *succeeds* $A_t \in P_t$ if $B_{t+1} \subset A_t$. Note that each set $A_t \in P_t$ is partitioned by its successors in P_{t+1}.

The tree associated to the partition has levels $0, \dots, T$. The nodes on level i correspond to the sets in P_i, and each set on level i is joined to it successors on level $i+1$. For instance, if

$$P_0 = \{\{\omega_1, \omega_2, \omega_3, \omega_4\}\}$$
$$P_1 = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\}$$
$$P_2 = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\},$$

we get the tree in the previous section.

We shall be working with a collection of assets $\mathbf{S} = (S_1, \dots, S_N)$ and times $0, \dots, T$. We write $S_n(t, \omega)$ for the price of S_n in state ω at time t.

asset price / share price

Definition 3. *We say that the price process* $\mathbf{S}$ *is* adapted to the filtration $(P_t)_{t=0}^T$ *if, for every* n*, every* t *and every* $A_t \in P_t$*,* $S_n(t, \omega)$ *is constant on* A_t*.*

If $\mathbf{S}$ is adapted to the filtration, then we will write $S_n(t, A)$ for this constant value. For instance, in the previous section we had $S(0, \Omega) = 4$, $S(1, \{\omega_1, \omega_2\}) = 8$, $S(2, \{\omega_1\}) = 16$, and so on.

adapted to a filtration "=" can be put on a tree

If ω_1 & ω_2 don't both cause 8 then cannot have a tree

11.3 Measurability and trading strategies

Our definition of an adapted price process is a little clumsy. Let's pause a moment and clean it up.

Consider an asset S_i: this evolves through time as $S_i(0), \ldots, S_i(T)$. At a given time t we have a partition P_t. Each set in the partition represents a collection of sample paths that are indistinguishable at that time, so S_i must be constant on that set. We can say this more cleanly.

Definition 4. *Let P be a partition of Ω. We say that a random variable X is* P-measurable *if X is constant on each set in P.*

We can now define an adapted process more succinctly.

Definition 5. *Let $\mathcal{P} = (P_t)_{t=0}^T$ be a filtration of Ω, and let $S_i(0), \ldots, S_i(T)$ be a sequence of random variables. We say that the asset S_i is* adapted to *$\mathcal{P}$ if $S_i(t)$ is P_t-measurable for each t. The price process $\mathbf{S} = (S_1, \ldots, S_T)$ is* adapted to *$\mathcal{P}$ if every asset S_i is adapted to $\mathcal{P}$.*

Our aim is to prove a generalization of the no-arbitrage theorem to a multiperiod model. We will want to construct portfolios that vary over time.

Let us write $H_n(i)$ for the volume of asset n bought at time $i-1$ and held until time i. We will allow $H_n(i)$ to depend on the state ω. At time $i-1$ we construct the portfolio

$$\big(H_1(i,\omega), \ldots, H_N(i,\omega)\big).$$

The quantity of each asset that we hold at any given time is a random variable: it depends on ω.

Definition 6. *A* trading strategy $\mathbf{H}$ *is a sequence of portfolios $\mathbf{H}(1), \ldots, \mathbf{H}(T)$ such that*

1. *$\mathbf{H}(i) = (H_1(i), \ldots, H_N(i))$ is a portfolio that we hold from time $i-1$ to time i; and*

2. *$H_n(i)$ is a random variable for each i and n*

We are interested in tracking the evolution of a portfolio over time, so we will want to know the value of the portfolio at each step, and the amount we gain between steps.

Definition 7. *Let* $\mathbf{H}$ *be a trading strategy. The* value *of* $\mathbf{H}$ *at time* t *is the vector*

$$V_t(\mathbf{H}) = \sum_{n=1}^{N} H_n(t) S_n(t).$$

Written out in terms of sample paths, this says

$$V_t(\mathbf{H}, \omega) = \sum_{n=1}^{N} H_n(t, \omega) S_n(t, \omega).$$

In other words, when we reach time t we are holding the portfolio $\mathbf{H}(t)$, and $V_t(\mathbf{H})$ is the value of this portfolio. The profit we make at each step is given by the gain vector.

Definition 8. *The* gain vector *at time* t *is given by*

$$\begin{aligned} G_t(\mathbf{H}, \omega) &= \sum_{n=1}^{N} H_n(t, \omega)(S_n(t, \omega) - S_n(t-1, \omega)) \\ &= \sum_{n=1}^{T} H_n(t, \omega) \Delta S_n(t, \omega), \end{aligned}$$

where

$$\Delta S_n(t, \omega) = S_n(t, \omega) - S_n(t-1, \omega)$$

is the change in price of S_n *between time* $t-1$ *and time* t.

But what information are we allowed to use in choosing our portfolio? Since $\mathbf{H}_i$ is constructed at time $i-1$, it seems reasonable that we should be allowed to take the prices $S_n(0), \ldots, S_n(i-1)$, for each asset, into account: after all, we know these before we have to construct the portfolio. On the other hand, it doesn't seem reasonable that we should be allowed to look into the future. (Otherwise it would be far too easy to make money!)

Definition 9. *Let* $\mathcal{P} = (P_t)_{t=0}^{T}$ *be a filtration of* Ω, *and let* $H_n(0), \ldots, H_n(T)$ *be a sequence of random variables. We say that* H_n *is a* previsible process *if* $H_n(i)$ *is* P_{i-1}*-measurable for every* i. *The trading strategy* $\mathbf{H} = (H_1, \ldots, H_N)$ *is* previsible *if* H_n *is a previsible process for every* n.

*

only one timestep

this is important

i.e. may be in exam

We also want our trading strategies to be self-financing: we don't want to add or remove money.

Definition 10. *A trading strategy* **H** *is* self-financing *if*

$$\sum_{n=1}^{N} H_n(t,\omega)S_n(t,\omega) = \sum_{i=1}^{N} H_n(t+1,\omega)S_n(t,\omega)$$

for $t \geq 1$ and every ω. (In other words, in every possible state of nature, the construction cost at time t of the portfolio we will carry to time $t+1$ is equal to the value at time t of the portfolio constructed at time $t-1$.)

value at time t = value at time $t+1$

11.4 Filtrations (industrial strength) non-examinable

We won't need these definitions now (the definitions we have work perfectly well when Ω is finite), but for later reference let's see the full version of partitions, filtrations, etc.

First of all, we must replace partitions by something more general.

Definition 11. *A (nonempty) collection $\mathcal{F}$ of subsets of Ω is called a* σ-algebra *if*

1. *If $A \in \mathcal{F}$ then $\Omega \setminus A \in \mathcal{F}$*
2. *If $A_1, A_2, \ldots \in \mathcal{F}$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.*

We state a couple of facts (which we leave as exercises):

1. The family $\{\emptyset, \Omega\}$ is a σ-algebra
2. The collection of all subsets of Ω is a σ-algebra
3. If P is a partition of Ω then the collection of sets obtained by taking all possible unions (including the empty union) of sets in P is a σ-algebra (called the σ-algebra *generated* by P)

We must also replace measurability by a more general definition.

Definition 12. *Let $X : \Omega \to \mathbb{R}$ and suppose that $\mathcal{F}$ is a σ-algebra on Ω. We say that X is* $\mathcal{F}$-measurable *if, for every interval $(a, b) \subset \mathbb{R}$, we have*

$$\{\omega \in \Omega : X(\omega) \in (a, b)\} \in \mathcal{F}.$$

Equivalently, we have $X^{-1}(I) \in \mathcal{F}$ for every open interval $I \subset \mathbb{R}$.

This extends the notion of measurability we already have (exercise!). We can now define filtrations and adapted sequences.

Definition 13. *A sequence $\mathcal{F}_0 \ldots, \mathcal{F}_T$ of σ-algebras is a* filtration *if $\mathcal{F}_i \subset \mathcal{F}_j$ whenever $i < j$.*

Definition 14. *A sequence of random variables $X_0, \ldots, X_T$ is* adapted *to the filtration $\mathcal{F}_0, \ldots, \mathcal{F}_T$ if X_i is $\mathcal{F}_i$-measurable for each i.*

30/10/14 (2)

12 Conditional expectation and martingales

12.1 Introduction

Consider a filtration $(P_t)_{t=1}^T$ and a probability measure $\mathbb{P}$ on Ω. If $A \in P_t$, the probability of being in A is

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega),$$

Since P_t is a partition, we have

$$\sum_{A' \in P_t} \mathbb{P}(A') = 1.$$

Now consider the sets in P_{t+1} that partition A, say these are $B_1, \ldots, B_k$. We have

$$\mathbb{P}(A) = \sum_{i=1}^{k} \mathbb{P}(B_i).$$

So if we define the conditional probability

$$\mathbb{P}[B_i|A] = \frac{\mathbb{P}(B_i)}{\mathbb{P}(A)},$$

we have

$$\sum_{i=1}^{k} \mathbb{P}[B_i|A] = 1.$$

The conditional probabilities give a probability measure on the successors of A: $\mathbb{P}[B_i|A]$ is the probability that B_i occurs given that A has occurred.

We can think of the conditional probabilities as being attached to edges in the information tree:

$$\begin{tikzcd} & B_1 \\ A \arrow[ur, "\mathbb{P}[B_1|A]"] \arrow[dr, "\mathbb{P}[B_2|A]"'] & \\ & B_2 \end{tikzcd}$$

$$\begin{tikzcd} & {} \\ 4 \arrow[ur, dash] \arrow[dr, dash] & \\ & 2 \end{tikzcd}$$

Conditional prob: what is prob of 4 given that we are at 2?

HW is just asking to work out probabilities

Going to prove:
risk neutral meas $\Leftrightarrow$ no arbitrage opport.
for multiple time steps

Note that

$$\mathbb{P}[B_i] = \mathbb{P}[A] \cdot \mathbb{P}[B_i|A].$$

Now suppose that Y is a random variable, in other words a contingent claim that pays $Y(\omega)$ in each state ω. The expectation of Y is of course

$$\mathbb{E}Y = \sum_{\omega \in \Omega} \mathbb{P}(\omega) Y(\omega).$$

If $A \in P_t$, and we know that $\omega \in A$, we can define a *conditional expectation* by

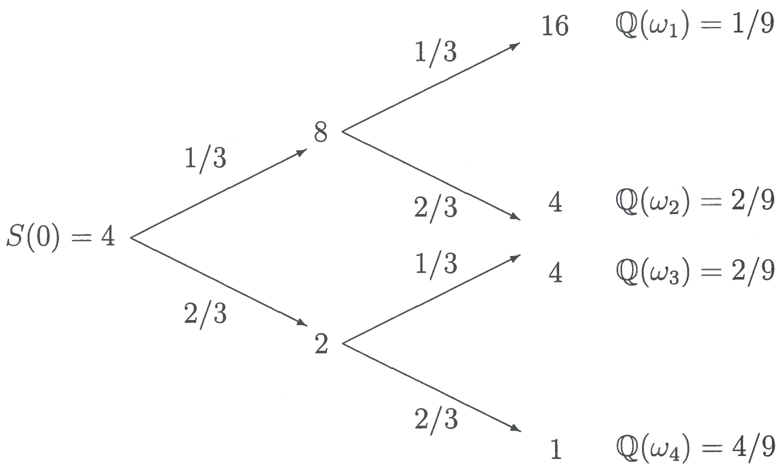
$$\mathbb{E}_{\mathbb{P}}[Y|A] = \sum_{\omega \in A} \frac{\mathbb{P}(\omega)}{\mathbb{P}(A)} Y(\omega).$$

We can evaluate this for each set $A \in P_t$, and gather together the results as a function.

Definition 1. *The* conditional expectation $\mathbb{E}_{\mathbb{P}}[Y|P_t]$ *is the function that, for each $A \in P_t$, has constant value on $\omega \in A$ equal to*

$$\mathbb{E}_{\mathbb{P}}[Y|P_t](A) = \sum_{\omega \in A} \frac{\mathbb{P}(\omega)}{\mathbb{P}(A)} Y(\omega).$$

For instance, consider our example with the measure $\mathbb{Q}$:



We have

$$\mathbb{Q}(\{\omega_1, \omega_2\}) = 1/3$$

and

$$\mathbb{E}_{\mathbb{Q}}(X(2)|P_1)(\omega_1) = \mathbb{E}_{\mathbb{Q}}(X|P_1)(\omega_2) = \frac{\mathbb{P}(\omega_1)X(\omega_1) + \mathbb{P}(\omega_2)X(\omega_2)}{\mathbb{P}(\{\omega_1, \omega_2\})}$$

which equals

$$\frac{(1/9)16 + (2/9)4}{1/3} = 8.$$

Similarly,

$$\mathbb{Q}(\{\omega_3, \omega_4\}) = 2/3$$

and

$$\mathbb{E}_{\mathbb{Q}}(X(2)|P_1)(\omega_3) = \mathbb{E}_{\mathbb{Q}}(X|P_1)(\omega_4) = \frac{\mathbb{P}(\omega_3)X(\omega_3) + \mathbb{P}(\omega_4)X(\omega_4)}{\mathbb{P}(\{\omega_3, \omega_4\})}$$

which equals

$$\frac{(2/9)4 + (4/9)1}{2/3} = 2.$$

So in fact, we have ended up with

$$\mathbb{E}_{\mathbb{Q}}[X(2)|P_1] = X_1.$$

What about $\mathbb{E}_{\mathbb{Q}}[X(2)|P_0]$? Since P_0 is the trivial partition, $\mathbb{E}_{\mathbb{Q}}[X(2)|P_0]$ has the constant value

$$\frac{\mathbb{P}(\omega_1)X(\omega_1) + \mathbb{P}(\omega_2)X(\omega_2) + \mathbb{P}(\omega_3)X(\omega_3) + \mathbb{P}(\omega_4)X(\omega_4)}{\mathbb{P}(\Omega)}$$

which equals

$$\frac{(1/9)16 + (2/9)4 + (2/9)4 + (4/9)1}{1} = 4.$$

So we have

$$\mathbb{E}_{\mathbb{Q}}[X(2)|P_0] = X_0,$$

and a similar calculation shows that

$$\mathbb{E}_{\mathbb{Q}}[X(1)|P_0] = X_0.$$

Could this be an accident?

12.2 Properties of conditional expectation

We shall need a few facts about conditional expectation. Throughout this section, we fix a probability measure $\mathbb{P}$ on Ω, and take conditional expectations with respect to this measure.

Lemma 2. *Let P be a partition of Ω and let X be a random variable. Then*

$$\mathbb{E}\big[\mathbb{E}(X|P)\big] = \mathbb{E}[X].$$

Proof. Recall that if $A \in P$, then for every $\omega \in A$ we have

$$\mathbb{E}(X|P)(\omega) = \mathbb{E}(X|A) = \sum_{\omega' \in A} \frac{\mathbb{P}(\omega')}{\mathbb{P}(A)} X(\omega'),$$

The lemma now follows by a straightforward calculation.

$$\begin{aligned}
\mathbb{E}\big[\mathbb{E}(X|P)\big] &= \sum_{\omega \in \Omega} \mathbb{P}(\omega)\mathbb{E}(X|P)(\omega) \\
&= \sum_{A \in P} \sum_{\omega \in A} \mathbb{P}(\omega)\mathbb{E}(X|P)(\omega) \\
&= \sum_{A \in P} \sum_{\omega \in A} \mathbb{P}(\omega)\mathbb{E}(X|A) \\
&= \sum_{A \in P} \sum_{\omega \in A} \mathbb{P}(\omega)X(\omega) \\
&= \sum_{\omega \in \Omega} \mathbb{P}(\omega)X(\omega) \\
&= \mathbb{E}X.
\end{aligned}$$

□

Thus the conditional expectation has the same expectation as the original random variable.

Another straightforward calculation shows that conditional expectation is linear.

Lemma 3. *Let P be a partition of Ω and let X and Y be random variables. Then*

$$\mathbb{E}(\lambda X + \mu Y|P) = \lambda\mathbb{E}(X|P) + \mu\mathbb{E}(Y|P).$$

Proof. If $A \in P$, we have

$$\begin{aligned}
\mathbb{E}(\lambda X + \mu Y|P)(A) &= \sum_{\omega \in A} \frac{\mathbb{P}(\omega)}{\mathbb{P}(A)}(\lambda X(\omega) + \mu Y(\omega)) \\
&= \lambda \sum_{\omega \in A} \frac{\mathbb{P}(\omega)}{\mathbb{P}(A)} X(\omega) + \mu \sum_{\omega \in A} \frac{\mathbb{P}(\omega)}{\mathbb{P}(A)} Y(\omega) \\
&= \lambda \mathbb{E}(X|P)(A) + \mu \mathbb{E}(Y|P)(A).
\end{aligned}$$

□

It is easy to check that if Y is *already* P-measurable then

$$\mathbb{E}(Y|P) = Y.$$

Slightly more is true.

Lemma 4. *Let P be a partition of Ω and let X and Y be random variables. If Y is P-measurable then*

$$\mathbb{E}(XY|P) = Y\mathbb{E}(X|P).$$

Proof. If $A \in P$, then Y has constant value $Y(A)$ on A. So

$$\begin{aligned}
\mathbb{E}(XY|P)(A) &= \sum_{\omega \in A} \frac{\mathbb{P}(\omega)}{\mathbb{P}(A)} X(\omega)Y(\omega) \\
&= Y(A) \sum_{\omega \in A} \frac{\mathbb{P}(\omega)}{\mathbb{P}(A)} X(\omega) \\
&= Y(A)\mathbb{E}(X|P)(A).
\end{aligned}$$

□

A particularly important property is known as the *Tower Law.*

Theorem 5. *(Tower Law) If $P_1 \leq P_2$ and X is a random variable then*

$$\mathbb{E}\big(\mathbb{E}(X|P_2)|P_1\big) = \mathbb{E}(X|P_1).$$

Thus if we condition X first on a partition P_2, and then on a less refined partition P_1, we get the same result as just conditioning X on P_1 in the beginning. Since this repeated conditioning is very common, it is often written in the form $\mathbb{E}(X|P_2|P_1)$.

Proof. Let A be a set in P_1 and suppose that P_2 partitions A into $B_1, \ldots, B_k$. Then

$$\mathbb{E}(X|P_2|P_1)(A) = \sum_{\omega \in A} \frac{\mathbb{P}(\omega)}{\mathbb{P}(A)} \mathbb{E}(X|P_2)(\omega)$$

(divide the sum into the separate sets B_i)

$$\begin{aligned}
&= \sum_{i=1}^{k} \sum_{\omega \in B_i} \frac{\mathbb{P}(\omega)}{\mathbb{P}(A)} \mathbb{E}(X|P_2)(B_i) \\
&= \sum_{i=1}^{k} \frac{\mathbb{P}(B_i)}{\mathbb{P}(A)} \mathbb{E}(X|P_2)(B_i)
\end{aligned}$$

(now expanding $\mathbb{E}(X|P_2)(B_i)$)

$$\begin{aligned}
&= \sum_{i=1}^{k} \frac{\mathbb{P}(B_i)}{\mathbb{P}(A)} \sum_{\omega \in B_i} \frac{\mathbb{P}(\omega)}{\mathbb{P}(B_i)} X(\omega) \\
&= \sum_{i=1}^{k} \sum_{\omega \in B_i} \frac{\mathbb{P}(\omega)}{\mathbb{P}(A)} X(\omega) \\
&= \sum_{\omega \in A} \frac{\mathbb{P}(\omega)}{\mathbb{P}(A)} X(\omega) \\
&= \mathbb{E}(X|P_1)(A).
\end{aligned}$$

This gives the required equality. □

12.3 Martingales

An important part in financial modelling is played by *martingales.*

Definition 6. *Let $(P_t)_{t=0}^T$ be a filtration, and suppose that $X = (X_t)_{t=0}^T$ is an adapted sequence of random variables. We say that X is a* martingale *if*

$$\mathbb{E}(X_{t+1}|P_t) = X_t$$

for every t.

In other words, at time t we expect that (on average) our position at time $t+1$ will be the same as our current position.

The martingale condition immediately implies a stronger condition.

Lemma 7. *If X is a martingale with respect to the filtration $(P_t)_{t=0}^T$ then*

$$\mathbb{E}(X_{t+s}|P_t) = X_t$$

for every $s, t \geq 0$.

Proof. We prove this by induction on s. If $s = 0$ then the statement is immediate (as X_t is already P_t-measurable). If $s = 1$, it follows from the martingale property.

Now suppose $s > 1$, and we have proved the statement for $s-1$. By the Tower Law and the martingale property, we have

$$\begin{aligned} \mathbb{E}(X_{t+s}|P_t) &= \mathbb{E}(X_{t+s}|P_{t+s-1}|P_t) \\ &= \mathbb{E}(X_{t+s-1}|P_t), \end{aligned}$$

which is equal to X_t by induction. □

With asset prices in mind, for any sequence $X_0, \ldots, X_T$ of random variables and $t \geq 1$, we define

$$\Delta X_t = X_t = X_{t-1}.$$

Recall that a sequence $H_1, \ldots, H_T$ of random variables is *previsible* (or *predictable*) if H_t is P_{t-1}-measurable for every t. Now consider the corresponding gains process: we can think of this as the outcome after we place a sequence of (previsible) bets.

previsible trading strategy "=" replicating portfolio

martingale "=" same expected value

i.e. value tomorrow = value today

Theorem 8. *If X is a martingale and H is a previsible process then the gains process*

$$G_t \equiv \sum_{i=1}^{t} H_i \Delta X_i$$

is a martingale.

Proof. Note first that

$$\begin{aligned}\mathbb{E}(\Delta X_t | P_{t-1}) &= \mathbb{E}(X_t - X_{t-1} | P_{t-1}) \\ &= \mathbb{E}(X_t | P_{t-1}) - \mathbb{E}(X_{t-1} | P_{t-1}) \\ &= 0,\end{aligned}$$

since $\mathbb{E}(X_t | P_{t-1}) = X_{t-1}$ by the martingale property, and $\mathbb{E}(X_{t-1} | P_{t-1}) = X_{t-1}$ by P_{t-1}-measurability.

Now

$$\begin{aligned}\mathbb{E}(G_t | P_{t-1}) &= \mathbb{E}(G_{t-1} + H_t \Delta X_t | P_{t-1}) \\ &= \mathbb{E}(G_{t-1} | P_{t-1}) + \mathbb{E}(H_t \Delta X_t | P_{t-1}).\end{aligned}$$

G_{t-1} is P_{t-1}-measurable, so we have $\mathbb{E}(G_{t-1} | P_{t-1}) = G_{t-1}$. Since H_t is also P_{t-1}-measurable, we have

$$\mathbb{E}(H_t \Delta X_t | P_{t-1}) = H_t \mathbb{E}(\Delta X_t | P_{t-1}) = 0.$$

So

$$\mathbb{E}(G_t (P_{t-1}) = G_{t-1}.$$

We conclude that the gains process G_t is a martingale. □

EXAM Q

12.4 Multiperiod No-Arbitrage Theorem

Consider once again a market model with assets $S_1, \ldots, S_N$ and sample paths Ω. We need to extend our definitions of arbitrage and risk-neutrality to a multiperiod model. Let's start with arbitrage.

don't need to add money

can see one time step ahead

Definition 9. *An* arbitrage opportunity $\mathbf{H}$ *is a self-financing, previsible trading strategy such that*

$$\begin{aligned} V_0(\mathbf{H}) &= 0 && \text{value today} = 0 \\ V_T(\mathbf{H}, \omega) &\geq 0 && \textit{for all } \omega \\ V_T(\mathbf{H}, \omega) &> 0 && \textit{for at least one } \omega \end{aligned}$$

Thus an arbitrage opportunity is a (self-financing, previsible) trading strategy that is guaranteed not to lose money, and makes a profit on at least one sample path.

We also need to deal with risk-neutral measures. Given an asset S_n, we have a corresponding sequence of prices $S_n(0), \ldots, S_n(T)$. We would like to say that "the expected price tomorrow is the same as the price today", but as with the one-period model we must allow for interest. We therefore work with the *discounted price process* $S_n^* = (S_n(0)^*, \ldots, S_n(T)^*)$, where all the prices have been discounted to time 0.

Definition 10. *A probability measure $\mathbb{Q}$ is a risk-neutral measure for the market with assets $S_1, \ldots, S_N$ if*

1. *$\mathbb{Q}(\omega) > 0$ for every $\omega \in \Omega$; and*

2. *For each asset S_n, the discounted price process S_n^* is a martingale with respect to $\mathbb{Q}$ and the filtration $(P_i)_{i=1}^{T}$ given by the information tree*

This is just like the one-period definition: every outcome has a positive probability of happening, and (by the martingale property) pricing by conditional expectation gives the correct answer.

Theorem 11 (Multiperiod No-Arbitrage Theorem)**.** *In a multiperiod model, there is no arbitrage opportunity if and only if there is a risk-neutral probability measure.*

Proof. Suppose first that $\mathbb{Q}$ is a risk-neutral probability measure. If $\mathbf{H}$ is an arbitrage opportunity with $V_T(\mathbf{H}, \omega_0) > 0$, then

prog by contradiction

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}} V_T(\mathbf{H})^* &= \sum_{\omega \in \Omega} \mathbb{Q}(\omega) V_T(\mathbf{H}, \omega)^* \\
&\geq \mathbb{Q}(\omega_0) V_T(\mathbf{H}, \omega_0)^* \\
&> 0.
\end{aligned}$$

On the other hand, consider an asset S_n. Since the price process S_n^* is a martingale with respect to $\mathbb{Q}$, and H_n is previsible, then the gains process

$$G_n(t) \equiv \sum_{i=1}^{t} H_n(i) \Delta S_n^*$$

from previous theorem

is also a martingle. So

$$\mathbb{E}_{\mathbb{Q}} G_n(T) = \mathbb{E}_{\mathbb{Q}} G_n(0) = 0.$$

(Note that the gains process is defined for $t \geq 0$, and $G_n(0) \equiv 0$.)

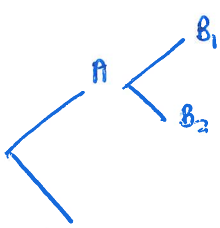
Summing over n, we have

$$\begin{aligned}
\mathbb{E} V_T(\mathbf{H})^* &= \mathbb{E}_{\mathbb{Q}}(V_T(\mathbf{H})^* - V_0(\mathbf{H})) \\
&= \sum_{n=1}^{N} \mathbb{E}_{\mathbb{Q}} G_n(T) \\
&= 0,
\end{aligned}$$

which gives a contradiction. We conclude that if there is a risk-neutral measure then there is no arbitrage opportunity.

Now for the second part of the proof. Suppose that there is no arbitrage opportunity in the multiperiod model. We claim first that there is no arbitrage opportunity in the separate one-period submodels defined between each node on the information tree and its children.

Suppose that there is a (one-period) arbitrage opportunity H_A at a node A on the tree, with children $B_1, \ldots, B_k$. We construct a multiperiod arbitrage opportunity H as follows. We begin with the empty portfolio, and at any other node that A or one of the B_i we leave the portfolio unchanged. At A, we construct the portfolio H_A, and at any of the B_i we convert the portfolio we are holding into the riskless asset (which is then held until time T). It is clear



that the trading strategy is self-financing (since H_A has zero construction cost), and also that it is previsible (our actions depend only on which node we're at in the information tree). What happens on a sample path ω? If ω does not pass through A then we have held the trivial portfolio throughout and end up with at least value 0. If ω passes through A, then we cashed in H_A after a single timestep and either end up with 0, or along at least one branch, with a positive profit.

We can therefore assume that none of the one-period models have an arbitrage opportunity, and so (by the single period No-Arbitrage Theorem) each has a (one-period) risk-neutral measure. We construct a measure $\mathbb{Q}$ on Ω by letting $\mathbb{Q}(\omega)$ be the product of the probabilities of the individual steps in ω according to the single-period risk-neutral measures at the nodes it visits.

Let us check that $\mathbb{Q}$ is a risk-neutral probability measure. Note first that $\mathbb{Q}(\omega) > 0$ for every ω, as all the individual steps have positive probability, so all we have to check is the martingale property.

Consider a node $A \in P_t$ with successors $B_1, \ldots, B_k$, and let $\mathbb{Q}_A$ be the one-step risk-neutral probability at A. It is easily checked that

$$\mathbb{Q}(B_i|A) = \mathbb{Q}_A(B_i),$$

just going to call it this

and in the one-period model at A we have

$$\mathbb{E}_{\mathbb{Q}_A} S_n(t+1)^* = S_n(A)^*.$$

Calculating, we find that

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}(S_n(t+1)^*|P_t)(A) &= \mathbb{E}_{\mathbb{Q}}(S_n(t+1)^*|A) \\
&= \sum_{k=1}^{n} \frac{\mathbb{Q}(B_i)}{\mathbb{Q}(A)} S_n(B_i)^* \\
&= \sum_{k=1}^{n} \mathbb{Q}(B_i|A) S_n(B_i)^* \\
&= \sum_{k=1}^{n} \mathbb{Q}_A(B_i) S_n(B_i)^* \\
&= \mathbb{E}_{\mathbb{Q}_A} S_n(t+1)^* \\
&= S_n(A)^*,
\end{aligned}$$

since $\mathbb{Q}_A$ is risk-neutral. It follows that $\mathbb{E}_{\mathbb{Q}}(S_n(t+1)^*|P_t) = S_t^*$. So S_n is a martingale.

We conclude that $\mathbb{Q}$ is a risk-neutral measure. $\square$

Thus we have deduced the Multiperiod No-Arbitrage Theorem from the single period version.

13 American options and stopping times

13.1 European and American options

Let $(P_t)_{t=0}^T$ be a filtrations. A *European option* is a random variable X, where $X(\omega)$ represents the payoff at time T. With an American option, we have the choice of exercising the option earlier.

Definition 12. *An* American option *is an adapted sequence* $(X_t)_{t=1}^T$ *of random variables.*

For European options in a multiperiod model, we can value using a risk-neutral probability measure (at least in the case when the one-period submodels are all complete): the price at time t of an option X is

$$V_{\text{euro}}(t) = \mathbb{E}_{\mathbb{Q}}(X(T)|P_t).$$

[The prices are enforced by the one-period risk-neutral measures; and the one-period measures fit together to give the multi-period risk-neutral measure.]

For American options the situation is different, and risk-neutral probability measures no longer provide a short-cut to pricing. We need a more sophisticated approach.

13.2 Dynamic programming

Dynamic programming describes the process of stepping backwards through time using a binomial tree model and consider the value of the claim using the one-period submodel. Different claims may have a one-period submodel value that differs from the value for a European. This method can be used to price a wide class of *exotic* options.

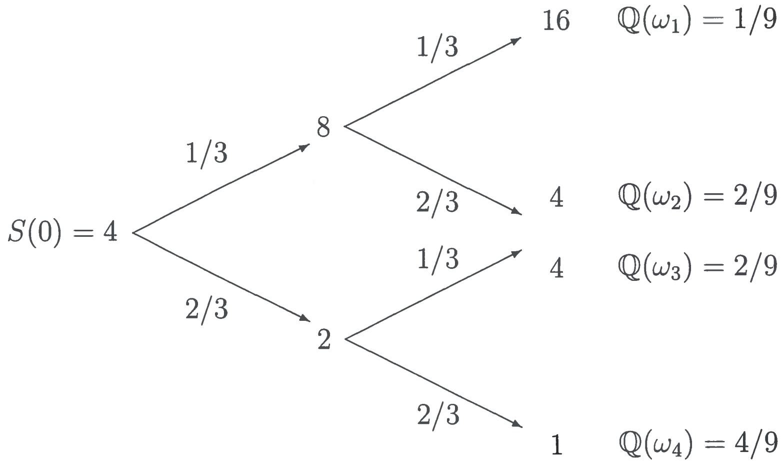
For example, consider a *barrier* call option where the option disappears if the asset goes above a given barrier level at any time during the life of the option (this is known as a *reverse knock-out* option). For each one-period submodel we know that if the asset's *path* ends higher than the barrier level then the value of the claim is zero.

We shall use dynamic programming to value american options with and without dividends:

Value of exercising = 8-3 = 5
option = 5

Exercise w/ Am. options because of dividends

Consider the following market model, where the (unique) risk-neutral measure has been marked.

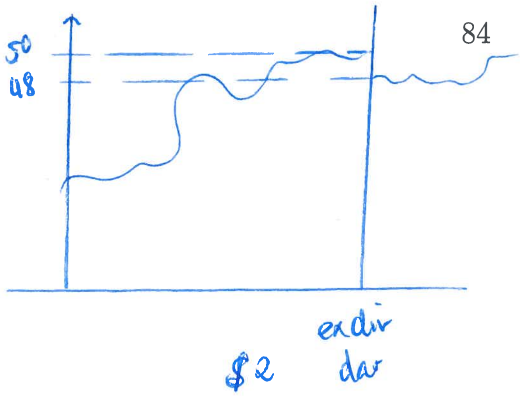


Example 13. *Consider an American call option with strike price* 3. *Let us value this using a* dynamic programming *method. We clearly have*

$$\begin{aligned} V_{\text{amer}}(\{\omega_1\}) &= 13 \\ V_{\text{amer}}(\{\omega_2\}) &= 1 \\ V_{\text{amer}}(\{\omega_3\}) &= 1 \\ V_{\text{amer}}(\{\omega_4\}) &= 0. \end{aligned}$$

What about $V_{\text{amer}}(\{\omega_1, \omega_2\})$? At $\{\omega_1, \omega_2\}$ we have two choices: either exercise the option, which is then worth $8 - 3 = 5$, or decide to wait. In that case, the option is equivalent to a one-period European call option, which has price $(1/3) \cdot 13 + (2/3) \cdot 1 = 5$. It follows that $V_{\text{amer}}(\{\omega_1, \omega_2\}) = 5$. At $\{\omega_3, \omega_4\}$, we do not profit if we exercise the option, but the associated one-period European option has price $(1/3) \cdot 1 + (2/3) \cdot 0 = 1/3$. It follows that we should not exercise the option, and $V_{\text{amer}}(\{\omega_1, \omega_2\}) = 1/3$

Finally, what about the time 0 price? If we exercise we receive a profit of $4 - 3 = 1$, but the associated one-step European option has value $(1/3) \cdot 5 + (2/3) \cdot (1/3) = 17/9$ so we should not exercise at this step. It follows that the time 0 value is $17/9$, which agrees with the European option, as we expected.

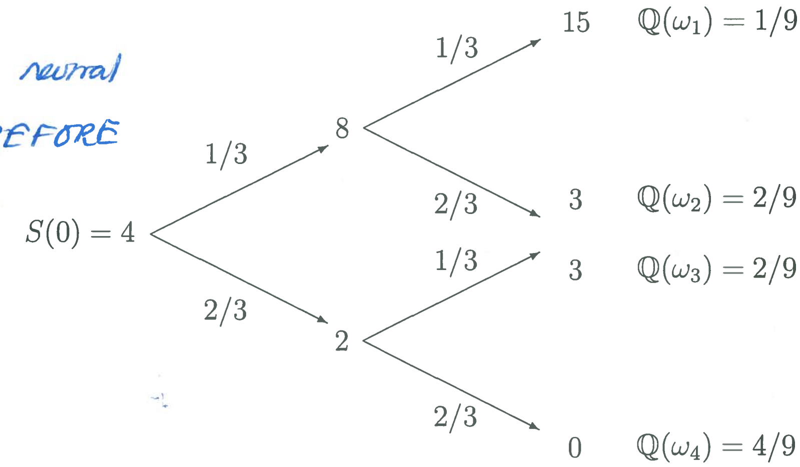


Company decides to pay $2 per share

Value drops by $2 almost instantaneously

Now consider the same market model, except that the asset pays a *dividend* of 1 at time 1.5. This means that the holder of each unit of the asset is paid 1 unit of cash, and the asset price correspondingly drops by 1. (The risk neutral probabilities don't change, as the total of asset price plus dividend is unchanged). So the asset prices are now:

Calculate risk neutral prob measures BEFORE you take off your dividend



Example 14. *Consider an American call option with strike price* 3 *in the dividend-paying example. Now we have*

$$V_{\text{amer}}(\{\omega_1\}) = 12$$

and

$$V_{\text{amer}}(\{\omega_2\}) = V_{\text{amer}}(\{\omega_3\}) = V_{\text{amer}}(\{\omega_4\}) = 0.$$

What about $V_{\text{amer}}(\{\omega_1, \omega_2\})$*? Exercising the option is worth* $8 - 3 = 5$*. Waiting is equivalent to a one-period European call option, which has price* $(1/3) \cdot 12 + (2/3) \cdot 0 = 4$*. It follows that we should exercise the option, and* $V_{\text{amer}}(\{\omega_1, \omega_2\}) = 5$*.*

Clearly $V_{\text{amer}}(\{\omega_1, \omega_2\}) = 4 - 3 = 1$*, and now at time 0, we have either* 1 *from exercising the option, or* $(1/3)5 + (2/3)0 = 5/3$ *from the one-period call option. So the time 0 value of the option is* $5/3$*.*

Note that the asset holder receives the divided, *but the holder of the option does not.*

13.3 Stopping times

At each step in the argument above, we had the option of stopping or continuing, and we made the decision based on the node we were at.

Definition 15. *A random variable* $\tau : \omega \to \{0, 1, \ldots\}$ *is called a* stopping time *if for every* $1 \le t \le T$ *we have*

$$\{\omega \in \Omega : \tau(\omega) = t\}$$

is a union of sets in P_t.

In other words, the sample paths on which we stop at time t correspond to a set of nodes at time t. Note that we allow $t > T$: this corresponds to a decision not to stop (or not to exercise the option).

In both of the examples above, the valuation corresponded to a stopping time. In the first example, we always stopped at time 2, so the stopping time is a constant function. In the second example, on sample paths ω_1 and ω_2 we stopped at time 1, and on the other sample paths we could continue to time 2.

Following through the dynamic programming approach, we get for any American option Y_t an associated value process Z_t, where (assuming that $r = 0$)

$$Z_T(\omega) = Y_T(\omega)$$

for every ω (at time T the value is immediately determined), and for $t < T$, and any $A \in P_t$,

$$Z_t(A) = \max_{\tau} \mathbb{E}_{\mathbb{Q}}[Y_\tau | P_t],$$

where τ is any stopping time with $t \le \tau(\omega) \le T$ for all ω.

The process Z_t satisfies (with $r = 0$) the condition $\mathbb{E}_{\mathbb{Q}}(Z_{t+1}|P_t) \le Z_t$ (as having the choice of whether to exercise at time t increases the value of the option). If we don't assume that $r = 0$, we get

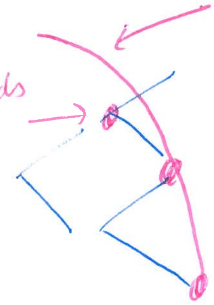
$$\mathbb{E}_{\mathbb{Q}}(Z^*_{t+1}|P_t) \le Z^*_t.$$

Processes that satisfy this inequality are known as *supermartingales*.

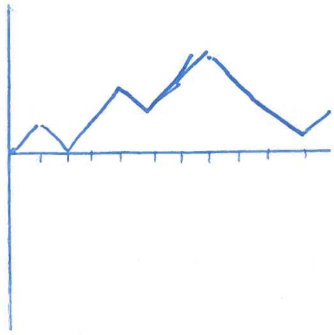
It can be shown that Z^*_t is the smallest supermartingale with $Z_t \ge Y^*_t$, but we won't pursue this.

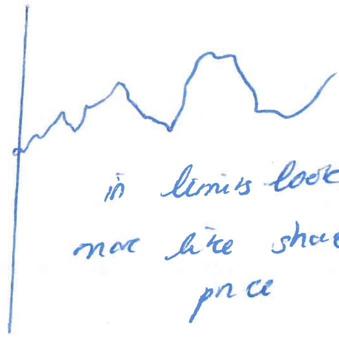
exercise boundary = stopping time

this tends to bdry as model gets bigger



This is more relevant to larger models





in limits looks more like share price

14 Brownian motion and stochastic calculus

14.1 Brownian motion

We now consider modelling the evolution of asset prices by continuous processes. The most important process for our purposes will be *Brownian motion*.

Let's begin by considering a (discrete time) random walk S_n, where successive steps are independent and occur at time intervals of length $1/n$. We define

$$S_n(0) = 0,$$

and at the the beginning of each time interval we increase or decrease S_n by $1/\sqrt{n}$. So, for $t \in [i/n, (i+1)/n)$, we have

$$S_n(t) = S_n\left(\frac{i}{n}\right) = S_n\left(\frac{i-1}{n}\right) \pm \frac{1}{\sqrt{n}},$$

where we take increments $+1/\sqrt{n}$ and $-1/\sqrt{n}$ with equal probability.

Why increments of size $1/\sqrt{n}$? The point is that if we take steps of size $\pm\rho$ then individual increments then have variance ρ^2. Since successive increments are independent, and we take nt steps by time t (at least when t is a multiple of $1/n$), we have

$$\text{var}(S_n(t)) = n\rho^2 \cdot t.$$

So in order to keep the variance fixed, we must scale the increments as $1/\sqrt{n}$.

What happens as $n \to \infty$? We can use the De Moivre-Laplace theorem, which tells us that for any fixed t, $S_n(t)$ converges (in a suitable sense) to a normal distribution with mean 0 and variance t. Similarly, for $t' > t$, $S_n(t') - S_n(t)$ converges to a normal distribution with mean 0 and variance $t' - t$, which is independent of $S_n(t)$ (since it depends on a different collection of random increments).

We won't analyze this convergence in any detail. Instead, we will describe more directly what we expect the limiting process to look like.

Definition 1. *A process* $\left(B(t)\right)_{t\geq 0}$ *is said to be a* (standard) Brownian motion *if the following four conditions are satisfied*

1. $B(0) = 0$ this is a convention - it is standard

2. *For* $s, t \geq 0$, *the random variable*

$$B(s+t) - B(s)$$

is normally distributed with mean 0 and variance t

3. *Whenever* $0 \leq t_0 < t_1 < \cdots < t_n$, *the quantities*

$$B(t_1) - B(t_0), B(t_2) - B(t_1), \ldots, B(t_n) - B(t_{n-1})$$

are independent whether it goes up or down - it is independent

4. *With probability 1,* $B(t)$ *is a continuous function of t*

The condition $B(0) = 0$ is a convention, saying that our sample paths start from the origin. The Brownian motion is *standard* because, for any s and t, the random variable $B(s+t) - B(s)$ is normal with mean 0 and variance t: we could have written down a more general definition with variance $\sigma^2 t$ at time t. The third condition says that *increments are independent*, and the fourth condition says that *sample paths are continuous.*

Brownian motion has been around for a long time. The idea of Brownian motion originated in 1828, in the work of the botanist Robert Brown. It was first introduced as a model of stock market prices by Louis Bachelier in 1900, in his thesis "Théorie de la spéculation", while the mathematical foundations of Brownian motion were developed by Norbert Wiener in the 1920s (in fact, Brownian motion is often referred to as a *Wiener process*). Stochastic integration (which we shall discuss later) was developed by Kiyoshi Itô in the 1940s.

Below we state some results without proof to show that Brownian motion is stable under various operations.

Lemma 2. *Suppose that* $\big(B(t)\big)_{t\geq 0}$ *is a Brownian motion.*

1. *For any* $s \geq 0$, $(B(t+s) - B(s))_{t\geq 0}$ *is a Brownian motion.*
2. *For any* $c > 0$, $\big(cB(t/c^2)\big)_{t\geq 0}$ *is a Brownian motion.*
3. $\big(tB(1/t)\big)_{t\geq 0}$ *is a Brownian motion, if we define* $tB(1/t) = 0$ *when* $t = 0$.

Although Brownian motion is continuous, it is far from smooth.

Theorem 3. *With probability 1, the sample path* $\big(B(t,\omega)\big)_{t\geq 0}$ *is nowhere differentiable.*

Note that showing that $B(t)$ fails with probability to be differentiable at a particular point is not the same as showing it is *nowhere* differentiable.

Probably the most important theorem about Brownian motion is the following:

Theorem 4. *Brownian motion exists.*

More precisely, it is possible to construct a random process that has all the properties of Brownian motion. For example, there is a construction due to Lévy, that works by building successively finer random walks.

14.2 Stochastic calculus

Let us consider an asset S whose price evolves in continuous time. Between time t and $t + \Delta t$, the price of the asset changes by

$$\Delta S(t) = S(t + \Delta t) - S(t).$$

If S is a riskless asset, then $\Delta S(t)$ is a function of the (continuously compounded) interest rate, and we have

$$\Delta S(t) \approx r\Delta t \cdot S(t).$$

This gives

$$dS = rSdt,$$

or more formally the familiar differential equation

$$\frac{dS}{dt} = rS,$$

which has solution

$$S(t) = S(0)e^{rt}.$$

More generally, we can write

$$dS = \mu(t)Sdt,$$

where $\mu(t)$ measures the return on the asset at time t.

We need a *random* element as well: the stock price is a random process. More precisely, we want to model the changes in stock price as

$$\Delta S(t) \approx r\Delta t \cdot S(t) + \sigma S \Delta W,$$

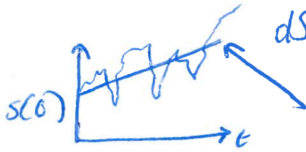
where ΔW is some sort of random increment and σ is a scaling factor (to fix the variance). We end up with the *stochastic differential equation*

W nowhere diff'able but

this is really shorthand for $\int .. dS = \int .. dt + \int dW$ $\longrightarrow$ * $$dS = rSdt + \sigma SdW.$$ This is a trick - stochastic Diff. Eq. SDE

It's not yet clear what this means, but note at least that if $\sigma = 0$ (if there is no random component) then we are back on firm ground with a normal differential equation.

$dS = rSdt + \sigma SdW$

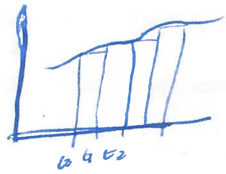


growth & random component around it

Switch from calling Brownian B to W (after Norman Wiener)

∅ Since W not diff'able
we need to define dW

We cannot write $\frac{dS}{dt} = rS + \sigma S \frac{dW}{dt}$



What should our random increments look like? The natural choice is to let $W(t)$ be a Brownian motion, so that

$$\Delta W(t) = W(t + \Delta t) - W(t)$$

is a normal random variable with mean 0 and variance Δt. Our random increments are then independent and normally distributed.

Unfortunately, we can't convert our stochastic differential equation into an expression involving $\frac{dS}{dt}$, as we would then have a term of form $\frac{dW}{dt}$: we can't expect to be able to do this, as we already know that Brownian motion is (with probability 1) nowhere differentiable.

What we can do is to interpret the stochastic differential equation as shorthand for an integral equation

$$S(T) = S(0) + \int_0^T r(t)S(t)dt + \int_0^T \sigma S(t)dW.$$

The first integral is the familiar Riemann integral. Recall that the usual definition of the Riemann integral $\int_0^T f(t)dt$ involves summing over partitions of an interval: for a function f and a partition π given by $0 = t_0 < t_1 < \cdots < t_N = T$, we approximate the integral by

$$\sum_{i=0}^{N-1} f(x_i) \cdot (t_{i+1} - t_i),$$

where x_i is any point in $[t_i, t_{i+1}]$. If the function is integrable, then this converges to $\int_0^T f(t)dt$ as $||P|| \to 0$, where $||P|| = \max |t_{i+1} - t_i|$.

What about the second integral? We can try doing the same thing as we did with the Riemann integral: for a function f and a partition π given by $0 = t_0 < t_1 < \cdots < t_N = T$, we approximate the *stochastic* integral

$$\int_0^T f(t)dW$$

by the sum

$$I(\pi) = \sum_j f(t_j) \cdot \underbrace{(W(t_{j+1}) - W(t_j))}_{\text{BROWNIAN MOTION}}.$$

multiply over the path it takes

The limit of these sums (in an appropriate sense, and if it exists) is known as the *Itô integral.*

"Eeto"

When we consider normal Riemannian integral - don't care which intervals we look at

Not true for Brownian motion: since we have convergence
Brownian motion path dependent - use LEFT HAND interval
which we know for previs. Brownian motion

Before we try evaluating an integral of this type, a few comments are in order.

Remark. *Unlike the Riemann integral, the Itô integral is a random variable: its value depends on the particular sample path we are following, and so we are really interested in its* distribution.

you cannot Riemann use this method for Ito

Remark. *In the Riemann integral, we could take x_j to be any point in $[t_j, t_{j+1}]$ without affecting the limit. In the Itô integral, we insist that $x_j = t_j$, i.e. we choose the left-hand end of the interval. This means that the value of $f(x_j)$ is* previsible, *in the sense that we know it already at the beginning of the interval. Choosing x_j to be (for instance) at the end of the interval can lead to a different result.*

Remark. *We can allow f to depend on the time t and the history of S up to time t (keeping the idea of previsibility).*

Let's try an example.

14.3 A first example

Integral of Brownian motion w.r.t. Brownian motion

Let's try to work out the integral

$$\int_0^T W(t)dW(t).$$

Let π be the partition given by $0 = t_0 < t_1 < \cdots < t_N = T$. Then

$$I(\pi) = \sum_{i=0}^{N-1} W(t_i) \cdot (W(t_{i+1}) - W(t_i)).$$

looking at end point

Neither of the resulting terms, $W(t_i)W(t_{i+1})$ or $W(t_i)^2$ looks very promising. So what can we do? If we were summing expressions of form $W(t_{i+1})^2 - W(t_i)^2$ we would be much happier, as we would get lots of cancellation. Similarly, if we were summing expressions involving just the increments $\Delta W(t_j) = W(t_{j+1}) - W(t_j)$ then we would have some chance of getting somewhere, as successive increments are independent.

So let's try to decompose the sum into terms of these types. After some thought, we find

$$
\begin{aligned}
I(\pi) &= \sum_{i=0}^{N-1} \left(W(t_i)W(t_{i+1}) - W(t_i)^2\right) \qquad * \\
&= \frac{1}{2}\sum_{i=0}^{N-1} \left(W(t_{i+1})^2 - W(t_i)^2\right) - \frac{1}{2}\sum_{i=0}^{N-1} \left(W(t_{i+1}) - W(t_i)\right)^2 \\
&= \frac{1}{2}W(T)^2 - \frac{1}{2}W(0)^2 - \frac{1}{2}\sum_{i=0}^{N-1} \Delta W(t_i)^2.
\end{aligned}
$$

What can we say about the last sum? We know that $\Delta W(t_i) = W(t_{i+1}) - W(t_i)$ is normal, with mean 0 and variance $t_{i+1} - t_i$. So

$$\mathbb{E}\left(\Delta W(t_i)^2\right) = t_{i+1} - t_i.$$

variance proportional to time

It follows that

$$\mathbb{E}\sum_{i=0}^{N-1} \Delta W(t_i)^2 = \sum_{i=0}^{N-1}(t_{i+1} - t_i) = T.$$

With a little more technical details it is possible to show that indeed the sum does converge to T as $||\pi|| \to 0$ as suggested by the expectation.

Since $W(0) = 0$, putting together the calculations above gives

$$\int_0^T W(t)dW(t) = \frac{1}{2}W(T)^2 - \frac{1}{2}T.$$

will justify in a moment

Note the difference between the Itô integral and the Riemann integral, where

$$\int_0^T tdt = \frac{1}{2}T^2.$$

In the Itô integral, we seem to end up with an additional term $-T/2$.

* to see this

$W_{t_j} = A \qquad W_{t_{j+1}} = B$

$\frac{1}{2}(B^2 - A^2) - \frac{1}{2}(B-A)^2 = \frac{1}{2}(2BA - 2A^2)$

$W_0 = 0$

value of function w.r.t partition
since it is dependent.

14.4 Filling in the details

In this section, we'll fill in some of the technical details in the definition of the Itô integral. The main aim is to indicate how the definitions and calculations in earlier sections could be made fully rigorous. We will then revert to a slightly more informal style for the rest of the course.

Consider first a function $f : [0, T] \to \mathbb{R}$. We can approximate f by taking a partition π given by $0 = t_0 < t_1 < \cdots < t_N = T$, and using the function f_π defined by

$$f_\pi(t) = f(t_j) \qquad \forall t \in [t_j, t_{j+1}).$$

Since

$$\int_0^T f_\pi(t) dt = \sum_{i=0}^{N} f(t_j)(t_{j+1} - t_j),$$

which is the estimate we use in the Riemann integral, we see that

$$\lim_{||\pi|| \to 0} \int_0^T f_\pi(t) dt = \int_0^T f(t) dt,$$

where as before we define

norm-limit of partition as → 0

$$||\pi|| = \min |t_{j+1} - t_j|.$$

Let's try the same approach with the Itô integral. Once again, we consider a real function f for times $t \in [0, T]$, but now the function also depends on the state $\omega \in \Omega$ (for instance, if f is the price of an asset, then for each possible world ω we have a sample path $(f(t, \omega))_{0 \le t \le T}$ giving the evolution of the asset price over time). So we actually have a function

$$f : [0, T] \times \Omega \to \mathbb{R}$$

(or, equivalently, for each possible world ω we have a function $f_\omega : [0, t] \to \mathbb{R}$).

Given the partition π, we define the simple function f_π as before, by

$$f_\pi(t, \omega) = f(t_j, \omega) \qquad \forall t \in [t_j, t_{j+1}).$$

(In other words, we take the same approximation as before, but for each possible world ω we are approximating a different function f_ω, and so get a different $f_{\pi,\omega}$.)

dependent on path we take

We *define*

$$\int_0^T f_\pi dW = \sum_{j=0}^{N-1} f_\pi(t_j)(W(t_{j+1}) - W(t_j)).$$

Again, both sides depend on ω (so both sides are random variables: note that this holds for both f_π and W). So for each ω we get a specific value (i.e. a real number) on each side.

The idea now is that the random variable $\int_0^T f_\pi dW$ should approach $\int_0^T f dW$. But first we need a measure of how well f_π approximates f. For a given ω, we measure approximation by

$$\int_0^T (f(t) - f_\pi(t))^2 dt.$$

function of path

But this depends on which possible world ω we are in, so we measure the approximation by taking an expectation:

$$\mathbb{E}\left[\int_0^T (f(t) - f_\pi(t))^2 dt\right].$$

We can now define the integral.

Definition 5. *If (π_n) is a sequence of partitions such that*

$$\mathbb{E}\left[\int_0^T (f(t) - f_\pi(t))^2 dt\right] \to 0$$

as $n \to \infty$, then

$$\int_0^T f_\pi(t) dW \to \int_0^T f(t) dW$$

as $n \to \infty$.

def of convergence looking at L^2 norm second order diff'able

There's still one more technical detail. We haven't said in what sense the integrals $\int_0^T f_\pi(t)dW$ converge to $\int_0^T f(t)dW$. Both sides are random variables (i.e. they depend on ω), so we need to use some measure of distance between random variables. The relevant measure is the distance in $L^2(\mathbb{P})$: we say that random variables $X_n \to X$ if

$$\mathbb{E}((X_n - X)^2) \to 0$$

This is measure theory & stochastic calculus can read about it in more detail.

as $n \to \infty$. (Or we could say that the mean and variance of the difference tend to 0.) Thus in our example, we are demanding

$$\mathbb{E}\left[\left(\int_0^T f(t)dW - \int_0^T f_\pi(t)dW\right)^2\right] \to 0$$

as $n \to \infty$. (This is more restrictive than convergence in $L^1(\mathbb{P})$ but less restrictive than uniform convergence.)

Remark. *Actually, there is slightly more technical detail, although we've now seen the more important parts (the remaining details are mostly best expressed in the language of measure theory).*

Let's see how this detail works out in practice, by filling in the details in our derivation of

$$\int_0^T W(t)dW(t).$$

Writing $f(t) = W(t)$, we have

$$\begin{aligned}\int_0^T (f(t) - f_\pi(t))^2 dt &= \sum_{i=0}^{N-1} \int_{t_j}^{t_{j+1}} (f(t) - f(t_j))^2 dt \\ &= \sum_{i=0}^{N-1} \int_{t_j}^{t_{j+1}} (W(t) - W(t_j))^2 dt.\end{aligned}$$

Now

$$\mathbb{E}\left[(W(t) - W(t_j))^2\right] = t - t_j,$$

variance

as $W(t) - W(t_j)$ is normal with mean 0 and variance $t - t_j$. So

$$\begin{aligned}\mathbb{E}\left[\int_0^T (f(t) - f_\pi(t))^2 dt\right] &= \sum_{i=0}^{N-1} \mathbb{E} \int_{t_j}^{t_{j+1}} (W(t) - W(t_j))^2 dt \\ &= \sum_{i=0}^{N-1} \int_{t_j}^{t_{j+1}} t - t_j dt \\ &= \sum_{i=0}^{N-1} \frac{1}{2}(t_{j+1} - t_j)^2 \\ &\to 0\end{aligned}$$

as $||\pi|| \to 0$.

It follows that
$$\int_0^T f_\pi dW \to \int_0^T f dW,$$
where the convergence is in $L^2(\mathbb{P})$.

Now, arguing as we did before, we see that
$$\begin{aligned}\int_0^T f_\pi dW &= \sum_j W(t_j)(W(t_{j+1}) - W(t_j)) \\ &= \frac{1}{2}\sum_j \left(W(t_{j+1})^2 - W(t_j)^2\right) - \frac{1}{2}\sum_j \left(W(t_{j+1}) - W(t_j)\right)^2 \\ &= W(T)^2 - \frac{1}{2}\sum_j (\Delta W(t_j))^2.\end{aligned}$$

In order to show that
$$\int_0^T f_\pi dW \to W(T)^2 - \frac{1}{2}T$$
in $L^2(\mathbb{P})$, it is enough to show that
$$\sum_j \Delta W(t_j)^2 \to T$$
in $L^2(\mathbb{P})$. As $\sum_j \Delta W(t_j)^2 - T = \sum_j (\Delta W(t_j)^2 - \Delta t_j)$, we have
$$\mathbb{E}\left[\left(\sum_j \Delta W(t_j)^2 - T\right)^2\right] = \sum_{i,j} \mathbb{E}\big[(\Delta W(t_i)^2 - \Delta t_i)(\Delta W(t_j)^2 - \Delta t_j)\big].$$
If $i \neq j$ then $\Delta W(t_i)^2 - \Delta t_i$ and $\Delta W(t_j)^2 - \Delta t_j$ are independent. Since
$$\mathbb{E}\big[\Delta W(t_i)^2 - \Delta t_i\big] = \Delta t_i - \Delta t_i = 0,$$
we have
$$\mathbb{E}\big[(\Delta W(t_i)^2 - \Delta t_i)(\Delta W(t_j)^2 - \Delta t_j)\big] = 0.$$
We are left with
$$\sum_i \mathbb{E}\big[(\Delta W(t_i)^2 - \Delta t_i)^2\big] = \sum_i \mathbb{E}\Delta W(t_i)^4 - 2\sum_i \mathbb{E}\Delta W(t_i)^2 \Delta t_i + \sum_i (\Delta t_i)^2.$$
It is now easily checked that each term in each of the sums is $O((\Delta t_i)^2)$ (note that you can calculate $\mathbb{E}\Delta W(t_i)^4$ exactly, as $\Delta W(t_i) \sim N(0, t_{i+1} - t_i)$) and so the sums all converge to 0.

This proves convergence of the whole thing

non EXAMINABLE

15 Itô's Lemma

— analogy of integration by parts

15.1 Itô's Lemma: time-independent version

The stochastic integral is an important idea, but calculating directly from the definition is quite a lot of work, as we have seen with $\int W(t)dW$. (Of course, a similar thing can be said for the usual differential calculus: usually we don't work directly from the definitions.) The tool we need is called *Itô's Lemma*.

Let us suppose that the asset price S satisfies a stochastic differential equation

$$dS = \mu dt + \sigma dW,$$

dropping the S at $\mu S dt$

where σ and μ are functions that depend only on t and the history of $W(s)$ for $s \leq t$. Recall that, in integral form, this says that

$$S(T) - S(0) = \int_0^T \mu dt + \int_0^T \sigma dW.$$

Now consider a function

$$f(S, t)$$

of asset price, where f has continuous second derivatives. What can we say about df?

For simplicity, let us first deal with the case when $f = f(S)$ is independent of time (we will come back to the time-dependent case in a later section). By Taylor's theorem, provided f has continuous second derivative, we have

$$f(S + \Delta S) = f(S) + f'(S)\Delta S + \frac{1}{2}f''(S)(\Delta S)^2 + o((\Delta S)^2).$$

take squared terms since this is L^2 norm

Now

$$\Delta S = \mu \Delta t + \sigma \Delta W$$

and so

$$(\Delta S)^2 = \mu^2(\Delta t)^2 + 2\mu\sigma\Delta t\Delta W + \sigma^2(\Delta W)^2.$$

The first two terms on the right are small compared with Δt; but since ΔW has order $\sqrt{\Delta t}$, the third term has order Δt. We get

$$(\Delta S)^2 = \sigma^2(\Delta W)^2 + o(\Delta t).$$

normally Q on Itô's Lemma
Q on Black scholes

It follows that

$$f(S+\Delta S) - f(S) = f'(S)(\mu\Delta t + \sigma\Delta W) + \frac{1}{2}f''(S)\sigma^2(\Delta W)^2 + o(\Delta t).$$

Now as W is Brownian motion and $\Delta W = W(t+\Delta t) - W(t)$, we have $\mathbb{E}\Delta W = 0$, while

$$\mathbb{E}(\Delta W)^2 = \operatorname{var}(\Delta W) = \Delta t.$$

In the limit, we can replace $(\Delta W)^2$ by Δt in our expression for $f(t+\Delta t) - f(t)$ (this step is valid as $\Delta t \to 0$, because we have many independent increments $(\Delta W)^2$; however, we won't give a detailed proof that this is correct).

(jumping over technical details)

We obtain the following.

Theorem 1 (Itô's Lemma: time-independent version)**.** *Let $f(S)$ be continuously twice-differentiable, and suppose that*

$$dS = \mu dt + \sigma dW.$$

Then

$$df = \frac{\partial f}{\partial S}(\mu dt + \sigma dW) + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial S^2}dt.$$

Itô's formula is extremely useful. Note that, when written out in integral form, it gives the following expression:

normal integral — ITO integral

$$f(S(t)) - f(S(0)) = \int_0^T \left(\mu\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial S^2}\right)dt + \int_0^T \sigma\frac{\partial f}{\partial S}dW.$$

Thus we get a relationship between a stochastic integral and a standard integral (with respect to time).

Remark. *We haven't given a proper proof of Itô's formula, but only a rough justification. In order to turn this into a proof, we need to show that our approximations converge in the correct sense, as in our rigorous proof that $\int_0^T W dW = \frac{1}{2}W(T)^2 - \frac{1}{2}T^2$ in the previous chapter.)*

We will skip the technical details for the moment (as well as the precise restrictions on f, and on the functions σ and μ).

We want ITO integral to be $\int W\,dW$ i.e. work backwards

15.2 Our first example, again

As an example, let us redo our calculation for $\int_0^T W(t)dW(t)$ using Itô's Lemma.

Define

$$S(t) = W(t),$$

which gives

$$dS = 0 \cdot dt + 1 \cdot dW,$$

so $\mu = 0$ and $\sigma = 1$.

We want to choose $f(S)$ so that the stochastic integral $\int_0^T \sigma \frac{\partial f}{\partial S} dW$ will be equal to $\int_0^T W dW$. Choosing $f(S) = \frac{1}{2}S^2$ gives

$$\frac{\partial f}{\partial S} = S(t) = W(t)$$
$$\frac{\partial^2 f}{\partial S^2} = 1$$

differentiated W
- skipping over details
We are doing this backwards
so integrating in a sense.

Itô's formula then gives

$$df = \frac{1}{2}dt + SdW.$$

Writing this in integral form, we get

$$f(S(t)) - f(S(0)) = \int_0^T \frac{1}{2}dt + \int_0^T W dW.$$

Now $f(S(t)) = \frac{1}{2}S(t)^2 = \frac{1}{2}W(t)^2$, and $W(0) = 0$, so this is equivalent to

$$\frac{1}{2}W(t)^2 = \frac{T}{2} + \int_0^T W dW,$$

or

$$\int_0^T W dW = \frac{1}{2}W(T)^2 - \frac{1}{2}T.$$

This is a bit simpler than our previous calculation!

more examples
in HW

15.3 A model for stock prices

Consider an asset with price $S(t)$ that evolves according to the stochastic differential equation

$$dS = \mu S dt + \sigma S dW.$$

This is a reasonable model for the behaviour of an asset's price over time: over a period Δt, the price changes by a deterministic quantity $\mu S \Delta t$ (representing some underlying deterministic growth) and a random quantity $\sigma S \Delta W$ (where σ measures the volatility of the asset).

It is useful to work in terms of $\log S(t)$, and so we define

$$f(S) = \log S,$$

and so clearly

$$\begin{aligned} f'(S) &= \frac{1}{S} \\ f''(S) &= -\frac{1}{S^2}. \end{aligned}$$

Itô's formula then gives

$$\begin{aligned} df &= f'(\mu S dt + \sigma S dW) + \frac{1}{2} f'' \sigma^2 S^2 dt \\ &= (\mu - \frac{1}{2}\sigma^2) dt + \sigma dW. \end{aligned}$$

In other words,

$$f(T) - f(0) = \int_0^T (\mu - \frac{1}{2}\sigma^2) dt + \int_0^T \sigma dW,$$

or

$$\log S(T) - \log S(0) = (\mu - \frac{1}{2}\sigma^2) T + \sigma W(T).$$

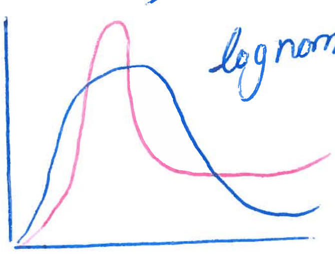
We conclude that

$$\log S(T) \sim \log S(0) + N\big((\mu - \frac{1}{2}\sigma^2) T, \sigma^2 T\big).$$

In other words, $\log(S(T)/S(0))$ is normally distributed with mean $(\mu - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$.

Note that this agrees with our model in Chapter 9.

stock price



lognormal dist.

price follows lognormal dist.

return is normally distributed

$f(S) = \log S$

15.4 Time-dependence

We get a slightly more complicated version of Itô's formula if we allow f to depend on t as well as $S(t)$.

Theorem 2 (Itô's Lemma: time-dependent version). *Let $f(S,t)$ have continuous second partial derivatives, and suppose that*

$$dS = \mu dt + \sigma dW.$$

Then

$$df = \left(\frac{\partial f}{\partial S}\mu + \frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial S^2} \right) dt + \frac{\partial f}{\partial S}\sigma dW.$$

no f', f'' because need to make diff'n explicitly clear i.e. w.r.t. S or t.

Once again, let us sketch a rough argument, following the same lines as in the time-independent case.

We have

$$f(S + \Delta S, t + \Delta t) = \frac{\partial f}{\partial S}\Delta S + \frac{\partial f}{\partial t}\Delta t + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}(\Delta S)^2 + \text{higher order terms}.$$

Now

$$\begin{aligned} (\Delta S)^2 &= (\mu \Delta t + \sigma \Delta W)^2 \\ &= \sigma^2 (\Delta W)^2 + \text{higher order terms}, \end{aligned}$$

as before. Replacing $(\Delta W)^2$ by Δt, and throwing away higher order terms, we get

$$\begin{aligned} \Delta f &= \frac{\partial f}{\partial S}(\mu \Delta t + \sigma \Delta W) + \frac{\partial f}{\partial t}\Delta t + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial S^2}\Delta t \\ &= \left(\mu \frac{\partial f}{\partial S} + \frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial S^2} \right)\Delta t + \frac{\partial f}{\partial S}\sigma \Delta W. \end{aligned}$$

16 Black-Scholes

16.1 Introduction

The most useful (and certainly the most famous) tool in evaluating derivative prices is the Black-Scholes formula. The formula gives a differential equation for the change in value over time of a derivatives contract written on an asset.

We assume that the asset price S follows the stochastic differential equation

$$dS = \mu S dt + \sigma S dW,$$

and that the riskless interest rate is r. We write $V(S,t)$ for the value of the contract at time t if the asset price is then S. [S and t are independent variables here. Of course, when we observe the asset over time, we see a sample path $(S(t))_{0 \le t \le T}$.]

The *Black-Scholes equation* for $V(S,t)$ is

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = rV.$$

Note that this doesn't yet give the value of a contract at any point: it just describes the change in price over time.

In order to solve the Black-Scholes equation in a specific instance, we need to specify *boundary conditions*, which depend on the value of the contract at expiry (and so on the type of contract we are considering). Note that we are here considering *European* contracts, i.e. contracts that have a specified expiry time.

We shall derive the Black-Scholes equation twice, presenting the proof in two different ways. In order to simplify the notation, we write

$$a(t) = \mu S(t), \qquad b(t) = \sigma S(t),$$

so that

$$dS = adt + bdW.$$

BS SDE $dS = \mu S dt + \sigma S dW$

PDE $\frac{dV}{dt} + rS\frac{dV}{dS} + \frac{1}{2}\sigma^2 \frac{d^2V}{dS^2}$ will derive this

will then solve to get this

formula $V = S\Phi(d_1) + e^{rt} K \phi(d_2)$

16.2 A self-financing strategy

know what's happening on next time step

Suppose that we can replicate the contract by a self-financing, previsible trading strategy, where at time t we have $H_1(t)$ units of the asset and $H_0(t)$ units of currency. [We could do this with the discrete-time model; let us assume that the same holds in continuous time.]

At time t, we have

$$V(t) = H_0(t) + H_1(t)S(t).$$

After time Δt, we get

$$V(t+\Delta t) = (1 + r\Delta t)H_0(t) + H_1(t)S(t+\Delta t),$$

so subtracting the first equation gives

$$\Delta V = rH_0\Delta t + H_1\Delta S.$$

So for replication, we need

$a = \mu S$
$b = \sigma S$

$$\begin{aligned}\Delta V &= rH_0\Delta t + H_1\Delta S\\ &= rH_0\Delta t + H_1(a\Delta t + b\Delta W).\end{aligned}$$

Rearranging, and letting $\Delta t \to 0$, we want

$$dV = (rH_0 + aH_1)dt + bH_1 dW.$$

V function whose value dependent on some underlying asset.

Now Itô's formula gives

$$dV = \left(\frac{\partial V}{\partial t} + a\frac{\partial V}{\partial S} + \frac{1}{2}b^2\frac{\partial^2 V}{\partial S^2}\right)dt + b\frac{\partial V}{\partial S}dW.$$

So we want

$$(rH_0 + aH_1)dt + bH_1 dW = \left(\frac{\partial V}{\partial t} + a\frac{\partial V}{\partial S} + \frac{1}{2}b^2\frac{\partial^2 V}{\partial S^2}\right)dt + b\frac{\partial V}{\partial S}dW.$$

It follows that we want

$$H_1 = \frac{\partial V}{\partial S},$$

which gives

$$\left(rH_0 + a\frac{\partial V}{\partial S}\right)dt = \left(\frac{\partial V}{\partial t} + a\frac{\partial V}{\partial S} + \frac{1}{2}b^2\frac{\partial^2 V}{\partial S^2}\right)dt$$

Want W bit to disappear

Brownian motion

and so

$$rH_0 = \frac{\partial V}{\partial t} + \frac{1}{2}b^2\frac{\partial^2 V}{\partial S^2}.$$

But $V = H_0 + H_1 S$, so this gives

$= H_1$

$$r\left(V - S\frac{\partial V}{\partial S}\right) = \frac{\partial V}{\partial t} + \frac{1}{2}b^2\frac{\partial^2 V}{\partial S^2}.$$

Rearranging gives

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}b^2\frac{\partial^2 V}{\partial S^2} = rV,$$

as required.

BLACK SCHOLES

Took replicating portfolio

Choose so that W cancels

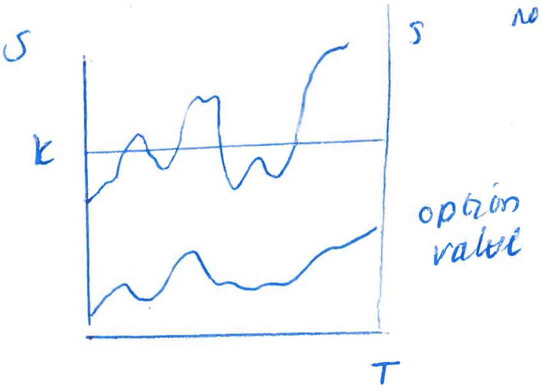
Using Ito's lemma

Note: no 'a' in BS PDE

so don't need to know what μ is

i.e. not dep. on future growth of asset

no linear relationship but some relationship



'prefers' this over previous
i.e. learn it

true for general options

16.3 Delta-neutral hedging

To present this derivation in a slightly different way, let us consider a portfolio with $h_1(t)$ units of the asset, combined with a short position of 1 unit of the derivatives contract.

At time t, the value $U(t)$ of the portfolio is

$$U(t) = h_1(t)S(t) - V(t).$$

After time Δt, the portfolio is worth

$$U(t) + \Delta U(t) = h_1(t)(S(t) + \Delta S(t)) - (V(t) + \Delta V(t)),$$

and so

$$\Delta U = h_1 \Delta S - \Delta V.$$

Now by Itô's formula, we have

$$dU = h_1(adt + bdW) - \left(\frac{\partial V}{\partial t} + a\frac{\partial V}{\partial S} + \frac{1}{2}b^2\frac{\partial^2 V}{\partial S^2}\right)dt - b\frac{\partial V}{\partial S}dW.$$

because Δ it is random

The right-hand side will be *riskless* (i.e. independent of the random change dW), if we choose

$$h_1 = \frac{\partial V}{\partial S}.$$

In this case, the portfolio must grow at the riskless rate, so $\Delta U = (r\Delta t)U$, or

$$dU = Urdt.$$

Substituting for h_1 and ΔU, we get

$$Urdt = (adt + bdW)\frac{\partial V}{\partial S} - \left(\frac{\partial V}{\partial t} + a\frac{\partial V}{\partial S} + \frac{1}{2}b^2\frac{\partial^2 V}{\partial S^2}\right)dt - b\frac{\partial V}{\partial S}dW$$

and so (after some cancellation)

$$Ur = -\frac{\partial V}{\partial t} - \frac{1}{2}b^2\frac{\partial^2 V}{\partial S^2}.$$

Since

$$U = h_1 S - V = S\frac{\partial V}{\partial S} - V,$$

this gives the Black-Scholes formula

$$\boxed{\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}b^2\frac{\partial^2 V}{\partial S^2} = rV.}$$

Δ = delta = $\dfrac{\partial V}{\partial S}$ change in value of derivative as asset price changes ie. the heat eq^n can be solved by an expectation

16.4 Feynman-Kac

Consider a process satisfying the equation

$$dS = \mu S dt + \sigma S dW.$$

If we assume that $S(t_0) = x$, then what can we say about $V(x, t_0)$? If the payoff at time T is $v(S(T), T)$ then the Feynman-Kac formula says that

$$V(x, t_0) = \mathbb{E}\big[v(S(T), T)e^{-r(T-t_0)}\big].$$

could be small v

More generally, if we have instantaneous payoffs at rate $u(S(t), t)$, we have

$$V(x, t_0) = \mathbb{E}\big[v(S(T), T)e^{-r(T-t_0)} + \int_{t_0}^{T} e^{-r(t-t_0)} u(S(t), t) dt\big].$$

Why should this be true? Feynman and Kac noticed that the equation satisfies the partial differential equation

$$\frac{\partial V}{\partial t} + \mu x \frac{\partial V}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V}{\partial x^2} = rV + u(x, t),$$

with boundary conditions

$$V(x, T) = v(x, t).$$

(The equation is then known as the *Kolmogorov* or *Dynkin equation.*)

If we take $\mu = r$, $\underline{x = S(t_0)}$ and $u = 0$, we see that V satisfies the Black-Scholes equation for a European contingent claim paying $v(S(T), T)$ at termination:

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = rV.$$

Next, we give an important application.

Showed $\log S(T)$ is normally distributed already

16.5 European claims

As a corollary of the Feynman-Kac formula, we have a formula for valuing European contingent claims.

For instance, consider a European call option paying $(S(T)-K)^+$ at time T. If the price of the asset at time t_0 is $S(t_0)$, then the time t_0 value of the call option is

$$\mathbb{E}\big[e^{-r(T-t_0)}(S(T)-K)^+\big].$$

In order to evaluate this, we need to know the distribution of $\log S(T)$. But we know from the previous chapter that $\log S(T)$ is normal with mean

$$m = \log S(t_0) + (r - \frac{1}{2}\sigma^2)(T - t_0)$$

and variance

$$s^2 = \sigma^2(T - t_0).$$

Let us therefore write

mean

z score $\quad z(T) = \dfrac{\log S(T) - m}{s}$ $\quad$ variance

and

$$w = \frac{\log K - m}{s}.$$

Since $z(T)$ is normal with mean 0 and variance 1, we know that it has density

$$\phi(t) = \frac{1}{\sqrt{2\pi}}e^{-t^2/2}.$$

So

$$\begin{aligned}\mathbb{E}\big[e^{-r(T-t_0)}(S(T)-K)^+\big] &= e^{-r(T-t_0)}\mathbb{E}\big[(e^{sz(T)+m}-K)^+\big]\\ &= e^{-r(T-t_0)}\int_w^\infty (e^{sx+m}-K)\phi(x)dx.\end{aligned}$$

Now

$$\begin{aligned}\int_w^\infty K\phi(x)dx &= K(1-\Phi(w))\\ &= K\Phi(-w).\end{aligned}$$

integrate normal - get cumulative normal

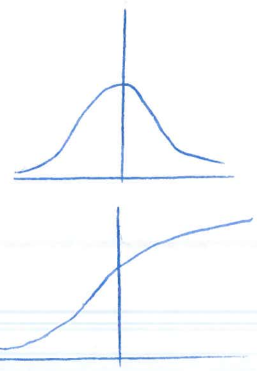
because of odd function

only interested in terms where this is strictly +ve

$e^{sz(T)+m} - K \geq 0$

$sz(T) + m \geq \log K$

$z(T) \geq \frac{\log K - m}{s} = w$



density function $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$

Also

$$\begin{aligned}
\int_w^\infty e^{sx+m}\phi(x)dx &= \frac{1}{\sqrt{2\pi}}\int_w^\infty e^{sx+m-\frac{1}{2}x^2}dx \\
&= \frac{1}{\sqrt{2\pi}}\int_w^\infty e^{-(1/2)(x-s)^2+s^2/2+m}dx \\
&= e^{s^2/2+m}\frac{1}{\sqrt{2\pi}}\int_w^\infty e^{-(1/2)(x-s)^2}dx \\
&= e^{s^2/2+m}\frac{1}{\sqrt{2\pi}}\int_{w-s}^\infty e^{-(1/2)x^2}dx \\
&= e^{s^2/2+m}(1-\Phi(w-s)) \\
&= e^{s^2/2+m}\Phi(s-w)).
\end{aligned}$$

diff. of two squares

substitution

Since

$$\begin{aligned}
s^2/2+m &= \frac{\sigma^2}{2}(T-t_0)+\log S(t_0)+(r-\frac{1}{2}\sigma^2)(T-t_0) \\
&= \log S(t_0)+r(T-t_0),
\end{aligned}$$

we get

$$\begin{aligned}
V(S(t_0),t_0) &= e^{-r(T-t_0)}\left[e^{s^2/2+m}\Phi(s-w)-K\Phi(-w)\right] \\
&= S(t_0)\Phi(s-w)-e^{-r(T-t_0)}K\Phi(-w).
\end{aligned}$$

Let us set

$$\begin{aligned}
d_2 &= -w \\
&= \frac{m-\log K}{s} \\
&= \frac{\log(S(t_0)/K)+(r-\frac{1}{2}\sigma^2)(T-t_0)}{\sigma\sqrt{T-t_0}}
\end{aligned}$$

and

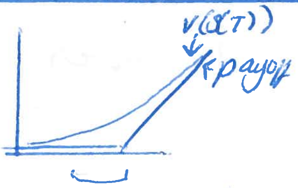
$$\begin{aligned}
d_1 &= s-w \\
&= \frac{m-\log K+s^2}{s} \\
&= \frac{\log(S(t_0)/K)+(r+\frac{1}{2}\sigma^2)(T-t_0)}{\sigma\sqrt{T-t_0}}.
\end{aligned}$$

We then obtain the following equation for the price at time t_0 of a European option with strike price K and expiry time T:

$$\boxed{V(S(t_0), t_0) = S(t_0)\Phi(d_1) - e^{-r(T-t_0)}K\Phi(d_2).}$$

Δ discussion

Call option

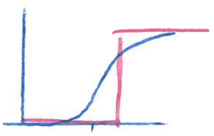


as $t \to T$ get convergence

$\Delta = \frac{\partial V}{\partial S}$

value of V changing most here

Δ

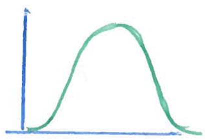


here we are showing how Δ changes

don't have to hedge every Δ on every option

asked exam q on this

$\Gamma = \frac{\partial^2 V}{\partial S^2}$ how does Δ change as asset price changes



$\Theta = \frac{\partial V}{\partial t}$

$\Theta + rS\Delta + \frac{\sigma^2}{2}\Gamma \leftarrow$ BS PDE

Can read about this in book

The exam

He can ask us anything - so maybe read the book?

1) Arbitrage covered interest parity, put-call parity, no-arbitrage theorem - this thm is asked almost every year
2) Binomial model - some tree, value something, binomial derivation
3) Americans / exotic options, dynamic programming type thing
4) Ito, stochastic calculus, Brownian motion
5) Black Scholes, PDE formula derivation, what assumptions have we made - not in the notes

Constant risk free

No arbitrage

Frictionless

Cont. hedging

Log normal dist

No dividends

Some years there is AT MOST ONE an essay style question
- this reduces amount of marking
- this is also the easiest question for someone who has understood course

Will involve deriving covered interest rate
put call parity
no arbitrage

For an essay q there is normally around 35 marks available

• There will be something in the paper you have not seen before

ask questions → walton@math.ucl.ac.uk ← do not believe maths website