Financial Derivatives

Giulia Iori

Contents

Overview of Financial Markets

Functions of Financial Markets:

Financial markets determine the prices of assets, provide a place for exchanging assets and lower costs of transacting. This aids the resource allocation process for the whole economy.

- price discovery process
- provide liquidity
- reduce search costs
- reduce information costs

Market Efficiency:

- Operational efficiency: fees charged by professional reflect true cost of providing those services.
- price efficiency: prices reflect the true values of assets.
  - Weak efficiency: current price reflect information embodied in past price movements.
  - Semistrong efficiency: current price reflect information embodied in past price movements and public information.
  - Strong efficiency: current price reflect information embodied in past price movements and all public and private information.

Brief history:

Birth of shareholding enterprise, Muscovy Company (1553), East India Company (1600), Hudson's Bay Company (1668). Trading starts on the shares of these company. Amsterdam stock exchange (1611), Austrian Bourse in Vienna (1771). In London coffee houses where brokers meet. New Jonathen's Coffee house becomes the Stock Exchange in 1773.

In New York dealers in stock met at 22 Wall Street (so called because of the building of a wall to keep livestock in and Indians out by early Dutch traders who founded NY). New York Stock Exchange and Board set up in 1817.

In France, the Societe des Moulin Du Balzac in Toulouse, which become the local electricity company, was the first shareholding company, and was quoted in Toulouse Stock exchange. First Bourse in Lyon in 1540. Bourse in Paris established in 1724, closed in 1791 during revolution and opened again under Napoleon in 1801.

Traded Financial Assets:

- Tangible: have physical existence.
- Intangible: legal claim to future cash flow, debt, equity, preferred stocks, convertible bonds.

Intermediaries:

- brokers: purchases on the behalf of a client.
- dealers or market makers: stand ready and willing to buy and sell on their own account, quote a bid and ask price. Keep an account. Provide immediacy. Bear risks: uncertainty about future prices of a stock, uncertainty about the time they must hold a position in a stock, possibility to trade with informed traders. Bid-ask spread reflects dealers processing costs and perceived risks.

Classification of Financial Markets:

By type of claim:

- Fixed Income: debt and preferred stocks
- Equity (or Stock) Market: equity
- Derivatives Markets: where futures, options and swaps are traded.

By maturity of claim:

- Money Market: short term debt, one year or less.
- Capital Market: long maturity assets, both debt and equity.

By geographical location of issuer:

- Domestic: where assets issued by domestic issuer are traded (example stock of American corporation traded in US).
- Foreign: where assets issued by foreign issuer are traded (example stock of Japanese corporation traded in London). Securities regulated by authorities in the country where they are issued.
- External: (also International Market or Euromarket) includes securities that are offered simultaneously to investors in a number of countries and are issued outside the jurisdiction of any single country.
- Emerging Markets: rapidly growing, markets of recently industrialized countries.

By type of market:

- Exchanges: physical places where only members can trade, highly regulated, only listed assets are traded who need to meet standards and pay fees.
- Over the counter markets: listed and not listed assets traded, transaction over the phone. Most bond trade here.
- Electronic Markets: growing thenks to market deregulation and technological advances.

Elements of trade:

- Security to trade
- quantity to be trade: round lot, odd lot, block trade.
- buy or sell: short sell and buy on margins. restriction on short selling: up-tick and zero-tick rules.
- Arrangements for payment and delivery: clearing house. Netting. Counterpart risk. Delivery agreements.
- type of orders: market, limit, stop.
  - market orders

    A market order is an order to buy or sell at the market price. A trader who submits a market order faces the risk that the order is executed at a price far from the desired one.
  - limit orders

    To guarantee that an order is executed only when the market price is below or above a certain threshold a trader can submit a limit order. By delaying transacting traders hope to trade at a more favourable price. Nonetheless, limit order traders face uncertainty over when the trade will execute and the estimated value of the asset may have changed since the order was placed. Traders, then, risk to be left with a negative expected profit, or, if the order is not matched during the trading window (given by the order lifetime), not to trade at all.
  - stop orders

    Stop orders are market orders which are triggered by the market price reaching a predetermined threshold. Even if they provide protection against unexpected losses or rises they do not guarantee that the order will be executed at a price close to the specified threshold. Stop orders can be submitted, for example to set up stop-loss hedging strategies.
- Trading Mechanisms:

- Quote-driven markets: competing market makers supply liquidity by quoting bid and ask prices and the number of shares at which they are willing to trade. Investors demand liquidity through the submission of market orders.
- Order-driven markets: investors can, but are not obliged to, submit limit orders. Orders are stored in the exchange's book and executed in the sequence they arrive to the market. A transaction occurs when a trader hits the quote on the opposite side of the market. Transactions are executed using time priority at a given price and price priority across prices. Limit orders provide liquidity to market orders.

Major Financial Markets in the world:

- USA:
  - NYSE (New York Stock Exchange) located in lower Manhattan, NYC. Primarily a stock exchange but also bonds, futures and options. Major exchange in the world. Uses DOT (1976) and (Designated Order Turnaround) as an automatic electronic order execution service. Super DOT (1984).
  - AMEX (American Stock Exchange) also stocks, bonds and options.
  - regional exchanges (Boston, Chicago, Cincinnati, Philadelphia, Midwest, Pacific). Stocks and options
  - CBOT (Chicago Board of Trade, 1848) future market.
  - CME (Chicago Mercantile Exchange, 1874, future markets.
  - IMM (International Monetary Market, 1972) futures in foreign currency.
  - CBOE (Chicago Board Options Exchange, 1973): main option markets.
  - NASDAQ (National Associations of Securities Dealers Automated Quotations). Over the counter. Listed and non listed stocks and bonds. Hybrid system.
  - NYFE (New York Future exchange)
  - NYMEX (New York Mercantile Exchange).
  - COMEX (Commodity exchange).
- UK:
  - LSE (London Stock Exchange): Uses SETS (Stock Exchange Electronic Trading Systems) for top 100 stocks and SEAQ (Stock exchange Automated Quotation) hybrid system similar to NASDAQ, for other shares. Also SEATS (Stock exchange Alternative Trading System)
  - LIFFE (London International Financial Futures Exchange)
  - LME (London Metal Exchange)

  - OMLX (London Security and Derivatives Exchange)
  - LCE (London Commodity Exchange)
  - IPE (International Petroleum exchange)

- Germany: Frankfurt.
- Canada: Toronto Stock Exchange, uses CATS(Computer Assisted Trading System). Also derivatives.
- France: Paris Bourse. Uses CAC (Cotation Assistee en Continu) and NSC (Nouveau Systeme de Cotation or supercac. MATIF and MONEP for futures and options.
- Japan: Tokio, Osaka and 6 others. TIFFE future exchange.
- Switzerland SOFFEX (Swiss Options and Financial Futures Exchange)
- Sydney SFE (Sydney Futures Exchange)
- Singapore SIMES (Singapore International Monetary Exchange)
- Hong Kong: HKFE Hong Kong Future exchange.

Stock Market Indexes:

Three factors differentiate stock market indexes:

- Universe of stocks represented (composition may change over time).
- Weights assigned to stocks.
- Method of averaging.

<u>Weights</u>

- by market value of company or market capitalization: Capitalization = Number of shares x Price of a share.

$$w_i = \frac{C_i}{\sum_{j=1}^{N} C_j}$$

- By the price of one unit of the company's stock.

$$w_i = \frac{p_i}{\sum_{j=1}^{N} p_j}$$

- Equal weight.

$$w_i = \frac{1}{N}$$

Method of averaging

- Arithmetic Average (all market index are constructed using arithmetic average)

$$I = C\sum_{i=1}^{N} w_i p_i$$

- Geometric Average

$$I = C(\Pi_{i=1}^{N} w_i p_i)^{1/N}$$

Exercise: Construct an index using the three different weighting scheme using the following data

APX	$P = 28.625\ US\$$	$N = 485.445\ million$
GE	$P = 105.250\ US\$$	$N = 852.935\ million$
$3M$	$P = 103.250\ US\$$	$N = 215.791\ million$
$Merck$	$P = 32.125\ US\$$	$N = 1282.316\ million$
$Exxon$	$P = 65.250\ US\$$	$N = 1241.618\ million$

Classification by producer:

- those produced by trading systems based on stocks traded on that system (example: NYSE composite index, American Stock Exchange Market Value Index, NASDAQ Composite Index, TOPIX).
- those produced by organizations that subjectively select the stock to be included in the index (example: $S\&P500$, DJIA or Dow Jones Industrial Average, Nikkei 225 Stock Average, FT100...)
- those produced by organizations that objectively select the stock to be included in the index.

Major Indexes:

- US
  - DJIA (1896 with 12 shares, 1928 with 30 hares): largest 30 blue chips traded on NYSE. Prices are averaged with equal weight. Calculated every minute. Accounts for splits and dividend.

- $S\&P500$ (1957): selected samples from of stocks traded on NYSE and AMEX and OTC market. Based on market capitalization.
- $S\&P100$ (1983): selected samples from of stocks traded on CBOE. Capitalization Index.
- NYSE Index (1966), consists of 1500 stocks. Based on capitalization.
- AMEX Index (1973) based on about 800 stocks. Capitalization Index.
- NASDAQ Composite Index (1984)
- NASDAQ Industrial Index (1984)
- NASDAQ 100 (1985).

• UK

- FT 30 (1935), geometric average index.
- FTSE 100 (1984). Top 100 shares. Capitalization index. Calculated every 15 seconds.
- FTSE 250 (1992), 250 shares after the FTSE 100.
- FTSE Actuaries 350.
- Financial Times Actuaries Index started in 1962 and widened to include 800 stocks in 1992.

• France

- CAC 40 (1987). Represent all major sectors. 40 large company traded on the exchange. Based on capitalization and measured every 30 seconds.
- SBF 240: older one. Calculated once per day using opening prices.
- SBF 250 (1993) replaced SBF 240.
- SBF 120 (1993): 40 shares in CAC 40 plus 80 more. Calculated every minute.
- MIDCAC (1995): index of middle capitalization stock.

• Germany

- FAZ 100 (1950) calculated once per day. 100 largest company company listed on the Frankfurt exchange.
- DAX (1987) 30 shares most actively traded on the Frankfurt exchange, Calculated continuously, include dividends.

• Italy

- Comit all Shares
- BCI
- MIB 30

• Spain

  - Madrid SE
  - IBEX 35

- Japan
  - Nikkei Dow 225, based on average prices
  - Nikkei 300 (1984) based on capitalization
  - Topix, based on capitalization. All shares listed on the first section of TSE.

- Hong Kong: Hang Seng Index

- Toronto: TSE 300

- International Equity Indexes: Morgan Stanley Word Index, FTSE Eurotrack 100, FTSE Eurotrack 200, Financial Times world Index, Salomon Brothers-Russel Global Equity Index, Global Index.

Introduction to Financial Derivatives

Derivatives can be seen as bets based on the behaviour of the underlying basic assets. A derivative can also be regarded as a kind of asset, the ownership of which entitles the holder to receive from the seller a cash payment or possibly a series of cash payments at some point in the future, depending in some pre-specified way on the behaviour of the underlying assets over the relevant time interval. In some instances, instead of a 'cash' payment another asset might be delivered instead. For example, a basic stock option allows the holder to purchase shares at some point in the future for a pre-specified price. In general an option is a derivative with a specified payoff function that can depend on the prices of one or more underlying assets. It will have specific dates when it can be exercised, that is, when the owner of the option can demand payment, based on the value of the payoff function.

Derivatives are used for a variety of purposes. They can be used to reduce risk by allowing the investor to hedge an investment or exposure, and hence function as a sort of insurance policy against adverse market movements. For example, if a firm needs a particular commodity, such as petroleum, on a regular basis, then they can guard against a rise in the price of oil by purchasing a call option. If the price of oil remains low, then the option is not exercised and the oil is bought at the current price in the market, while if the price rises above the strike, then the option is exercised to buy oil at a below-market value.

Derivatives can also be used to gain extra leverage for specialized market speculation. In other words, if an investor has reason to believe that the market is going to move in a particular way, then a larger profit per dollar invested can be made by buying suitable derivatives, rather than the underlying asset. But similarly, if the investment decision is wrong, the investor runs the risk of making a correspondingly larger loss.

In London, organised derivatives trading takes place at the London International Financial Futures and Options Exchange (LIFFE).

A derivative contract is defined by

- its payoff function $f(t, S(t))$
- its purchase price f_0.

The most common types of derivatives are forwards, futures, swaps and options. This course focuses mainly on options.

Most options can only be exercised once, and have a fixed expiration date, after which the option is no longer valid. There are many different schemes for prescribing when an option can be exercised. The most common examples are the so-called European options, which can only be exercised on the expiration date, and American options, which can be exercised at any time up to the expiration date. However you are never forced to 'exercise the option'.

The two most common options are the call option, which gives the owner the right to buy a designated underlying asset at a set price (called the strike price K), and the put option which allows the owner to sell the underlying asset at a given strike price. If we consider European options, then the payoff function depends only on the value of the underlying assets S_t at the expiry date, $t = T$. This can be expressed mathematically as

$$C_T = [S_T - K]^+$$

where the compact notation $[x]^+ = max[x, 0]$ has been used.
The payoff of a long forward contract is instead

$$F_T = [S_T - K]$$

In the case of a put option, the payoff is only non-zero if the asset price at expiration is less than the strike price. This is given by

$$P_T = [K - S_T]^+.$$

Now we need to determine the price f_0 that someone should pay at time zero to buy a derivative that pays off $f(t, S_t)$ dollars at time t. A plausible guess is

$$f_0 = e^{-rT} E[f(T, S_T)]$$

which represents the discounted expected payoff of the derivative, that is, the probability weighted average of the possible payoff.

This guess is a typical ex- ample of the 'expectation hypothesis'. But we will see that this is wrong.

Time value of money:

Note that in general we assume that the value now of a dollar promised at time T is given by e^{-rT} where r is the continuously compounded interest rate (that we have here assumed constant could be itself random).

Interest rates and compounding frequencies

Suppose we invest an amount I for one year and compound interest periodically (and reinvest interest) m times during the year at a rate

$$r_m/m$$

where r_m is the annual rate compound m times. The terminal value of this investment is

$$I(1 + r_m/m)^m.$$

What is the equivalent annual rate r_1 compounded only once? We need to impose

$$I(1 + r_m/m)^m = I(1 + r_1)$$

which gives

$$r_1 = (1 + r_m/m)^m - 1.$$

In general if we want to compare compounding m_1 versus m_2 times during the year we have to impose

$$I(1 + r_{m_1}/m_1)^{m_1} = I(1 + r_{m_2}/m_2)^{m_2}.$$

The annual rate compounding m_1 times equivalent to a annual rate compounding m_2 times is given by

$$r_{m_1} = ((1 + r_{m_2}/m_2)^{m_2/m_1} - 1)m_1.$$

Note that for $m_1 = 1$ this reduces to previous case.

Consider now the case of an investment which pays a rate r, compounded m times during the year, for n years. The terminal value is

$$I(1 + r_m/m)^{nm}.$$

Remember that

$$\lim_{m\to\infty}(1 + x/m)^m = e^x.$$

hence

$$\lim_{m\to\infty}(1 + r_m/m)^{nm} = e^{r_c m}.$$

where r_c is called continuously compounding rate and correspond to r_m compounded $m = \infty$ times.

To convert from discrete compounding to continuous compounding we impose

$$e^{r_c n} = \left(1 + \frac{r_m}{m}\right)^{mn}$$

or

$$r_c = m\ln(1 + \frac{r_m}{m})$$

and

$$r_m = m(e^{r_c/m} - 1)$$

Hedging

So far we have talked about investors that buy derivatives, but there must likewise be a financial institutions selling them. These sellers are generally investment banks, stock exchanges, and other large institutions. When selling a derivative, the issuer makes an initial gain up-front from the fee that they charge. The issuer who has sold a derivative can do better than just wait and see, he can hedge the contract.

They issuer will use the up-front money, possibly in conjunction with borrowing, to hedge the derivative that they have sold by buying other instruments in the market to form a hedging portfolio, in such a manner that, regardless of the way that the prices of the underlying assets change, they neither gain nor lose money.

The ability of the issuer to hedge the contract should be reflected in the price.

Abitrage pricing

Arbitrage is the key to understanding the mathematics of derivative pricing.

No arbitrage means that it is not possible to construct a strategy that on average makes a profit higher than the risk free rate without taking some risks.

This also imples that is not possible to construct a strategy that requires no cash input but has some positive probability of making profits without any risk of a loss.

No arbitrage also implies that given two strategies with the same initial position, and guaranteed final positions, then these final positions must be equal. Otherwise, by going long the strategy with the higher final value and short the other we would generate an arbitrage. Similarly if two strategies have the same final value and involve the same risk they need to have the same initial value.

Any no arbitrage argument for pricing a derivative is ultimately based on a replication strategy, which is a trading strategy that uses market instruments to 'replicate' the initial and final positions required by the derivative.

No arbitrage implies the value of the derivative contract must be equal to the cost of setting up the hedging portfolio.

The principle of no arbitrage may be the key to understanding derivative pricing, but what kind of law is it? It is clearly not a fundamental law of nature, and is not even always obeyed by the markets. In some ways it is similar to Darwin's theory of natural selection. An institution that does not price by arbitrage arguments the derivatives that it sells will suffer relative to institutions that do. If the price is set too high, then competitors will undercut it; if the price is too low, then the institution will be liable to market uncertainty as a hedging portfolio cannot be properly constructed.

But how is the value of the initial payment to be calculated? What is the composition of the hedging portfolio? These are the two big questions in derivatives pricing.

The Binomial Model

The simplest assumption we can make about prices is that they are descrete and follow a random walk.

A random walk is a formalization of the intuitive idea of taking successive steps, each in a random direction. A random walk is a simple stochastic process sometimes called a "drunkard's walk".

In general we assume the time horizon is T and the set of dates in our market is $t = 0, 1, \cdots, T$.

Also we assume the market consist of

- the money market account:

$$B_0 = 1, \quad B_t = (1+r)^t$$

  were we assumed a constant interest rate r is paid on the money market account. r is the risk free rate (compounded over the same period).

- a stock

$$S_{t+1} = uS_t \quad \textit{with probability} \quad p$$
$$S_{t+1} = dS_t \quad \textit{with probability} \quad 1-p$$

  with $0 < d < u$ (on average stock value increases).

- a derivative contract specified the maturity date T, and the payoff function $f(t, S_t)$.

Assumptions:

- No market frictions, i.e. no transaction costs, no bid-ask spreads, no margin requirements, no restrictions on short sell, no taxes.
- No counterpart risk. Same lending and borrowing rates.
- Markets are liquids, i.e. traders can buy or sell any amount of a security without affecting it price.
- Markets are complete.
- Prices have adjusted so that there are no arbitrage opportunities.

One step model: $T = 1$

$$S_1 = uS_0 = S^u \quad \textit{with probability} \quad p$$
$$S_1 = dS_0 = S^d \quad \textit{with probability} \quad 1-p$$

We want to price the derivative by using a no arbitrage argument. The investor will buy a derivative from a trader, who will then take the proceeds of this sale and invest in the stock and money market so that the randomness in his stock and derivative positions cancel. Let x be the number (possibly fractional, possibly negative) of units in the money market account that the trader buys, and y the number of units in the stock S. How does the trader choose the values of x and y?

- Consider a portfolio composed of x unit of cash and y stock. Its initial value is

  $$\pi_0 = xB_0 + yS_0$$

  where $B_0 = 1$. The value of the portfolio at the end of the first period is

  $$\pi_1 = x(1+r) + yS_1$$

  or

  $$\pi_1 = x(1+r) + yuS_0 \quad \textit{with probability} \quad p$$
  $$\pi_1 = x(1+r) + ydS_0 \quad \textit{with probability} \quad 1-p$$

- Consider a second portfolio composed by one option contract f on the same stock.

  The value of the option at the end of the first period is

  $$f_1 = f_u \quad \textit{with probability} \quad p$$
  $$f_1 = f_d \quad \textit{with probability} \quad 1-p$$

  We want to choose x, y so that the value of the two portfolios coincide at the end of the period,

  $$x(1+r) + yuS_0 = f_u$$
  $$x(1+r) + ydS_0 = f_d$$

  Because the bond and stock portfolio exactly duplicates the payoff function of the derivative, we call it a replicating strategy. The existence of a replicating strategy means that the derivative can be constructed from the underlying assets and hence an investor need never buy the derivative. In this case we say that the market is **complete**. In the real world, however, a derivative has lower transaction and maintenance costs and that is why investors will purchase them.

$$x(1+r) + yuS_0 = f_u$$
$$x(1+r) + ydS_0 = f_d$$

Subtracting the two equations above we obtain

$$yS_0(u-d) = f_u - f_d$$

or

$$y = \frac{f_u - f_d}{S_0(u-d)}$$

and

$$x(1+r) = f_u - u\frac{f_u - f_d}{(u-d)}$$

Also, since the purchase of the portfolio is entirely funded by the money received from the sale of the derivative, we have $f_0 = x + yS_0$ or

$$f_0 = x + \frac{f_u - f_d}{S_0(u-d)}S_0.$$

Multiplying both sides by $(1+r)$

$$f_0(1+r) = x(1+r) + f_u\frac{1+r}{u-d} - f_d\frac{1+r}{u-d}$$

Replacing $x(1+r)$ from above we obtain

$$f_0(1+r) = f_u - u\frac{f_u - f_d}{(u-d)} + f_u\frac{1+r}{u-d} - f_d\frac{1+r}{u-d} =$$
$$= f_u\frac{1+r-d}{u-d} + f_d\frac{u-(1+r)}{u-d}$$

This formula has an interesting interpretation. Define

$$p^* = \frac{1+r-d}{u-d}$$

then

$$1 - p^* = 1 - \frac{1+r-d}{u-d} = \frac{u-(1+r)}{u-d}$$

Hence

$$f_0 = \frac{1}{1+r}[f_u p^* + f_d(1-p^*)]$$

p^* can be interpreted as a probability measure. The total probability is one. To be a probability we need to impose that each event has positive probability, i.e.

$$\frac{1+r-d}{u-d} > 0$$

$$\frac{u-(1+r)}{u-d} > 0$$

From where

$$d < 1 + r < u$$

p^* is called the **risk neutral measure**. This can be seen observing that under p^* the growth rate of the asset is r.

Proof:

$$E^*[S_1] = p^* S_0 u + (1-p^*) S_0 u =$$

$$\frac{1+r-d}{u-d} S_0 u + \frac{u-(1+r)}{u-d} S_0 d =$$

$$S_0 \frac{u + ru - du + ud - d - rd}{u-d} = S_0(1+r)$$

Hence

$$f_0 = \frac{1}{1+r} E^*[f_1]$$

The value of the derivatives can be expressed as the present value of the expected future payoff, where the expectation is calculated in a risk neutral world.

The price of a call European option is then

$$C_0 = \frac{1}{1+r} E^*[(S_T - K)^+] =$$

$$\frac{1}{1+r}[(uS_0 - K)^+ p^* + (dS_0 - K)^+ (1-p^*)]$$

The price of a put European option is then

$$P_0 = \frac{1}{1+r} E^*[(K - S_T)^+] =$$

$$\frac{1}{1+r}[(K - uS_0)^+ p^* + (K - dS_0)^+ (1-p^*)]$$

Multi-period binomial model: $T = N$

A very useful special model is obtained, by letting the branches recombine to form a lattice of prices, and hence is called a 'lattice model' or 'recombining tree'. At time t = n the number of different states is only n + 1 which grows much more slowly than the 2n nodes of the 'basic' tree.

Define

$$Z_t = \frac{S_{t+1}}{S_t}$$

Z_t take only two values u and d. Zs are independent and identically distributed Bernoulli random variables, and

$$S_t = S_0 \prod_{\tau=1}^{t} Z(\tau).$$

The price of a European Call is:

$$C_0 = \frac{1}{(1+r)^T} E^*[(S_T - K)^+] =$$

$$\frac{1}{(1+r)^T} E^*[(S_0 \prod_{i=1}^{T} Z_i - K)^+] =$$

$$\frac{1}{(1+r)^T} \sum_{j=0}^{T} \binom{T}{j} p^{*j} (1-p^*)^{T-j} [S_0 u^j d^{T-j} - K]^+$$

In practice the pricing is done by going backward on the tree.

Examples:

Continuous time pricing

We will model asset prices as Itô's process $x(t)$ satisfiyng

$$dx(t) = a(t)dt + b(t)dW(t)$$

$a(t)$ is called the drift and $b(t)$ is called the volatility. They can be stochastic. In the special case

$$dx(t) = a(t, x(t))dt + b(t, x(T))dW(t)$$

$x(t)$ is called a diffusion process and the equation above is called a stochastic differential equation or SDE.

Stochastic process

A stochastic process is a sequence of r.v. $X = (X_t(\omega), t \in T, \omega \in \Omega)$ defined on some probability space Ω. T can be a finite set, or countably infinite set (discrete-time process) or an interval (a, b) or (a, ∞) (continuous-time process). The index t is usually refereed to as time. A stochastic process $X_t(\omega)$ is a function of two variables:

- at fixed time t, $X_t(\omega), \omega \in \Omega$ it is a random variable
- for a given random outcome ω, $X_t(\omega), t \in T$ it is a function of time, called a realization or a trajectory, or sample path of the process X.

Wiener process or Brownian motion

A stochastic process $W = (W(t) : t \geq 0)$ is a standard Brownian motion on some probability space $(\Omega, P, \mathcal{F})$ if

- $W(0) = 0$
- W has independent increments: $W(t+u) - W(t)$ is independent of $(W(s) : s \leq t)$ for $u \geq 0$.
- W has stationary increments: the law of $W(t+u) - W(t)$ depends only on u.
- W has Gaussian increments: $W(t+u) - W(t) \sim N(0, u)$.

As the step size Δt in the random walk tends to 0 (and the number of steps increased comparatively) the random walk converges to Brownian motion in an appropriate sense.

Properties of Wiener processes:

- W has continuous sample paths.
- Sample path of Brownian motion are nowhere differentiable.
- Sample path are of unbounded variation. This implies that the process covers an infinite distance as it evolves over any time interval, no matter how short the interval is.
- Wiener processes are Markov (or memoryless):

  A process X is said to be Markov if for each t, each $A \in \sigma(X(s) : s > t)$ (the future) and $B \in \sigma(X(s) : s < t)$ (the past),

  $$P(A|X(t), B) = P(A, X(t)).$$

Moments of s.B.M

If $X \sim N(\mu, \sigma)$ its moment generating function is

$$M(t) = E[e^{tX}] = e^{\mu t + 1/2\sigma^2 t^2}.$$

We know that the n-moment is given by

$$m_X^n = \frac{dM(t)}{dt^n}|_{t=0}.$$

Given that $dW(t) = W(t+dt) - W(t) \sim N(0, dt)$ it is easy to show that

$$E[dW(t)] = E[W(t+dt) - W(t)] = E[W(dt)] = 0$$

$$E[dW^2(t)] = E[W^2(dt)] = var[W(dt)] = dt$$
$$E[dW^4(t)] = E[W^4(dt)] = 3dt^2$$
$$var[dW^2(t)] = 3dt^2 - dt^2 = 2dt^2$$

Note that $W(0) = 0$ and
$$W(t) = W(t) - W(0)$$
represents the increment from the initial position. Hence it also holds

$$E[W(t)] = 0$$

$$E[W^2(t)] = t$$
$$E[W^4(t)] = 3t^2$$
$$var[W^2(t)] = 3t^2 - t^2 = 2t^2$$

Itô's Lemma:

Consider an Itô's process $x(t)$ following

$$dx(t) = a(t)dt + b(t)dW(t)$$

Take a smooth function $F(t, x(t))$. It$\hat{o}$'s lemma says that

$$dF(t, x(t)) = (F_t + a(t)F_x + \frac{1}{2}b^2(t)F_{xx})dt + b(t)F_x dW(t)$$

where
$$F_t = \frac{\partial F(t,x)}{\partial t} \quad F_x = \frac{\partial F(t,x)}{\partial x} \quad F_{xx} = \frac{\partial^2 F(t,x)}{\partial x^2}$$

So F follows an Itô process with drift

$$F_t + aF_x + \frac{1}{2}b^2 F_{xx}$$

and volatility

$$bF_x$$

In integral form (which is the only thing which has a well defined meaning)

$$F(t,x(t)) = F(0,x(0)) + \int_0^t (F_t + a(s)F_x + \frac{1}{2}b^2(s)F_{xx})ds + \int_0^t b(s)F_x dW(t)$$

Non rigorous proof:

Expand $F(t,x(t))$ in a Taylor series to the first order in dx and dt:

$$dF \sim F_t dt + F_x dx + \frac{1}{2}F_{xx}dx^2 + \cdots$$

Now $dx^2(t) = [a(t)dt + b(t)dW(t)]^2$. But we have seen that

$$E[dW^2(t)] = E[W^2(dt)] = var[W(dt)] = dt$$

$$var[dW^2(t)] = 3dt^2 - dt^2 = 2dt^2$$

hence the variance of $dW^2(t)$ is negligible compared to its mean for dt small. In the limit $dt \to 0$ we can approximate $dW^2(t)$ with its mean i.e.

$$\lim_{dt\to 0} dW^2(t) \sim dt$$

.

Hence, to the first order in dt

$$dx^2(t) = [a(t)dt + b(t)dW(t)]^2 = b^2(t)dt$$

Replacing (and removing for simplicity the explicit dependence from t)

$$dF = F_t dt + F_x dx + \frac{1}{2}F_{xx}dx^2 =$$

$$F_t dt + F_x[adt + bdW] + \frac{1}{2}F_{xx}b^2 dt$$

and

$$dF = (F_t + aF_x + \frac{1}{2}b^2F_{xx})dt + bF_x dW$$

$$dF = (F_t + aF_x + \frac{1}{2}b^2F_{xx})dt + bF_x dW$$

A special case:

$$dx(t) = \mu x(t)dt + \sigma x(t)dW(t)$$

i.e. $a(t) = \mu x(t)$ and $b(t) = \sigma x(t)$ then

$$dF = (F_t + \mu x F_x + \frac{1}{2}\sigma^2 x^2 F_{xx})dt + \sigma x F_x dW$$

Example:

$$F(t) = \ln x(t)$$

$$F_t = 0$$

$$F_x = \frac{1}{x}$$

$$F_{xx} = -\frac{1}{x^2}$$

$$d\ln x = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW$$

and

$$\ln(x(t)) = \ln(x(0)) + \int_0^t (\mu - \frac{1}{2}\sigma^2)du + \int_0^t \sigma dW(u)$$

with solution

$$x(t) = x(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)}$$

Hence the unique solution of

$$dx(t) = \mu x(t)dt + \sigma x(t)dW(t)$$

with initial condition $x(0)$ is

$$x(t) = x(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)}$$

The Black-Scholes equation.

The Black-Scholes-Merton option pricing model (BSM) is a continuous time model for the pricing of a European option on an asset which pays no dividends. The model was developed by Fisher Black and Myron Scholes in 1973 and formalized and extended by Robert Merton in the same year. In the model it is assumed that the stock price follows a geometric Brownian motion,

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t) \tag{1}$$

where μ is a constant drift coefficient, σ is a constant and positive volatility coefficient, and $S_0 \in R^+$ is the initial stock price. W_t is a one-dimensional Brownian motion defined on a filtered probability space $(\Omega, P, \mathcal{F})$.

We will assume that trading takes place continuously in time, unrestricted borrowing or lending of funds is possible at the same constant rate and the market is frictionless (no transaction costs, no taxes, no short selling constraints).

Using Itô formula we can integrate the previous equation and obtain:

$$S(t) = S(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)}$$

A risk-free security, called saving account or money market account, is also available in the market whose price process is given by

$$dB(t) = rB(t)dt$$

or

$$B(t) = B(0)e^{rt}$$

where by convention $B(0) = 1$. This security is called the market account.

The derivation of the Black and Scholes equation for the option price relies on the observation that the option payoff can be replicated by holding a continuously rebalanced position in the underlying stock and cash. If the replicating strategy is self-financing, and if the market is arbitrage-free then the arbitrage price of the option is given by the value process of the replicating portfolio.

Consider a derivative with maturity date T and payoff $h = f(S(T), T)$

At time $t = 0$ construct a portfolio consisting of $a(0)$ units of stocks and $b(0)$ units of cash,

$$\Pi(0) = a(0)S(0) + b(0)B(0).$$

Suppose that a replicating strategy exists such that

$$a(T)S(T) + b(T)B(T) = h.$$

We also impose that the portfolio is self-financing, i.e. we do not add or take cash away from the portfolio at any time. The strategy $(a(t), b(t))$ is self-financing if

$$d\Pi = a(t)dS(t) + b(t)dB(t)$$

i.e. the change in value of the portfolio only depends on the change of the asset prices.

Then

$$d\Pi(t) = a(t)dS(t) + b(t)dB(t) = a(t)[\mu S(t)dt + \sigma S(t)dW(t)] + b(t)B(t)rdt$$

But if the two portfolios are the same at maturity and self-financing then, by no-arbitrage, at any time $t \leq T$, it must also holds

$$\Pi(t) = f(S(t), t)$$

and also

$$d\Pi(t) = df(S(t), t)$$

For the derivative, using Itô's lemma we get

$$df(S(t), t) = [f_t(S(t), t) + \mu S(t)f_S(S(t), t) + \frac{1}{2}\sigma^2 S^2(t)f_{SS}(S(t), t)]dt +$$
$$\sigma S(t)f_S(S(t), t)dW(t)$$

Imposing

$$a(t)[\mu S(t)dt + \sigma S(t)dW(t)] + b(t)B(t)rdt =$$
$$[f_t(S(t), t) + \mu S_t f_S(S(t), t) + \frac{1}{2}\sigma^2 S^2 f_{SS}(S(t), t)]dt + \sigma S f_S(S(t), t)dW(t)$$

and equating terms in dW we find

$$a(t) = f_S(S(t), t)$$

while equating terms in dt we find

$$[f_t(S(t), t) + \mu S(t)f_S(S(t), t) + \frac{1}{2}\sigma^2 S^2 f_{SS}(S(t), t)] = rb(t)B(t) + a(t)\mu S(t) =$$
$$r[-a(t)S(t) + f(S(t), t)] + a(t)\mu S(t) =$$
$$r[-f_S(S(t), t)S(t) + f(S(t), t)] + f_S(S(t), t)\mu S(t)$$

where we used

$$b(t)B(t) = -a(t)S(t) + f(S(t), t)$$

Finally, taking all time dependences away we obtain

$$f_t + rSf_S + \frac{1}{2}\sigma^2 S^2 f_{SS} = rf$$

which is the Black and Scholes equation. The solution of the Black and Sholes equation gives for a call and a put

$$C(0) = S(0)N(d_1) + KB^{-1}(T)N(d_2)$$
$$P(0) = KB^{-1}(T)N(-d_2) - S(0)N(-d_1)$$

with

$$d_1 = \frac{\ln(S(0)/K) + (r + 1/2\sigma^2)T}{\sigma\sqrt{T}}$$

and

$$d_2 = \frac{\ln(S(0)/K) + (r - 1/2\sigma^2)T}{\sigma\sqrt{T}}$$

Feynman-Kac Theorem

We anticipate that the solution to the partial differential equation

$$F_t + rSF_S + \frac{1}{2}\sigma^2 S^2 F_{SS} - rF = 0$$

with final condition $F(T, S(T))$ admits the representation

$$F(0, S(0)) = e^{-rT} E[F(T, S(T))]$$

where

$$dS(t) = rS(t)dt + \sigma S(t)dW(t)$$

Hence Feynman-Kac theorem this provides a link between the portfolio replication approach and the risk-neutral valuation approach, according to which we can price a derivative by discounting it's expected future payoff in the risk-neutral world.

The Greeks

In the formulas below T is time to maturity and should be replaced with $T - t$ if the current time is different from $t = 0$.

- for a call

$$\Delta = \frac{\partial C}{\partial S} = N(d_1) > 0$$

$$Vega = \frac{\partial C}{\partial \sigma} = S\sqrt{T}n(d_1) > 0$$

$$\Theta = \frac{\partial C}{\partial T} = -\frac{S\sigma}{2\sqrt{T}}n(d_1) - Kre^{-rT}N(d_2)$$

usually negative

$$\rho = \frac{\partial C}{\partial r} = TKe^{-rT}N(d_2) > 0$$

$$\Gamma = \frac{\partial^2 C}{\partial S^2} = \frac{n(d_1)}{S\sigma\sqrt{T}} > 0$$

- for a put

$$\Delta = \frac{\partial C}{\partial S} = N(d_1) - 1 < 0$$

$$Vega = \frac{\partial C}{\partial \sigma} = S\sqrt{T}n(d_1) > 0$$

$$\Theta = \frac{\partial C}{\partial T} = -\frac{S\sigma}{2\sqrt{T}}n(d_1) + Kre^{-rT}N(-d_2)$$

usually negative

$$\rho = \frac{\partial C}{\partial r} = -TKe^{-rT}N(-d_2) < 0$$

$$\Gamma = \frac{\partial^2 C}{\partial S^2} = \frac{n(d_1)}{S\sigma\sqrt{T}} > 0$$

Example: Delta for a Call

To prove formula we use identity

$$n\left(\frac{\alpha}{\sqrt{s}} - \beta\sqrt{s}\right) = e^{2\alpha\beta} n\left(\frac{\alpha}{\sqrt{s}} + \beta\sqrt{s}\right)$$

with $\alpha = \log(S(t)/K)/\sigma$, $\beta = (r - 1/2\sigma^2)/\sigma$ and $s = T - t$.

$$C(t) = S(t)N(d_1) + KB^{-1}(T-t)N(d_2)$$

$$N(d_{1,2}) = \int_{-\infty}^{d_{1,2}} n(x)\ dx$$

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

$$d_{1,2} = \frac{\ln(S(0)/K) + (r \pm 1/2\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

$$\Delta = \frac{\partial C}{\partial S} = N(d_1) + S(t)\frac{\partial N(d_1)}{\partial S} + KB(T-t)\frac{\partial N(d_2)}{\partial S}$$

Using Leibniz's rule for differentiating an integral

$$\frac{d}{dz}\int_{a(z)}^{b(z)} f(x,z)\ dx = \int_{a(z)}^{b(z)} \frac{\partial f(x,z)}{\partial z}\ dx + f(b(z),z)\frac{db(z)}{dz} - f(a(z),z)\frac{db(z)}{dz}$$

Let's do it at $t = 0$

$$\Delta = N(d_1) + S(0)n(d_1)\frac{\partial d_1}{\partial S(0)} - Ke^{-rT}n(d_2)\frac{\partial d_2}{\partial S(0)}$$

$$\frac{\partial d_{1,2}}{\partial S(0)} = \frac{1}{S(0)\sigma\sqrt{T}}$$

$$\Delta = N(d_1) + \frac{1}{\sigma\sqrt{T}}\left[n(d_1) - \frac{K}{S(0)}e^{-rT}n(d_2)\right]$$

$$n(d_2) = \frac{1}{\sqrt{2\pi}}\ e^{-d_2^2/2} = \frac{1}{\sqrt{2\pi}}\ e^{-(d_1 - \sigma\sqrt{T})^2/2} =$$

$$\frac{1}{\sqrt{2\pi}}\ e^{-d_1^2 - \sigma^2 T/2 + d_1\sigma\sqrt{T}} = n(d_1)\ e^{-\sigma^2 T/2 + d_1\sigma\sqrt{T}}$$

$$\Delta = N(d_1) + n(d_1)\frac{1}{\sigma\sqrt{T}}\left[1 - \frac{K}{S(0)}e^{-rT}e^{-\sigma^2 T/2 + d_1\sigma\sqrt{T}}\right]$$

but

$$\left[1 - \frac{K}{S(0)}e^{-rT}e^{-\sigma^2 T/2 + d_1\sigma\sqrt{T}}\right] = \left[1 - e^{-rT + \ln(K/S(0)) - \sigma^2 T/2 + d_1\sigma\sqrt{T}}\right] =$$

$$\left[1 - e^{-(rT+\sigma^2 T/2) - ln(S(0)/K) + \ln(S(0)/K) + (r+1/2\sigma^2)(T-t)}\right] = 0$$

and

$$\Delta = N(d_1)$$

From Black-Scholes PDE

$$f_t + rSf_S + \frac{1}{2}\sigma^2 S^2 f_{SS} = rf$$

we obtain

$$rf = \frac{1}{2}S^2\sigma^2\Gamma + rS\Delta - \Theta$$

Hedging a Portfolio of derivatives

The variation of a portfolio $\Pi(t, S(t))$ of derivatives contract is given by

$$d\Pi = \Pi_t dt + \Pi_S dS + \frac{1}{2}\Pi_{SS}dS^2$$

Define

$$\Delta_\Pi = \Pi_S$$
$$\Gamma_\Pi = \Pi_{SS}$$
$$\Theta_\Pi = \Pi_t$$

If we construct the portfolio so that it's Delta nutral i.e. $\Delta_\Pi = 0$ we find

$$d\Pi = \Theta_\Pi dt + \frac{1}{2}\Gamma_\Pi dS^2. \qquad (*)$$

Also from B&S equation

$$r\Pi = \frac{1}{2}S^2\sigma^2\Gamma_\Pi + rS\Delta_\Pi + \Theta_\Pi$$

which gives, if $\Delta_\Pi = 0$,

$$\frac{1}{2}S^2\sigma^2\Gamma_\Pi = r\Pi - \Theta_\Pi.$$

Hence, if Θ is large and positive, Γ tends to be large and negative (the term with r is usually trascurable). Hence the root of eq (*) are both real.

Exercise

Plot $d\Pi$ as a function of dS and discuss how to opimally Gamma-hedge the portfolio in the different cases of positive or negative Theta.

More on Stochastic process

Partition:

A partition $\{F_i\}_{i=1}^K$ of a set A is a family of mutually disjoint subsets of A whose union is A.

Example

Consider a two step the binomial model. At time 0 it is impossible to tell any of the elements of the sample space Ω apart, so the members of the set $F_0 = \{UU; UD; DU; DD\}$ all appear the same. At time 1 it is possible to divide the sample space up into two distinguishable 'partitions' $F_1^u = \{UU; UD\}$ and $F_1^d\{DU; DD\}$

depending on whether the initial movement in the market was up or down.

Finally at time 2, there are four different market states, that we can differ- entiate between

$F_2^{uu} = \{U, U\}$ $F_2^{ud} = \{U, D\}$ $F_2^{du} = \{D, U\}$ $F_2^{dd} = \{D, D\}$

Each of these collection of sets divides up or 'partitions' the sample space at the relevant time. From our example we see that associated with the random movements of the market there is a natural sequence of partitions of the sample space in terms of F_0, F_1, F_2. We call such a sequence a filtration.

Filtration: A filtration $\mathcal{F}$ is a family $\{F_i\}_{i=1}^K$ of partitions of Ω. An additional constraint that we require is that the partitions at later times respect the earlier partitions. That is, if $i < k$, then every partition at the earlier time is equal to the union of some set of partitions at the later time. In our example we have $F_0 = F_1^u \cup F_1^d$.

We call the probability system $(\Omega, \Sigma, \mathcal{F}, P)$ a filtered probability space. The filtration gives some sort of time ordering to the probability space.

In general we will think of the filtration $\mathcal{F}_t$ as the information available up to time t. We also assume that we never forget, thus if $s < t$

$$\mathcal{F}_s \subset \mathcal{F}_t$$

The information we are concerned with will typically be the information on the values of some stochastic process $X(t)$, i.e. the filtration generated by the process $X(t)$.

If we want to be more formal we would say that a filtration $\mathcal{F}$ is an increasing family of σ-algebra $\mathcal{F}_t$ where $\mathcal{F}_t$ is the σ-algebra of all events that have occurred or not by time t.

The natural filtration of a stochastic process X is defined by

$$\mathcal{F}_t^X = \sigma(\cup_{u\in[0,t]}\sigma(X_u))$$

Note that the union of σ-algebra is not necessarily a σ-algebra, this is why we need to take the σ-algebra of the union.

Example

Remember that the σ-algebra of the event iΣ s a set of subsets of the sample space Ω, satisfying the following conditions:

- $\Omega \in \Sigma$
- if $A, B \in \Sigma$, then $A \cup B \in \Sigma$
- if $A \in \Sigma$, then $\Omega - A \in \Sigma$

Also remember that the smallest σ-algebra containing a set A is $\{\Omega, \phi, A, A^C\}$.

Take $\Omega = \{1, 2, 3, 4\}$ and

$$\mathcal{F} = \{\phi, \Omega, \{1\}, \{2, 3, 4\}\}$$

$$\mathcal{G} = \{\{\phi, \Omega, \{4\}, \{1, 2, 3\}\}$$

Then

$$\mathcal{G} \cup \mathcal{F} = \{\{\phi, \Omega, \{1\}, \{4\}, \{1, 2, 3\}, \{2, 3, 4\}\}$$

is not a σ-algebra. In fact the set

$$\{1\} \cup \{4\} \notin \mathcal{G} \cup \mathcal{F}$$

but

$$\sigma(\mathcal{G} \cup \mathcal{F}) = \{\{\phi, \Omega, \{1\}, \{4\}, \{1, 2, 3\}, \{2, 3, 4\}, \{2, 3\}, \{1, 4\}\}$$

is a σ-algebra. It is in fact the smallest σ-algebra containing $\mathcal{G} \cup \mathcal{F}$.

Conditional Expectation

The conditional expectation

$$E[X|\mathcal{F}_t]$$

denotes the expectation conditional to the information available at time t.

Properties

- **Iterated conditional expectations property or tower property**

  If X is a r.v. and $\mathcal{F}_1 \subseteq \mathcal{F}_2$ then

  $$E[E[X|\mathcal{F}_2]|\mathcal{F}_1] = E[E[X|\mathcal{F}_1]|\mathcal{F}_2] = E[X|\mathcal{F}_1]$$

- **Conditional mean formula**

  $$E[E[X|\mathcal{F}]] = E[X]$$

- If X is independent of $\mathcal{F}$ then

$$E[X|\mathcal{F}] = E[X]$$

- If X and Y are two r.v. and $X \in \mathcal{F}$ then

$$E[XY|\mathcal{F}] = XE[Y|\mathcal{F}]$$

- A r.v. X is said to be *measurable* with respect the σ-algebra $\mathcal{F}$ if for any real number x the subset $\{\omega \in \Omega : X(\omega) = x\} \subset \mathcal{F}$. We usually use an abbreviated notation $X \in \mathcal{F}$. It means that the value of X is known given $\mathcal{F}$, i.e.

$$E[X|\mathcal{F}] = X$$

- A stochastic process on a filtered probability space $(\Omega, \Sigma, P, \mathcal{F})$ is a sequence of r.v. $X = \{X(t); t \geq 0\}$.
- If X is such that for each t, $X(t)$ is a $\mathcal{F}(t)$-measurable r.v., i.e. $E[X(t)|\mathcal{F}(t)] = X(t)$, then we say that X is *adapted* to the filtration $\mathcal{F} = \{\mathcal{F}(t); t > 0\}$.
- A stochastic process X is said to be *predictable* with respect to the filtration $\mathcal{F}$ if each r.v. $X(t)$ is measurable respect to $\mathcal{F}_{t-1}$, i.e. $E[X(t)|\mathcal{F}_{t-1}] = X(t)$ (this obviously implies $E[X(t)|\mathcal{F}_t] = X(t)$ given that $\mathcal{F}_{t-1} \subseteq \mathcal{F}_t$).

Martingales

Given a filtered probability space and an adapted stochastic process $Z = \{Z(t) : t > 0\}$ such that $E[|Z(t)|] < \infty$ for all t (integrability condition), we call it a

- martingale if $E[Z(t+s)|\mathcal{F}(t)] = Z(t)$ for each $s > 0$
- super-martingale if $E[Z(t+s)|\mathcal{F}(t)] \leq Z(t)$ for each $s > 0$
- sub-martingale if $E[Z(t+s)|\mathcal{F}(t)] \geq Z(t)$ for each $s > 0$

It also holds that $E[Z(t)] = E[E[Z(t)|\mathcal{F}(0)]] = E[Z(0)]$ for each $t > 0$, i.e. a martingale is constant on average and model a fair game (while a supermartingale is decreasing on average and models an unfavorable game, a submartingale is increasing on average and models a favorable game).

Martingale Properties of Wiener processes:

- Wiener processes are martingales:

$$E[W(t)|\mathcal{F}(s)] = E[W(s) + (W(t) - W(s))|\mathcal{F}(s)] = W(s)$$

- $W^2(t) - t$ is a martingale.

$$E[W^2(t) - t|\mathcal{F}(s)] = E[(W(s) + W(t) - W(s))^2|\mathcal{F}(s)] - t =$$

$$W^2(s) + 2W(s)E[W(t) - W(s)|\mathcal{F}(s)] + E[(W(t) - W(s))^2|\mathcal{F}(s)] - t =$$

$$W^2(s) + 0 + (t - s) - t = W^2(s) - s$$

which is the martingale condition.

- $e^{\theta W(t) - \frac{1}{2}\theta^2 t}$ is a martingale.

$$E[e^{\theta W(t) - \frac{1}{2}\theta^2 t}|\mathcal{F}(s)] = e^{-\frac{1}{2}\theta^2 t}E[e^{\theta(W(t) - W(s) + W(s))}|\mathcal{F}(s)] =$$

$$e^{-\frac{1}{2}\theta^2 t}e^{\theta W(s)}E[e^{\theta(W(t) - W(s))}|\mathcal{F}(s)] =$$

$$e^{-\frac{1}{2}\theta^2 t}e^{\theta W(s)}E[e^{\theta(W(t) - W_s)}] = e^{-\frac{1}{2}\theta^2 t}e^{\theta W(s)}e^{\frac{1}{2}\theta^2(t-s)^2} = e^{\theta W(s) - \frac{1}{2}\theta^2 s}$$

which is the martingale condition.

Equivalent measures

Given two measures P and Q defined on the same probability space $(\Omega, \mathcal{F})$ we say that P is absolutely continuous with respect to Q if $P(A) = 0$ whenever $Q(A) = 0$, for $A \in \mathcal{F}$. If also $P(A) = 0$ whenever $Q(A) = 0$, for $A \in \mathcal{F}$, i.e. P and Q have the same null sets, then we say that the two measure are equivalent and write $P \sim Q$.

Definition:

A probability measure P^* on $(\Omega, \mathcal{F})$ equivalent to P is called a martingale measure for a process S if S follows a P^*-martingale with respect to the filtration $\mathcal{F}$, i.e.

$$E^*[S(t+s)|\mathcal{F}(t)] = S(t)$$

for each $s > 0$.

Fundamental theorems of Asset Pricing

- We say that a market is *viable* if it does not admit arbitrage strategies.

  No-Arbitrage Theorem:

  A market is viable if and only if there exists a probability measure P^* equivalent to P under which discounted prices are martingale, i.e.

  $$\tilde{S}(t) = E^*[\tilde{S}(t+s)|\mathcal{F}(t)] \quad for \;\; all \quad s > 0$$

  One side of the proof, i.e. that if P^* exists than a market is viable is easy. The other side is difficult.

  We do not give any of the two proofs here.

- **Completeness Theorem:**

  A market is said to be *complete* if all contingent claims are attainable.

  A market is complete if and only if there exists a **unique** probability measure P^* equivalent to P under which discounted prices are martingales.

- **Fundamental Theorem of Asset Pricing**

  In an arbitrage free, complete market there exists a unique equivalent martingale measure P^*, equivalent to P, under which discounted prices are martingales.

- **Risk-Neutral pricing Formula**

  In an arbitrage free, complete markets arbitrage prices of contingent claims are their discounted expected values under the risk-neutral measure P^*.

  This comes from the fact that in a risk-free complete market the price of a contingent claim, h, is given by the value process of any replicating strategy $V_\phi(t)$. A replicating strategy exists if the market is complete, i.e.

  $$V_\phi(T) = h$$

  But we know that if P^* is an equivalent martingale measure and ϕ any self-financing strategy than the value process $V_\phi(t)$ is a P^*-martingale with respect to the filtration $\mathcal{F}$. We also know that a self-financing strategy remains self-financing after a numeraire change.

  Hence

  $$V_\phi(t) = B(t)\tilde{V}_\phi(t) = B(t)E^*[\tilde{V}_\phi(T)|\mathcal{F}(t)] = B(t)E^*[h/B(t)h|\mathcal{F}(t)]$$

  and

  $$V_\phi(0) = E^*[B^{-1}(t)h]$$

  as stated.

  As the equivalent martingale measure is unique than also prices of contingent claims are uniquely defined in complete markets.

Radon-Nikodym theorem

Proposition

Let $W(t)$ be a d-dimensional s.B.m. defined on a probability space $(\Omega, P, \mathcal{F})$. If $\mathcal{F}$ is the Brownian filtration for any probability measure Q equivalent to P on $(\Omega, \mathcal{F})$ there exists a d-dimensional adapted process θ such that

$$\frac{dQ}{dP}|_{\mathcal{F}(t)} = \eta(t) = \exp\left(\int_0^t \theta(u) dW(u) - \frac{1}{2}\int_0^t |\theta(u)|^2 du\right)$$

Given sufficient integrability on the process θ (Novikov's condition) $\eta(t)$ is a martingale.

If θ is constant and unidimensional

$$\eta(t) = e^{\theta W(t) - \frac{1}{2}\theta^2 t}$$

and it is easy to prove that $\eta(t)$ is a P martingale (we have already proved this earlier).

Note that $\eta(t)$ is a follows

$$d\eta(t) = \theta\eta(t) dW(t)$$

and

$$E^P[\eta(t)] = E^P[\eta(0)] = 1$$

Bayes Formula

Let's P and Q be two equivalent probability measures defined on a common probability space $(\Omega, \mathcal{F})$. Define

$$\eta = \frac{dQ}{dP}$$

Let's ψ be a r.v. integrable with respect to Q and $\mathcal{G} \subset \mathcal{F}$. Then

$$E_Q[\psi|\mathcal{G}] = \frac{E_P[\psi\eta|\mathcal{G}]}{E_P[\eta|\mathcal{G}]}$$

Note that, given $E_P[\eta] = 1$, Bayes formula reduces in the unconditional to the obvious case

$$E_P[\eta\psi] = E_Q[\psi]$$

Girsanov Theorem

If Q and P are two equivalent measures s.t. the Radon-Nikodym derivative of Q respect to P equals

$$\eta(t) = \frac{dQ}{dP} = \exp\left(\int_0^T \theta(u)dW(u) - \frac{1}{2}\int_0^T |\theta(u)|^2 du\right)$$

and $W^P(t)$ is a standard Brownian motion on the space $(\Omega, P, \mathcal{F})$, then the process

$$\bar{W}^Q(t) = W^P(t) - \int_0^t \theta(u)du$$

follows a standard Brownian motion on the space $(\Omega, Q, \mathcal{F})$.

If θ is constant it becomes

$$\bar{W}^Q(t) = W^P(t) - \theta t$$

Proof:

Use moment generating function:

$$E^Q[e^{\alpha W^P(t)}] = E^P[\eta(t)e^{\alpha W^P(t)}] = E^P[e^{\theta W^P(t) - \frac{1}{2}\theta^2 t}e^{\alpha W^P(t)}] = e^{-\frac{1}{2}\theta^2 t}E^P[e^{(\theta+\alpha)W^P(t)}] =$$

$$e^{-\frac{1}{2}\theta^2 t}e^{\frac{1}{2}(\theta+\alpha)^2 t} = e^{\theta\alpha t + \frac{1}{2}\alpha^2 t}$$

This shows that under Q, $W^P(t) \sim N(\theta t, t)$, i.e. it takes a drift θt. Thus to have a s.B.m under Q we need to subtract this drift. The s.B.m. under Q is

$$W^Q(t) = W^P(t) - \theta t$$

In this case the change of measure correspond to a change of drift. In finance $-\theta$ is called market price of risk (respect to P). The risk-neutral world is a world where the market price of risk is zero, while in the real world the market price of risk is $-\theta$.

Note that the volatility of the Brownian motion does not change when we change measure, it is only the drift that changes.

Stochastic Integral

The integral cannot be defined in the usual sense, i.e. pathways (i.e. trajectory by trajectory) because the W-trajectories are of locally unbounded variation. But it is possible to give a global $\mathcal{L}^2$ definition of the integral.

Consider a function $X(s, W_s)$ s.t.

$$\int_0^t E[X^2(s, W_s)]ds < \infty]$$

i.e. X is in $\mathcal{L}^2$. Take the interval $(0, t)$ and subdivide it in n subintervals. Define $u_k = \frac{kt}{n}$. The stochastic integral is defined as

$$I_t(X(s, W_s)) = \int_0^t X(s, W_s)dW(s) = \lim_{n\to\infty} \sum_{k=0}^{n-1} X(u_k)[W(u_{k+1}) - W(u_k)]$$

Note that we have taken the forward increments of the Wiener process (this is the Itô convention).

The quantity above is a random variable and the procedure of taking this limit is tricky.

Properties

- Under some regularity assumption for the function X the stochastic integral is a martingale, i.e.

  $$E[I_t(X)|\mathcal{F}_s] = I_s(X)$$

  and

  $$E[I_t(X)] = E[I_0(X)] = 0$$

  This means

  $$E[\int_0^t X(u, W_u)dW(u)|\mathcal{F}_s] = \int_0^s X(u, W_u)dW(u)$$

  and

  $$E[\int_0^t X(u, W_u)dW(u)] = E[\int_0^0 X(u, W_u)dW(u)] = 0$$

  and

  $$E[\int_t^s X(u, W_u)dW(u)|\mathcal{F}_t] = \int_t^t X(u, W_u)dW(u) = 0$$

- It$\hat{o}$ isometry

  $$E[I_t^2(X)] = \int_0^t E[X^2(s, W_s)]ds$$

- If $X(t)$ is a deterministic function $I_t(X)$ is a normal random variable with zero mean and variance $\int_0^t X^2(s)]ds$.

Example Let's calculate

$$var[I_T(T-t)] = E[(\int_0^T (T-t)dW(t))^2] - (E[\int_0^T (T-t)dW(t)])^2 = E[\int_0^T (T-t)^2 dt] =$$

$$\int_0^T (T^2 - 2tT + t^2)dt = T^3 - \frac{2T^3}{2} + \frac{T^3}{3} = \frac{T^3}{3}$$

Ito's Lemma and stochastic integral

Consider a process following

$$dX(t) = a(t)dt + b(t)dW(t)$$

Take a smooth function $F(t, X(t))$ of it. Itô's lemma says that

$$dF = (F_t + aF_X + \frac{1}{2}b^2 F_{XX})dt + bF_X dW$$

Example 1:

Take $X(t) = W(t)$, i.e. a=0, b=1, and $F = tW_t$ and use Itô's Lemma to calculate dF.

$$F_t = W_t, \quad F_W = t, \quad F_{WW} = 0$$

and

$$dF = d(tW_t) = tdW + W_t dt$$

Let's calculate

$$I_T(t) = \int_0^T t dW_t$$

we have seen that

$$d(tW_t) = tdW + W_t dt$$

hence

$$\int_0^T t dW = \int_0^T d(tW_t) - \int_0^T W_t dt = TW_T - \int_0^T W_t dt.$$

Hence the arithmetic average of $W(t)$ defined as

$$A_W(t) = \frac{1}{t}\int_0^t W_s ds = \frac{1}{t}\int_0^t (t-s)dW = \frac{1}{t}I_t(t-s)$$

and has mean zero and variance $\frac{1}{t^2}\frac{t^3}{3} = \frac{t}{3}$,

In fact $A_W(T) \sim N(0, \frac{t}{3})$

Example 2:

Take $X(t) = W(t)$ and $F(t, X(t)) = W^2(t)$. Use Itô's Lemma to calculate $dF = dW^2$.

$$F_t = 0, \quad F_W = 2W, \quad F_{WW} = 2$$

$$dF = dW_t^2 = 2W_t dW_t + \frac{1}{2}\ 2\ dt = 2W_t dW_t + dt$$

Let's calculate

$$I_T(W_t) = \int_0^T W_t dW_t$$

We know that $W_t dW_t = \frac{1}{2}(dW_t^2 - dt)$

Hence,

$$\int_0^T W_t dW_t = \frac{1}{2}\int_0^T dW_t^2 - \frac{1}{2}\int_0^T dt = \frac{1}{2}W_T^2 - \frac{1}{2}T$$

Risk-neutral pricing approach

Binomial Tree

Assume time horizon is T and set of dates in our market is $t = 0, 1, \cdots, T$. Also assume the money market account process is:

$$B_t = (1+r)^t$$

and

$$S_{t+1} = uS_t \quad with \;\; probability \quad p$$
$$S_{t+1} = dS_t \quad with \;\; probability \quad 1-p$$

with $0 < d < u$.

Define

$$Z_t = \frac{S_{t+1}}{S_t}$$

Z_t take only two values u and d. They are independent and identically distributed random variables.

First we need to find the risk neutral measure, i.e. a measure P^* such that

$$E^*[\tilde{S}_{t+1}|\mathcal{F}_t] = \tilde{S}_t$$

where $S_t = B_t\tilde{S}_t$.

This implies

$$E^*[\tilde{S}_{t+1}/\tilde{S}_t|\mathcal{F}_t] = E^*[Z_{t+1}B_t/B_{t+1}|\mathcal{F}_t] = E^*[Z_{t+1}/(1+r)|\mathcal{F}_t] = 1$$

or

$$1 + r = E^*[Z_{t+1}|\mathcal{F}_t] = E^*[Z_{t+1}] = up^* + d(1-p^*)$$

from where we find

$$p^* = \frac{1+r-d}{u-d}$$

We have already seen that in the risk-neutral world the growth rate of the asset is the risk free rate

$$E^*[S_{\Delta t}] = S_0(1+r)$$

Nonetheless with the standard choice

$$u = 1/d = e^{\sigma\sqrt{\Delta t}}$$

one finds that

$$var[S_{\Delta t}] = p^*u^2 + (1-p^*)d^2 - [p^*u + (1-p^*)d]^2 = \sigma^2\Delta t$$

which says that the variance in the real world is the same as in the risk neutral world.

Black and Scholes Model

In the model it is assumed that the stock price follows a geometric Brownian motion,

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{2}$$

where μ is a constant drift coefficient, σ is a constant and positive volatility coefficient, and $S_0 \in R^+$ is the initial stock price. W_t is a one-dimensional Brownian motion defined on a filtered probability space $(\Omega, P, \mathcal{F})$.

NB: In this section we use the convention of denoting time dependence as S_t, B_t, W_t etc. instead of $S(t), B(t), W(t)$.

We will assume that trading takes place continuously in time, unrestricted borrowing or lending of funds is possible at the same constant rate and the market is frictionless (no transaction costs, no taxes, no short selling constraints).

Using Itô formula we can integrate the previous equation and obtain:

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}$$

A risk-free security is also available in the market whose price process is given by

$$dB_t = rB_t dt$$

or

$$B_t = B_0 e^{rt}$$

where by convention $B_0 = 1$. This security is called the money market account.

The derivation makes use of the martingale approach and does not explicitly shows how to construct the replication portfolio. It starts from the observation (no arbitrage theorem) that a market is arbitrage free if and only if there exists a probability measure P^* equivalent to P under which discounted asset prices are martingale.

We need first to find the equivalent martingale measure.

We want to find a constant θ such that the discounted price

$$\tilde{S}_t = S_t / B_t = S_t e^{-rt}$$

follows a martingale respect to the measure P^* i.e.

$$\tilde{S}_t = \tilde{S}_0 e^{-\frac{1}{2}\sigma^2 t + \sigma W_t^*}$$

where P^* is defined by

$$\eta_T = \frac{dP^*}{dP} = e^{\theta W_T - \frac{1}{2}\theta^2 T}$$

Lemma: The unique martingale measure P^* for the discounted price process is given by the Radon-Nikodym derivative

$$\frac{dP^*}{dP} = e^{\theta W_T - \frac{1}{2}\theta^2 T}$$

with

$$\theta = \frac{r - \mu}{\sigma}$$

Proof

Let's use Girsanov Theorem:

Under P

$$\tilde{S}_t = \tilde{S}_0 e^{(\mu - r - \frac{1}{2}\sigma^2)t + \sigma W_t}$$

Under P^*

$$W_t^* = W_t - \theta t$$

hence

$$\tilde{S}_t = \tilde{S}_0 e^{(\mu - r - \frac{1}{2}\sigma^2)t + \sigma(W_t^* + \theta t)} = \tilde{S}_0 e^{(\mu - r + \sigma\theta)t - \frac{1}{2}\sigma^2 t + \sigma W_t^*}$$

if we choose θ s.t.

$$\mu - r + \sigma\theta = 0 \qquad \rightarrow \qquad \theta = \frac{r - \mu}{\sigma}$$

then

$$\tilde{S}_t = \tilde{S}_0 e^{-\frac{1}{2}\sigma^2 t + \sigma W_t^*}$$

which is a martingale.

Also

$$d\tilde{S}_t = \tilde{S}_t \sigma W_t^*.$$

This also implies that under P^*

$$dS_t = rS_t dt + \sigma S_t dW_t^*$$

In fact:

$$dS_t = \mu S_t dt + \sigma S_t dW_t = \mu S_t dt + \sigma S_t (dW_t^* + \theta dt) =$$
$$\mu S_t dt + \sigma S_t \left(dW_t^* + \frac{r - \mu}{\sigma} dt \right) = rS_t dt + \sigma S_t dW_t.$$

Define the market price of risk $\lambda = -\theta$. We can rewrite the equation for S_t as

$$dS_t = (\mu - \lambda\sigma) S_t dt + \sigma S_t dW_t^*$$

This notation is normally used in incomplete markets, when the value of λ is not known, as we will see later.

We can now price a derivative using the risk-neutral pricing methodology.

Let's take a simple vanilla call option whose payoff is

$$h_T = (S_T - K)^+.$$

The option value is given by the risk-neutral formula

$$C_t = B_t E^*[(S_T - K)^+ B_T^{-1} | \mathcal{F}_t].$$

where $B_0 = 1$ and $B_t = e^{rt}$. Under P^*

$$\frac{dS_t}{S_t} = rdt + \sigma dW_t^*.$$

At $t = 0$ notation simplifies as

$$C_0 = E^*[B_T^{-1}(S_T - K)^+] = E^*[B_T^{-1}(S_T - K)I_{\{S_T>K\}}] =$$

$$E^*[(S_T B_T^{-1} I_{\{S_T>K\}}] - e^{-rT} K E^*[I_{\{S_T>K\}}] = J_1 - J_2.$$

Note that in general

$$E[I_D] = \int_\Omega I_D dP = \int_D dP = P(D).$$

Let's calculate J_2 first.

$$J_2 = e^{-rT} K E^*[I_{\{S_T>K\}}] = e^{-rT} K P^*\{S_T > K\} =$$

$$e^{-rT} K P^*\{S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma W_T^*} > K\} =$$

$$e^{-rT} K P^*\{(r - \frac{1}{2}\sigma^2)T + \sigma W_T^* > \log(K/S_0)\} =$$

$$e^{-rT} K P^*\{\sigma W_T^* > -\log(S_0/K) - (r - \frac{1}{2}\sigma^2)T\} =$$

rewrite $W_T = \epsilon\sqrt{T}$ with $\epsilon \sim N(0,1)$

$$J_2 = e^{-rT} K P^* \left\{ \epsilon > -\frac{\log(S_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \right\}$$

But because of the symmetry of the normal density,

$$P^*(\epsilon > -d) = P^*(\epsilon < d)$$

thus

$$J_2 = e^{-rT} K P^* \left\{ \epsilon < \frac{\log(S_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \right\} = e^{-rT} K N(d_2)$$

where

$$d_2 = \frac{\log(S_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

and $N(\cdot)$ is the cumulative normal distribution.

Now let's calculate J_1

$$J_1 = E^*[S_T e^{-rT} I_{\{S_T>K\}}] = S_0 E^*[e^{-\frac{1}{2}\sigma^2 T + \sigma W_T^*} I_{\{S_T>K\}}].$$

Let's introduce a new measure Q defined by

$$\eta = \frac{dQ}{dP^*} = e^{-\frac{1}{2}\sigma^2 T + \sigma W_T^*}$$

η is a martingale s.t. $E^*[\eta] = 1$ so it is a valid change of measure. Now we can use formula

$$E_P[\psi_t \eta_t] = E_Q[\psi_t]$$

and get

$$J_1 = S_0 E^*[\eta I_{\{S_T > K\}}] = S_0 E^Q[I_{\{S_T > K\}}] = S_0 Q\{S_T > K\} =$$
$$S_0 Q\{S_0 e^{-\frac{1}{2}\sigma^2 T + \sigma W_T^*} > K e^{-rT}\}$$

Now let's use Girsanov. Under Q the s.B.m. is

$$\bar{W}_t = W_t^* - \sigma t$$

Hence we can replace in equation above

$$W_t^* = \bar{W}_t + \sigma t$$

and obtain

$$J_1 = S_0 Q\{e^{\frac{1}{2}\sigma^2 T + \sigma \bar{W}_T} > K e^{-rT}\} =$$
$$S_0 Q\{\sigma \bar{W}_T > \log(K/S_0) - rT - \frac{1}{2}\sigma^2 T\} = S_0 Q\{-\sigma \bar{W}_T < \log(S_0/K) + (r + \frac{1}{2}\sigma^2)T\}$$

Take $\xi = -\frac{\bar{W}_T}{\sqrt{T}} \sim N(0,1)$ and we get

$$J_1 = S_0 Q\left\{\xi < \frac{\log(S_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right\} = S_0 N(d_1)$$

where

$$d_1 = \frac{\log(S_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

Putting the pieces together we obtain

$$C_0 = S_0 N(d_1) - K e^{-rT} N(d_2)$$

The value of the put can be found using put-call parity

$$P_t = C_t - S_t + K e^{r(T-t)}$$

Exercise: Show the formula above holds.

Feynman-Kac Theorem

Define $X(s)$ to be the solution, for $t \leq s \leq T$, of the stochastic differential equation

$$dX(s) = \tilde{\mu}(X(s), s)ds + \tilde{\sigma}(X(s), s)dW(s)$$

with initial condition

$$X(t) = x$$

Define

$$\mathcal{A} = \tilde{\mu}(t, x)\frac{\partial}{\partial x} + \frac{1}{2}\tilde{\sigma}^2(t, x)\frac{\partial^2}{\partial^2 x}$$

Assume the function $F(t, x)$ is a solution, for $0 \leq t \leq T$, of the problem

$$\frac{\partial F}{\partial t}(t, x) + \mathcal{A}F(t, x) = 0 \qquad (*)$$

with boundary condition $F(T, x) = h(x)$

If

$$\int_t^T E[\tilde{\sigma}(s, X(s))F_x(s, X(s))]ds < \infty$$

then F admits the representation

$$F(0, x) = E[h(X(T)]$$

Proof:

Using Itô's formula

$$F(T, X(T)) = F(0, X(0)) + \int_t^T \left[\frac{\partial F}{\partial t}(s, X(s)) + \mathcal{A}F(s, X(s))\right] ds +$$
$$\int_t^T \tilde{\sigma}(s, X(s))\frac{\partial F}{\partial x}(s, X(s))dW(s).$$

Under the above assumptions the stochastic integral is a martingale with constant expectation (which is zero). Hence replacing assumption (*) and taking expectation on both sides

$$E[F(T, X(T))] = F(0, X(0))$$

or

$$F(0, x) = E[h(X(T))].$$

Analogously the solution to the partial differential equation

$$F_t + \tilde{\mu}F_x + \frac{1}{2}\tilde{\sigma}^2 F_{xx} - rF = 0$$

with final condition $F(T, x) = h(x)$ admits the representation

$$F(0,x) = e^{-rT}E[h(X(T))].$$

Proof: Take

$$G(t,X(t)) = e^{-rt}F(t,X(t))$$
$$G(0,X(0)) = F(0,X(0))$$
$$G_t = -re^{-rt}F + e^{-rt}F_t$$
$$G_x = e^{-rt}F_x$$
$$G_{xx} = e^{-rt}F_{xx}$$

and

$$dG = e^{-rt}[-rF + F_t + \tilde{\mu}F_x + \frac{1}{2}\tilde{\sigma}^2 F_{xx})]dt + \tilde{\sigma}e^{-rT}F_x dW$$

If

$$-rF + F_t + \tilde{\mu}F_x + \frac{1}{2}\tilde{\sigma}^2 F_{xx} = 0$$
$$G(T,X(T)) = G(0,X(0)) + e^{-rT}\int_0^T \sigma(X(u))F_x(u,X(u))dW(u)$$

Taking expectation

$$E[G(T,X(T))] = G(0,X(0))$$

Replacing $G(t,X(t)) = e^{-rt}F(t,X(t))$

$$F(0,X(0)) = e^{-rT}E[F(T,X(T))] = e^{-rT}E[h(X(T)].$$

Remember the Black and Scholes equation is

$$f_t + rSf_x + \frac{1}{2}\sigma^2 S^2 f_{xx} - rf = 0$$

If we identify F above with the price of a derivative and X with the price of a stock and replace

$$\tilde{\mu}(S(t),t) = rS(t)$$
$$\tilde{\sigma}(S(t),t) = \sigma S(t)$$

this theorem provides the link between risk neutral pricing and the Black-Scholes partial differential equation:

We can price a derivative by discounting it's expected future payoff

$$f(0) = e^{-rT}E^*[h]$$

in a world where the stock price $X(t)$ follows the SDE

$$dX(t) = rX(t)dt + \sigma X(t)dW(t)$$

i.e. in the risk-neutral world

Spot and forward interest rates

Term Structure of Interest rates

In financial markets the price for credit is referred to as the interest rate. It is determined by demand and supply of credit. It is an inter-temporal price. i.e. the price today for money that is to be returned at some future date. Fisher (1930) develop a theory on how interest rates can be derived from the consumption and saving decisions of individuals. Most people prefer to consume now relative to postpone gratification to a future time. As a reward for being patient people ask for an extra amount of future consumption. make any payment in between.

Different rates of interest are quoted for different lengths of time $r(t, T)$. The term structure of interest rates is the name given to the pattern of interest rates available on instruments of similar credit risk but with different terms to maturity.

Forward rates

The forward rate is the interest rate, implied by zero rates, that you could earn (pay) for an investment (loan) made at a given time in the future till another given time in the future. This is a deterministic rate, agreed in advance by the lender and borrower.

Forward rates are defined by the no arbitrage condition

$$(1 + r(t, t_n))^n = (1 + r(t, t_{n-1}))^{n-1}(1 + f(t, t_{n-1}, t_n)),$$

from where

$$f(t, t_{n-1}, t_n) = \frac{(1 + r(t, t_n))^n}{(1 + r(t, t_{n-1}))^{n-1}} - 1$$

Hence for example

$$f(t, t, t_1) = r(t, t_1)$$

$$f(t, t_1, t_2) = \frac{(1 + r(t, t_2))^2}{(1 + r(t, t_1))} - 1$$

With continuous compounding:

$$Ie^{r(t,t_2)(t_2-t)} = Ie^{r(t,t_1)(t_1-t)+f(t,t_1,t_2)(t_2-t_1)}$$

Hence

$$f(t, t_1, t_2) = \frac{r(t, t_2)(t_2 - t) - r(t, t_1)(t_1 - t)}{t_2 - t_1} =$$

$$= r(t, t_2) + (t_1 - t)\frac{r(t, t_2) - r(t, t_1)}{t_2 - t_1}.$$

So if $r(t, t_2) > r(t, t_1)$ then $f(t, t_1, t_2) > r(t, t_2)$ and if $r(t, t_2) < r(t, t_1)$ then $f(t, t_1, t_2) < r(t, t_2)$.

In the limit $t_2 \to t_1$

$$f(t,t_1) = f(t,t_1,t_1) = r(t,t_1) + (t_1 - t)\frac{\partial r(t,t_1)}{\partial t}$$

This is the instantaneous forward rate applicable to a very short period starting at t_1.

Bonds

These are fixed income securities which obligate the issuer (borrower) to make a pre-specified set of payments to the purchaser (lender).

Term to maturity date on which the debt will be redeemed by borrower.

Par value or face value: amount the issuer agrees to redeem at the maturity date.

Coupon: is the periodic interest payment made to the bond-holder during the life of the bond. Usually paid semiannually.

Floating rate security: describes security in which the coupon rate is reset periodically.

Zero coupon bond: make no periodic interest payments. Only return the face amount at maturity N.

Coupon bond: make periodic interest payments and return face amount at maturity.

Bonds evaluation

How do we evaluate the price of the bond? This is the amount you are willing to pay today to receive the face value of the bond, plus the coupons, at futures times. The price of the bond is the present value of this stream of future cash flows. Assume a bond last m years and pays coupons semi-annually. If we use continuously compounded rates we would obtain

$$B(0,m) = \sum_{i=1}^{2m} C_i e^{-r_c(0,t_i)t_i} + Me^{-r_c(0,m)m}$$

Coupon bonds can be thought of as a collection of zero coupon bonds:

$$B(0,m) = \sum_{t=1}^{2m} C_t p(0,t) + Mp(0,M)$$

<u>Example: zero coupon bond</u>

$m = 10$ years

face value: $M = 1000\$$

$$P(0,m) = \frac{M}{(1 + r(0,m))^m}$$

Example: coupon bond

$m = 10$ years

Face value: $M = 1000\$$

Coupon rate: 8% of face value in semiannual coupon installments. Assume semiannually compounded rates

$$C = r_c/2 * M = 0.04 * 1000 = 40\$.$$

$$B(0,m) = \sum_{t=1}^{20} \frac{40}{(1 + r_s(0,t/2)/2)^t} + \frac{1000}{(1 + r_s(0,m)/2)^{2m}}$$

Yield to Maturity:

This is the discount rate at which the present value of all future payments would equal the present price of a bond. Also called Internal Rate of Return. If maturity is m and coupons are paid semiannually

$$P(0,m) = \sum_{t=1}^{2m} C e^{-y\frac{t}{2}} + M e^{-ym}$$

I have assumed a continuously compounded yield. The yield to maturity is unique to a given bond at that time.

For a zero-coupon bond yield coincides with spot rate

$$y_a = r(0,m) = \left(\frac{M}{P(0,m)}\right)^{1/m} - 1$$

I have assumed an annually compounded yield.

Bootstrapping

This is the procedure to extract spot rates from bond prices. If we want risk free rates the issuer need to be credit worthy. Usually government bonds (also called treasury bonds) are used for this purpose.

How do we extract rates form coupon bonds?

If bonds were all zero coupon bonds this process would be very easy

$$r(0,m) = \left(\frac{M}{P(0,m)}\right)^{1/m} - 1$$

Nonetheless most long maturity bonds are coupon bonds.

Exercise

Say we want the spot rates for year one, two and three. Assume a zero coupon bond is traded with maturity one year $P(0,1)$ while two coupon bonds are traded, paying annual coupons c, with maturity two years $B(0,2)$ and three years $B(0,3)$.

We do this for continuously compounded rates.

$$
\begin{aligned}
P(0,1) &= Me^{-r(0,1)} \\
r(0,1) &= -\log\frac{P(0,1)}{M} \\
B(0,2) &= ce^{-r(0,1)} + (c+M)e^{-r(0,2)2} \\
r(0,2) &= -\frac{1}{2}\log\left[\frac{B(0,2) - ce^{-r(0,1)}}{c+M}\right] \\
B(0,3) &= ce^{-r(0,1)} + ce^{-r(0,2)2} + (c+M)e^{-r(0,3)3} \\
r(0,3) &= -\frac{1}{3}\log\left[\frac{B(0,3) - ce^{-r(0,1)} - ce^{-r(0,2)2}}{c+M}\right]
\end{aligned}
$$

Forward rates, spot rates and zero coupon bond prices

We can define forward rates in terms of zero coupon bonds that deliver one dollar at maturity T, i.e. $p(T,T) = 1$, and

$$p(t,T) = e^{-r(t,T)(T-t)}$$

via the relationship

$$e^{r(t,t_2)(t_2-t)} = e^{r(t,t_1)(t_1-t)+f(t,t_1,t_2)(t_2-t_1)}$$

or

$$e^{f(t,t_1,t_2)(t_2-t_1)} = \frac{e^{r(t,t_2)(t_2-t)}}{e^{r(t,t_1)(t_1-t)}} = \frac{e^{-r(t,t_1)(t_1-t)}}{e^{-r(t,t_2)(t_2-t)}} = \frac{p(t,t_1)}{p(t,t_2)}$$

where we assumed continuous compounding. Then

$$f(t,t_1,t_2) = -\frac{\log p(t,t_2) - \log p(t,t_1)}{(t_2-t_1)}$$

The instantaneous forward rate ($t_2 \to t_1$)

$$f(t,t_1) = -\frac{\partial \log p(t,t_1)}{\partial t_1}$$

The future short (or spot) rate ($t = t_1$)

$$r(t_1,t_2) = -\frac{\log p(t_1,t_2)}{(t_2-t_1)}$$

The current short (or spot) rate ($t_1 = t$)

$$r(t, t_2) = -\frac{\log p(t, t_2)}{(t_2 - t)}$$

The instantaneous short rate (or spot rate) is defined by

$$r(t) = f(t, t)$$

If rates are stochastic the bank account evolves as

$$B(T) = B(t)e^{\int_t^T r(s)ds}$$

and the zero coupon price

$$p(t, T) = e^{-\int_t^T f(t,s)ds}$$

Furthermore the risk-neutral formula tells us that

$$p(t, T) = B(t, T)E^Q[p(T, T)/B(T, T)|\mathcal{F}(t)] = E^Q[1B(t, T)/B(T, T)|\mathcal{F}(t)] = E^Q[e^{-\int_t^T r(s)ds}|\mathcal{F}(t)]$$

where $r(t)$ is the instantaneous spot rate and in this contest we call the risk neutral measure Q (instead of P^*).

Short rate models

We have seen that we can express the price of zero-coupon bonds either in term of spot interest rates, in term of forward rates, or in term of the instantaneous spot rate:

$$p(t, T) = e^{-r(t,T)(T-t)} = e^{-\int_t^T f(t,s)ds} = E^Q[e^{-\int_t^T r(s)ds}|\mathcal{F}(t)]$$

Thus we can choose to either model the dynamics of each bond, or the dynamics of forward rates, or the dynamics of the instantaneous short rate.

If we choose to model the instantaneous short rate

$$dr(t) = \mu(r(t), t)dt + \sigma(r(t), t)dW(t)$$

we can recover bonds price using risk-neutral evaluation formulas

$$p(t, T) = E^Q[e^{-\int_t^T r(s)ds}|\mathcal{F}_t]$$

and the payoff of a generic contingent claim with payoff $\Phi(r(T), T)$ is given by

$$\Phi(r(t), t) = E^Q[e^{-\int_t^T r(s)ds}\Phi(r(T), T)|\mathcal{F}_t]$$

Instantaneous forward rates then can be obtained by

$$f(t, T) = -\frac{\partial \log p(t, T)}{\partial T} = \frac{E^Q[r(T)e^{-\int_t^T r(s)ds}|\mathcal{F}_t]}{E^Q[e^{-\int_t^T r(s)ds}|\mathcal{F}_t]}$$

Proof:

$$f(t,T) = -\frac{\partial \log p(t,T)}{\partial T} = -\frac{\partial}{\partial T}\log(E^Q[e^{-\int_t^T r(s)ds})|\mathcal{F}_t] = -\frac{1}{E^Q[e^{-\int_t^T r(s)ds}|\mathcal{F}_t]}\frac{\partial}{\partial T}E^Q[e^{-\int_t^T r(s)ds}|\mathcal{F}_t]$$

We use Fubini theorem to take the derivative inside the expectation:

$$f(t,T) = -\frac{1}{E^Q[e^{-\int_t^T r(s)ds}|\mathcal{F}_t]}E^Q\left[\frac{\partial}{\partial T}e^{-\int_t^T r(s)ds}|\mathcal{F}_t\right] =$$

$$-\frac{1}{E^Q[e^{-\int_t^T r(s)ds}|\mathcal{F}_t]}E^Q\left[e^{-\int_t^T r(s)ds}\frac{\partial}{\partial T}\left(-\int_t^T r(s)ds\right)|\mathcal{F}_t\right]$$

We now use Liebnitz rule of differentiation of an integral to find

$$f(t,T) = \frac{1}{E^Q[e^{-\int_t^T r(s)ds}|\mathcal{F}_t]}E^Q\left[e^{-\int_t^T r(s)ds}r(T)|\mathcal{F}_t\right] = \frac{E^Q[r(T)e^{-\int_t^T r(s)ds}|\mathcal{F}_t]}{E^Q[e^{-\int_t^T r(s)ds}|\mathcal{F}_t]}$$

Model Calibration

Under the risk neutral measure Q the short rate dynamics becomes

$$dr(t) = (\mu(r(t),t) - \lambda(t)\sigma(t,S(t))dt + \sigma(r(t),t)dW^Q(t) =$$

$$dr(t) = m(r(t),t)dt + \sigma(r(t),t)dW^Q(t)$$

where λ is the market price of risk and we defined

$$m(r(t),t) = \mu(r(t),t) - \lambda(t)\sigma(t,S(t)$$

Problem: Bond market models are incomplete.

This is because interest rates are not traded and it is not possible to replicate the argument of Black and Scholes by constructing an hedging portfolio that replicate the payoff of the contingent clam (here the bond is the contingent claim). Nonetheless we need to guarantee that the market is arbitrage free. We can only work with interest rate models that admit an equivalent martingale measure but we cannot guarantee that this measure is unique. We are in the situation where there are several equivalent martingale measures, or several possible choices of the market price of risk process $\lambda(t)$, and for each choice we end up with a different no arbitrage price.

But then which measure (or which λ) do we choose?

The one chosen by the market! This is done by inverting the term structure i.e. by matching the theoretical prices of bonds predicted by the model (which will depend on λ) with the actual bond prices in the market. Then we can use the "calibrated" value of λ to price more complex structure, for example interest rate derivatives.

Affine Models

For calibration to be easy to implement we can only work with some relatively simple short rate models, in particular we want models that allow to write the bond price in the form

$$p(t,T) = e^{A(t,T)-B(t,T)r(t)}$$

where A and B are deterministic functions. These models are said to posses an affine term structure. Affine here refers to the fact that the term structure is a linear function of a small set of common factors, in this case

$$r(t,T) = B(t,T)r(t) - A(t,T)$$

If

$$p(t,T) = e^{A(t,T)-B(t,T)r(t)}$$

then, using Ito's formula

$$dp(t,T) = p_t dt + p_r dr + \frac{1}{2}p_{rr} d < r >= \left(p_t + mp_r + \frac{1}{2}\sigma^2 p_{rr}\right) dt + \sigma p_r dW^Q(t) =$$

$$p(t,T)\left[A_t(t,T) - B_t(t,T)r(t) - B(t,T)m(r(t),t) + \frac{1}{2}B(t,T)\sigma^2(r(t),t)\right] dt + p(t,T)B(t,T)\sigma(r(t),t)dW^Q(t)$$

But under the risk-neutral measure it also has to hold that

$$dp(t,T) = p(t,T)r(t)dt + p(t,T)\sigma_P(t,T,r(t))dW^Q(t)$$

It is by imposing that

$$A_t(t,T) - B_t(t,T)r(t) - B(t,T)m(t,r(t)) + \frac{1}{2}B(t,T)\sigma^2(t,r(t)) = r(t)$$

that we can find the possible functional forms for $m(t,r(t))$ and $\sigma(t,r(t))$. The kind of models that have an affine term structure are of the form

$$dr(t) = m(r(t),t)dt + \sigma(r(t),t)dW^Q(t)$$

with

$$m(t,r(t)) = \alpha(t)r(t) + \beta(t)$$

$$\sigma(t,r(t)) = \sqrt{\gamma(t)r(t) + \delta(t)}$$

Also the volatility of the bond price is given by

$$\sigma_P(t,T,r(t)) = B(t,T)\sigma(r(t),t)$$

Finally the forward at dynamic can be found from

$$f(t,T) = -\frac{\partial}{\partial T}\log p(t,T) = -\frac{\partial}{\partial T}[A(t,T) - B(t,T)r(t)]$$

Short rate models

The most well known affine model (there are many more non affine) are

- Vasicek
$$dr(t) = c(\mu - r(t))dt + \sigma dW^Q(t)$$
- CIR
$$dr(t) = a(b - r(t))dt + \sigma\sqrt{r(t)}dW^Q(t)$$
- Ho and Lee
$$dr(t) = \mu(t)dt + \sigma dW^Q(t)$$
- Hull and White
$$dr(t) = (b(t) - a(t)r(t))dt + \sigma(t)dW^Q(t)$$

Vasicek Model

Let the interest rate $r(t)$ follow under the risk-measure Q the process

$$dr(t) = c(\mu - r(t))dt + \sigma dW^Q(t),$$

where c, μ are constant, σ is constant and positive, and $W(t)$ is a standard Brownian motion. The initial value of r is $r(0)$.

The process is mean reverting because the drift term changes sign above or below the threshold level μ pushing the process back to this level with a speed that is proportional to c.

Remember that Itô's lemma states that given an Itô's process $X(t)$ following

$$dX(t) = a(t)dt + b(t)dW(t),$$

and a smooth function $F(t, X(t))$, then

$$dF(t, X(t)) = (F_t + a(t)F_X + \frac{1}{2}b^2(t)F_{XX})dt + b(t)F_X dW(t),$$

where

$$F_t = \frac{\partial F(t, X)}{\partial t}, \quad F_X = \frac{\partial F(t, X)}{\partial X}, \quad F_{XX} = \frac{\partial^2 F(t, X)}{\partial X^2}.$$

We can define $Y(t) = e^{ct}r(t)$ and use Itô's lemma to derive the SDE for $Y(t)$.

The derivatives are $Y_t = ce^{ct}r(t)$, $Y_r = e^{ct}$, $Y_{rr} = 0$.
Hence

$$dY(t) = [ce^{ct}r(t) + (c\mu - cr(t))e^{ct}]dt + \sigma e^{-ct}dW^Q(t) = c\mu e^{ct}dt + \sigma e^{ct}dW^Q(t).$$

Integrating

$$Y(t) = Y(0) + c\mu \int_0^t e^{cs} ds + \sigma \int_0^t e^{cs} dW^Q(s)$$

and replacing back $Y(t) = e^{ct} r(t)$

$$e^{ct} r(t) = r(0) + \mu(e^{ct} - 1) + \sigma \int_0^t e^{cs} dW^Q(s)$$

or

$$r(t) = e^{-ct} r(0) + \mu(1 - e^{-ct}) + \sigma e^{-ct} \int_0^t e^{cs} dW^Q(s).$$

The stochastic integral is a martingale with zero mean which implies

$$E[r(t)] = r(0)e^{-ct} + \mu(1 - e^{-ct})$$

Also

$$var(r(t)) = E[r^2(t)] - E[r(t)]^2$$

where

$$E[r^2(t)] = E\left[e^{-ct} r(0) + \mu(1 - e^{-ct}) + \sigma e^{-ct} \int_0^t e^{cs} dW^Q(s)\right]^2 =$$

$$[r(0)e^{-ct} + \mu(1 - e^{-ct})]^2 + E[(\sigma e^{-ct} \int_0^t e^{cs} dW^Q(s))^2] +$$

$$2[r(0)e^{-ct} + \mu(1 - e^{-ct})] E[\sigma e^{-ct} \int_0^t e^{cs} dW^Q(s)]$$

But $E[\sigma e^{-ct} \int_0^t e^{cs} dW^Q(s)] = 0$, because the stochastic integral is a zero-mean martingale. Hence

$$E[r^2(t)] = E[r(t)]^2 + E[(\sigma e^{-ct} \int_0^t e^{cs} dW^Q(s))^2]$$

Now remember the It$\hat{o}$'s isometry property of stochastic integrals:

$$E[I_t^2(X)] = \int_0^t E[X^2(s, W_s)] ds$$

Then

$$E[r^2(t)] - E[r(t)]^2 = \sigma^2 e^{-2ct} \int_0^t e^{2cs} ds = \sigma^2 e^{-2ct} \, \frac{e^{2ct} - 1}{2c} = \frac{\sigma^2}{2c}(1 - e^{-2ct}).$$

and $var(r(t)) = \frac{\sigma^2}{2c}(1 - e^{-2ct})$.

The long term mean of this process is μ and the long term variance is $\frac{\sigma^2}{2c}$.

From this expression is now possible to find closed form solution for bond prices, forward rates and bond options (see Bjork to see formulas).

Bond prices

$$p(0,T) = E^Q[e^{-\int_0^T r(t)dt}]$$

Let's define $X(T) = \int_0^T r(t)dt$ and calculate it:

$$X(T) = \int_0^T r(t)dt = \int_0^T \left[e^{-ct}r(0) + \mu(1-e^{-ct}) + \sigma e^{-ct}\int_0^t e^{cs}dW^Q(s)\right]dt =$$

$$\mu T + (r(0)-\mu)\frac{1-e^{-cT}}{c} + \sigma\int_0^T\int_0^t e^{-c(t-s)}dW^Q(s)dt =$$

$$\mu T + (r(0)-\mu)\frac{1-e^{-cT}}{c} + \sigma\int_0^T\left(\int_0^t e^{-c(t-s)}dt\right)dW^Q(s) =$$

$$\mu T + (r(0)-\mu)\frac{1-e^{-cT}}{c} + \sigma\int_0^T \frac{1-e^{-c(T-s)}}{c}dW^Q(s)$$

$X(T)$ is a normal random variable. Following the previous calculation its mean and variance are

$$E[X(T)] = \mu T + (r(0)-\mu)\frac{1-e^{-cT}}{c}$$

$$var[X(T)] = \frac{\sigma^2}{c^2}\int_0^T (1-e^{-c(T-s)})^2 ds =$$

$$\frac{\sigma^2}{c^2}\int_0^T (1+e^{-2c(T-s)} - 2e^{-c(T-s)})ds =$$

$$\frac{\sigma^2}{c^2}\left(T + \frac{1-e^{-2cT}}{2c} - \frac{2(1-e^{-cT})}{c}\right) = \frac{\sigma^2}{2c^3}\left(2cT - e^{-2cT} - 3 + 4e^{-cT}\right)$$

Then using moment generating function of normal random variables we find

$$p(0,T) = E^Q[e^{-\int_0^T r(t)dt}] = e^{(-E[X]+\frac{1}{2}var[X])} =$$

$$\exp\left\{-\mu T - (r(0)-\mu)\frac{1-e^{-cT}}{c} + \frac{\sigma^2}{4c^3}\left(2cT - e^{-2cT} - 3 + 4e^{-cT}\right)\right\} = e^{A(0,T)-B(0,T)r(0)}$$

with

$$A(0,T) = -\mu T + \mu\frac{1-e^{-cT}}{c} + \frac{\sigma^2}{4c^3}\left(2cT - e^{-2cT} - 3 + 4e^{-cT}\right)$$

$$B(0,T) = \frac{1-e^{-cT}}{c}$$

At a generic time we would find

$$p(t,T) = e^{A(t,T)-B(t,T)r(t)}$$

with

$$A(t,T) = -\mu(T-t) + \mu\frac{1-e^{-c(T-t)}}{c} + \frac{\sigma^2}{4c^3}\left(2c(T-t) - e^{-2c(T-t)} - 3 + 4e^{-c(T-t)}\right)$$

$$B(t,T) = \frac{1-e^{-c(T-t)}}{c}$$

The instantaneous forward prices are

$$f(t,T) = -\frac{\partial}{\partial T}[A(t,T) - B(t,T)r(t)] = r(t)e^{-c(T-t)} + \mu(1-e^{-c(T-t)}) - \frac{\sigma^2}{2c^2}[1+e^{-2c(T-t)} - 2e^{-c(T-t)}] =$$

$$r(t)e^{-c(T-t)} + \mu(1-e^{-c(T-t)}) - \frac{\sigma^2}{2c^2}(1-e^{-c(T-t)})^2$$

The volatility of the bond prices are

$$\sigma_p(t,T,r(t)) = \frac{\sigma}{c}[1-e^{-c(T-t)}]$$

and the term structure of interest rate is

$$r(t,T) = \frac{1}{T-t}[r(t)B(t,T) - A(t,T)] =$$

$$\frac{1}{T-t}\left[\mu(T-t) + (r(t)-\mu)\frac{1-e^{-c(T-t)}}{c} + \frac{\sigma^2}{4c^3}\left(2c(T-t) - e^{-2c(T-t)} - 3 + 4e^{-c(T-t)}\right)\right]$$

Two ways of calibrating the model

- use historical data of instantaneous spot rates to estimate the parameters c, μ, σ, then calculate bond prices $p(t,T)$. Normally the prices calculated this way never match the observed ones $p_{obs}(t,T)$.
- use the current bond prices $p_{obs}(t,T)$, to estimate the parameters. This works better but given we normally have more bonds than parameters still we cannot fit very accurately. Furthermore the calibration has to be redone daily because the bond the term structure changes and we end up with parameters that fluctuate over time, while the assumption of the model is that the parameter are fixed and not stochastic.

Term structure models

Short rate models are relatively easy and closed form analytical formulas can be obtained for bond prices and interest rates derivatives. Nonetheless it is difficult to obtain a realistic term structure for the forward rates.

Heath Jarrow and Morton (HJM) suggested to model the entire forward curve by assuming for every maturity T

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW^Q(t)$$

where $f(0,T)$ is given the current forward rates. Bond prices are then given by

$$P(t,T) = e^{-\int_t^T f(t,s)ds}$$

the instantaneous spot rate is obtained by taking $\lim t \to T$

$$r(T) = f(T,T) = f(0,T) + /int_0^T \alpha(t,T)dt + \sigma(t,T)dW^Q(t)$$

HJM do not suggest any specific model but a framework for modeling bond markets.

Forward neutral measure and Zero-coupon bonds as numeraire

Historically the money market account has been chosen as a numeraire but the zero coupon bond would be the natural choice if one looks at the price of an asset which only gives a single payoff at a well defined future time T. The value of the zero coupon bond which pays one dollar at time T, i.e. $P(T,T) = 1$, can be treated as the value of any other contingent claim. The risk-neutral valuation formula gives:

$$P(t,T) = B(t)E^Q[P(T,T)B^{-1}(T)|\mathcal{F}_t]$$

and $P(t,T)/B(t)$ is a Q-martingale.

Then we can use the change of numeraire theorem above and take $P(t,T)$ as the numeraire. The new equivalent martingale measure respect to which security prices expressed in units of the zero-coupon bonds are martingales is

$$\frac{dQ_T}{dQ} = \frac{P(T,T)B(0)}{P(0,T)B(T)} = \frac{1}{P(0,T)B(T)}$$

Security prices can then be expressed as

$$Z(t) = P(t,T)E^{Q_T}[Z(T)/P(T,T)|\mathcal{F}_t] = P(t,T)E^{Q_T}[Z(T)|\mathcal{F}_t]$$

Q_T is called the forward-neutral martingale measure. The relative price process of a security $Z(t)$ t with respect to $P(t,T)$, $Z(t)/P(t,T)$, is called the forward price $F_Z(t)$ of the security Z. $F_Z(t)$ is a Q_T-martingale so

$$F_Z(t) = Z(t)/P(t,T) = E^{Q_T}[Z(T)/P(T,T)|\mathcal{F}_t] = E^{Q_T}[Z(T)|\mathcal{F}_t]$$

The forward price of a security which pays no dividend up to time T is equal to the expectation of the value at time T of this security under the forward neutral martingale measure.

The value of a contingent claim is given by

$$C(0) = B(0)E^Q[h/B(T)] = P(0,T)E^{Q_T}[h]$$

Note that if r is constant then

$$P(t,T) = e^{-r(T-t)}$$

and

$$\eta = \frac{dQ_T}{dQ} = 1$$

Hence $\theta = 0$. This means that in this case the Bownian motion has the same drift under the two measures Q and Q_T.

Exotic Derivatives:

Binary:

- Cash or Nothing:

$$C_T = XI_{\{S_T>K\}}$$
$$P_T = XI_{\{S_T<K\}}$$

- Asset or Nothing:

$$C_T = S_T I_{\{S_T>K\}}$$
$$P_T = S_T I_{\{S_T<K\}}$$

Gap Options:

$$C_T = (S_T - X)I_{\{S_T>K\}} = BAC_T - BCC_T$$
$$P_T = (X - S_T)I_{\{S_T<K\}} = BCP_T - BAP_T$$

Supershare Options:

Payoff:

$$h_T = \frac{S_T}{K_1} I_{\{K_1<S_T<K_2\}} \qquad for \quad K_1 < K_2$$

Chooser Options:

$$h_T = (S_T - K)^+ I_A + (K - S_T)^+ I_{A^C}$$

where A is the event

$$A = \{C(S_{T_0}, T - T_0, K) > P(S_{T_0}, T - T_0, K)\}$$

and A^C is the complement of A.

Power Options

$$C_T = (S_T^\alpha - K)^+$$

$$P_T = (K - S_T^\alpha)^+$$

Powered Options

$$C_T = [(S_T - K)^+]^\alpha$$

$$C_T = [(K - S_T)^+]^\alpha$$

Option depending on several underlying

Exchange Options

$$h_T = (X_T - Y_T)^+$$

Options on the best/worst of two assets

$$C_{max}(T) = \max(S_1(T), S_2(T)) = S_2(T) + [S_1(T) - S_2(T)]^+$$
$$C_{min}(T) = \min(S_1(T), S_2(T)) = S_1(T) - [S_1(T) - S_2(T)]^+$$

So they can be expressed in terms of exchange options.

Spread Options

$$C_T = [(S_T^1 - S_T^2) - K]^+$$

Basket Options

$$C_T = [I_T - K]^+$$

where

$$I_t = \sum_{i=1}^{n} w_i S_t^i$$

Foreign exchange options

Option struck in foreign currency

$$C_T^1 = Q_T(S_T^f - K^f)^+$$

Option struck in domestic currency

$$C_T^2 = (S_T^f Q_T - K^d)^+$$

Quanto Options

$$C_T^3 = \bar{Q}(S_T^f - K^f)^+$$

Path dependent Options

Asian Options

Asian options have a payoff which depends on the average value of the underlying asset over the maturity of the options.

We have two cases

- $h_T = (S_T - \bar{S}_T)^+$
- $h_T = (\bar{S}_T - K)^+$

The average can be defined as arithmetic or geometric and in continuous or discrete time:

- Arithmetic Average

  discrete time:
  $$A_D = \frac{1}{N}\sum_{i=1}^{N} S_{t_i}$$
  continuous time:
  $$A_C = \frac{1}{T}\int_0^T S_t dt$$
- Geometric Average

  discrete time:
  $$G_D = (\prod_{i=1}^{N} S_{t_i})^{1/N}$$
  continuous time:
  $$G_C = e^{\frac{1}{T}\int_0^T \ln S_t dt}$$

Lookback Options

$$C_T = (S_T - m_T^S)$$
$$P_T = (M_T^S - S_T)$$

where
$$m_T^S = \min\{S_t; 0 \le t \le T\}$$
and
$$M_T^S = \max\{S_t; 0 \le t \le T\}$$

Barrier Options

$$c_T^{uo} = (S_T - K)^+ I_{\{M_T^S \leq H\}}$$
$$c_T^{ui} = (S_T - K)^+ I_{\{M_T^S \geq H\}}$$
$$c_T^{uo} + c_T^{ui} = (S_T - K)^+ (I_{\{M_T^S \leq H\}} + I_{\{M_T^S \geq H\}}) = c_T$$

from where

$$c_T^{uo} = c_T - c_T^{ui}$$

$$c_T^{do} = (S_T - K)^+ I_{\{m_T^S \leq H\}}$$
$$c_T^{di} = (S_T - K)^+ I_{\{m_T^S \geq H\}}$$
$$c_T^{do} + c_T^{di} = (S_T - K)^+ (I_{\{m_T^S \leq H\}} + I_{\{m_T^S \geq H\}}) = c_T$$

from where

$$c_T^{do} = c_T - c_T^{di}$$

Energy Derivatives

Domestic oil and petroleum prices were deregulated in the 1980s, and natural gas prices were partially deregulated. Before price deregulation, the market for domestic oil and gas derivatives was limited. Under price regulation, the U.S. Department of Energy (DOE), the Federal Energy Regulatory Commission (FERC), and the State public utility commissions (PUCs) directly or indirectly controlled the prices of domestic crude oil, petroleum products, wellhead natural gas, pipeline transmission, and retail gas service. Government was also deeply involved in deciding the merits of pipeline investment and siting. The immediate effect of price controls was to stabilize price.

Unfortunately, price certainty was paid for with shortages in some areas and surplus elsewhere and by complex cross-subsidies from areas where prices would have been lower to areas where prices would have been higher, with accompanying efficiency costs. Currently, the prices of crude oil, natural gas, and all petroleum products are free from Federal regulation. The FERC continues to impose price ceilings on pipeline services and has approval authority for new pipeline construction. Most States continue to regulate prices for small users of natural gas (residences and commercial enterprises), but large usersparticularly, power plants, which accounted for about 21 percent of the Nations natural gas consumption in 2001, and petrochemical plantsare generally free to make their best deals.

Most of the energy futures and options on futures are traded on the New York Mercantile Exchange (NYMEX). However, the trading volume of electricity futures is less than electricity forwards traded in the over-the-counter (OTC) markets. A large variety of energy derivatives are traded among market participants in the OTC markets, including forward contracts, swaps, plain vanilla options, and exotic (i.e., non-standard) options like spark spread options, swing options and swaptions.

The light, sweet crude oil futures contract is the most actively traded commodity in the world. Since oil is an inherently non-standard commodity, the various exchanges have chosen several reference grades (such as WTI) to form the basis of the futures and options market. Light, sweet crude oil futures contracts traded on NYMEX are written on units of 1,000 barrels of oil for delivery in Cushing, Oklahoma. These contracts may be either futures contracts of various maturities, calendar spread, crack spread, or average price options.

Calendar spreads are a risk management device that allows the holder to purchase (or sell) oil based on the price difference between 2 delivery dates. Crack spread options are written on the price difference between two related commodities. The crack spread options traded on NYMEX for WTI light, sweet crude oil are written on the price differentials of crude oil futures and heating oil futures and crude oil futures and gasoline futures. Average price options (or APO's) are Asian type options where the holder is allowed to buy or sell oil based on the average futures price for a given period of time.

While the crude oil market may be the most actively traded commodity, the growth rate of the U.S. natural gas market has far exceeded that of oil. The NYMEX natural gas futures contract has seen the largest percentage trade volume gains of any product launched in the exchange's history. Currently, natural gas composes nearly one-fourth of the total U.S. energy consumption with expected gas consumption tripling to 13 trillion cubic feet by 2020.

Structurally, the natural gas sector is very different from crude oil. As opposed to having a small number of firms which have high levels of vertical integration (i.e. firms that wholly own several steps in the production, refining, and transportation processes), the North American natural gas sector is very disjoint; with different firms controlling drilling, refining and pipeline distribution.

Natural gas futures contracts are traded on the NYMEX in units of 10,000 million Btu's (British thermal energy units) for delivery at the Henry Hub in Louisiana. These contracts trade with maturity dates ranging from 1 to 72 months. Additionally, NYMEX offers vanilla calls and puts, calendar spread and spark spread options. Due to the increased reliance of electricity suppliers on natural gas, the spark spread options allows one to manage the risk associated with the spread between natural gas futures and electricity futures prices.

Swaps:

OTC financially settled where on each swap settlement date oneside pays a fixed amount and the other side pays the average price the commodity forward had over a predetermined set of averaging dates. There is usually one net payment.

Example: Cal-07 swap on WTI front contract during each US business day, monthly settled. If U is a set of business days, $F(t,T)$ the forward price at time t for a forward with expiry T, and t_u denote the swap settlement dates (end of each month here), and $D(t,t_u)$ the corresponding discount factor, then the swap price is given by

$$R_s(t) = \frac{\sum_u w_u D(t,t_u) F(t,T_u)}{\sum_u w_u D(t,t_u)}$$

Exotic swaps: Participating swaps, Basis swaps which involve two or more underlying assets (cracks, calendar spread, refinery margins, heat-rate), Total or Excess Return swaps on Commodity indices, Swaps on customised indices and publications (e.g. achieved export or import prices), Swaps paid in different currencies than the underlying commodity pricing, Swaps with inflation linked payments.

Options

European Options: Black-Scholes pricing.

American options: trade on futures above. At expiry they settle into futures. They are usually priced with on a binomial or trinomial tree.

Average Price Options (Asians, APOs) on the front nearby monthly forward: for a given averaging period consider a set U of averaging (usually business) days t_i , $i = 1, ..., n$, and $F(t,T)$ denote the forward price at time t for a contract maturing at time T. For each fixing date, let T_i be the expiration date of the front contract for the date i. Then a call option has a payoff:

$$A_{APO} = (A(T) - K)^+ \qquad\qquad A(T) = \frac{1}{n}\sum_{i=1}^{n} F(t_i, T_i)$$

There are both analytical (Turnbull-Wakeman, Curran and Rogers-Shi, Geman-Yor), as well as numerical methods (Monte Carlo and Tree) for pricing APOs on one underlying contract as well as numerical methods.

However, the above APO is on a option on a basket of underlying forward contracts. For example an annually settled APO involves 12 futures/forwards, all of which are correlated.

An appropriate pricing model needs to take this correlation into account.

Swaptions:

These are options on a swap as defined previously. Again this is an option on a basket of underlying forward contracts.

Spread Options:

In the petroleum industry, refinery managers are more concerned about the difference between their input and output prices than about the level of prices. Refiners profits are tied directly to the spread, or difference, between the price of crude oil and the prices of refined products. Because refiners can reliably predict their costs other than crude oil, the spread is their major uncertainty. NYMEX in 1994 launched the crack spread contract. One type of crack spread contract bundles the purchase of three crude oil futures (30,000 barrels) with the sale a month later of two unleaded gasoline futures (20,000 barrels) and one heating oil future (10,000 barrels). The 3-2-1 ratio approximates the real-world ratio of refinery output2 barrels of unleaded gasoline and 1 barrel of heating oil from 3 barrels of crude oil.

Calendar Spread Options Storage facilities play an important role in the crude oil and refining supply chain. Heating oil dealers build inventories during the summer and fall for winter delivery. For most non-energy commodities, the cost of storage is one of the key determinants of the differential between current and future prices. Although storage plays a smaller role in price determination in some energy markets (most notably, for electricity), it can be important for heating oil and natural gas. If the price differential between winter months and shoulder months substantially exceeds storage expenses, traders can buy and store gas and sell gas futures. Such arbitrage tends to narrow the price differential.

NYMEX offers calendar spread options on crude oil, heating oil, and unleaded gasoline. Buying a call on the calendar spread options contract will represent a long position (purchase) in the prompt months of the futures contract and a short position (sale) in the further months of the contract.

Thus, the storage facility can buy a call on a calendar spread that will allow it to lock in a storage profit or to arbitrage a spread that is larger than its cost of storage.

Spark Spread options (options on the spread between Power and Gas)

Dark spread options (options on the spread between Power and Coal)

$$C_{SP} = (S_1(T) - S_2(T) - K)^+$$

Barrier option, Outside barrier options on basket of refined products with the barrier (usually more than one) based on crude prices.

Swing options

Swing options are volumetric options. Typical of gas, but also seen in Power markets. Assume an exercise period $[0, T]$ and a strike k (GBp/them or USD/mmBTU). On each day in the $[0, T]$ period the option holder has the right to recall a quantity $q(t)$ such that:

$$V_{min} \leq q(t) \leq V_{max} \qquad\qquad A \leq \sum_{k \in [0,T]} q(t) \leq B$$

where A, B represent the allowed min and max volumes during the entire duration of the trade. If the buyer (option holder) takes more or less volume then he/she pays a penalty. This is why it is also called take-or-pay options. The number of swing rights can be less than or equal to the total number of days in the period N.

Electricity Derivatives

Forward and Futures

Electricity forward contracts represent the obligation to buy or sell a fixed amount of electricity at a pre-specified contract price, known as the forward price, at certain time in the future (called maturity or expiration time). In other words, electricity forwards are custom tailored supply contracts between a buyer and a seller, where the buyer is obligated to take power the seller is obligated to supply.

Generators such as independent power producers (IPPs) are the natural sellers (or, shortside) of electricity forwards while LSEs such as utility companies often appear as the buyers (or, long-side). The maturity of an electricity forward contract ranges from hours to years although contracts with maturity beyond two years are not liquidly traded.

Some electricity forwards are purely financial contracts, which are settled through financial payments based on certain market price index at maturity, while the rest are physical contracts as they are settled through physical delivery of underlying electricity.

The payoff of a forward contract promising to deliver one unit of electricity at price K at a future time T is:

$$S_T - K$$

where S_T is the electricity spot price at time T.

Although the payoff function appears to be same as for any financial forwards, electricity forwards differ from other financial and commodity forward contracts in that the underlying electricity is a different commodity at different times. The settlement price S_T is usually calculated based on the average price of electricity over the delivery period at the maturity time T.

First traded on the NYMEX in March 1996, electricity futures contracts have the same payoff structure as electricity forwards. However, electricity futures contracts, like other financial futures contracts, are highly standardized in contract specifications, trading locations, transaction requirements, and settlement procedures. The most notable difference between the specifications of electricity futures and those of forwards is the quantity of power to be delivered. Usually settled in cash. The delivery quantity specified in electricity futures contracts is often significantly smaller than that in forward contracts

Swaps

Electricity swaps are financial contracts that enable their holders to pay a fixed price for underlying electricity, regardless of the floating electricity price, or vise versa, over the contracted time period. Electricity swaps are widely used in providing short- to medium-term price certainty up to a couple of years. Electricity locational basis swaps are also commonly used to lock in a fixed price at a geographic location that is different from the delivery point of a futures contract. That is, a holder of an electricity locational basis swap agrees to either pay or receive the difference between a specified futures contract price and another locational spot price of interest for a fixed constant cash flow at the time of the transaction. These swaps are effective financial instruments for hedging the basis risk on the price difference between power prices at two different physical locations.

Plain call and put options

Electricity call and put options offer their purchasers the right, but not the obligation, to buy or sell a fixed amount of underlying electricity at a pre-specified strike price by the option expiration time. They have similar payoff structures as those of regular call and put options on financial securities and other commodities.

The underlying of electricity call and put options can be exchange-traded electricity futures or physical electricity delivered at major power transmission inter-ties, like the ones located at California-Oregon Border and Palo Verde in the Western U.S. power grid. The majority of the transactions for electricity call and put options occur in the OTC markets.

Spark spread options

An important class of non-standard electricity options is the spark spread option (or, spark spread). Spark spreads are cross-commodity options paying out the difference between the price of electricity sold by generators and the price of the fuels used to generate it. The amount of fuel that a generation asset requires to produce one unit of electricity depends on the asset's fuel efficiency or heat rate (Btu/kWh). The holder of a European- spark spread call option written on fuel G at a fixed heat rate K_H has the right, but not the obligation, to pay at the option's maturity K_H times the fuel price at maturity time T and receive the price of one unit of electricity. Thus, the payoff at maturity time T is

$$[S_T - K_H G_T]^+$$

where S_T and G_T are the electricity and fuel prices at time T, respectively.

Swing options

Electricity swing options are adopted from their well-known counterparts in the natural gas industry. Also known as flexible nomination options, swing options have the following defining features. First, these options may be exercised daily or up to a limited number of days during the period in which exercise is allowed. Second, when exercising a swing option, the daily quantity may vary (or, swing) between a minimum daily volume and a maximum volume.

However, the total quantity taken during a time period such as a week or a month needs to be within certain minimum and maximum volume levels. Third, the strike price of a swing option may be either fixed throughout its life or set at the beginning of each time period based on some pre-specified formula. Last, if the minimum-take quantity of any contract period is missed by the buyer, then a lump sum penalty or a payment making up the sellers revenue shortfall needs to be paid (i.e., take-or-pay).

Pricing energy derivatives

Option valuation is held by binomial trees, MonteCarlo methods, risk-neutral prcing

Commodities (and hence price processes) are different from financial assets in that:

- Electricity is non-storable, and agricultural products are fungible.
- Operational constraints and market structure does not allow to capture arbitrage efficiently
- Seasonality (Heating oil, Natural Gas, even intraday power prices)
- Liquidity constraints: occasional squeezes, outages, or even geopolitical factors affect directly commodity prices. These affect in turn our assumptions of commodity price processes (lognormal returns, jumps, etc).
- Even though financial instruments are linked to the underlying physical assets, this can give a distorted picture: the size of the financial contracts trading in NYMEX today is 1000 times larger than the physical transactions per day.

Forward price

$$F(t,T) = S(t)e^{(r+u-y)(T-t)}$$

where u is storage costs rate and y is the convenience yield. If y was deterministic forward prices could be modeled as martingales.

This remarkable relationship allows to interpret the convenience yield as a continuous dividend payment made to the owner of the commodity. Hence, under the additional assumption that the price of the underlying commodity is driven by a geometric Brownian motion, Mertons (1973) formula for options on dividend-paying stocks, with dividedn yield q, provides the price of a plain vanilla call option written on a commodity with price S, namely

$$C(t) = S(t)e^{-q(T-t)}N(d_1) - Ke^{-r(T-t)}N(d_2)$$

Pricing power options

In the case of Asian options which represent the majority of options written on oil (since most indices on oil are defined as arithmetic averages over a given period), it is well known that the valuation problem becomes much more difficult and several approximations for the call price have been offered in the literature. Geman-Yor (1993) were able, using stochastic time changes and Bessel processes, to provide the Laplace transform of the exact price of the Asian option. Eydeland-Geman (1995) inverted this Laplace transform and showed the superiority of this approach over Monte Carlo simulations, in particular in terms of hedging accuracy. These results were established under the general assumptions of dividend payments for stocks or convenience yield for commodities.

Pricing electricity derivatives

Since the value of electricity derivatives are based on the underlying electricity prices, modeling electricity price is the most critical component in pricing electricity derivatives. Due to the unique physical and operational characteristics of electricity production and transmission processes, electricity price exhibits different behaviors than other financial prices which can be often described by Geometric Brownian Motion. There has been a growing literature addressing mainly two competing approaches to the problem of modeling electricity price processes:

(a) "Fundamental approach" that relies on simulation of system and market operation to arrive at market prices; and

(b) "Technical approach" that attempts to model directly the stochastic behavior of market prices from historical data and statistical analysis.

While the first approach provides more realistic system and transmission network modeling under specific scenarios, it is computationally prohibitive due to the large number of scenarios that must be considered. Therefore, we shall focus our attentions on the second approach and review the corresponding methodologies for pricing electricity derivatives.

Approaches to characterize market prices include discrete-time time series models such as GARCH and its variants, Markov regime-switching models, continuous-time diffusion models such as mean-reversion, jump-diffusion, and other diffusion models. There are also models proposed for direct modeling of electricity forward curves.

a) According to the definition, convenience yield is the difference between two quantities: the positive return from owning the commodity for delivery and the cost of storage. Because of the impossibility of storing power, these two quantities cannot be specified.

b) The non-storability of electricity leads to the breakdown of the relationship which prevails at equilibrium between spot and future prices on stocks, equity indices, currencies, etc. The no arbitrage argument used to establish this relation is not valid in the case of power, since it requires that the underlying instrument be bought at time t and held until the expiration of the futures contract.

c) There is another important consequence of non-storability: using the spot price evolution models for pricing power options is not very helpful, since hedges involving the underlying asset, i.e., the famous delta hedging, cannot be implemented, as they require buying and holding power for a certain period of time.

In the classical Black-Scholes-Merton world, the key quantity in the option valuation is the spot price of the underlying asset, since it provides, in particular, the hedging portfolio. Hence, the first and important step is to model the spot price dynamics. In the case of electricity, the main problems that one faces while modeling spot dynamics are the difficult issues of matching fat tails of marginal and conditional distributions and the spikes in spot prices. There are a number of techniques addressing these issues; below, we describe two models that appear most relevant.

i) A diffusion process with stochastic volatility, namely

$$dS_t = \mu_1(t, S_t)dt + \sigma_t S_t dW_t^1$$

$$d\Sigma = \mu_2(t, \Sigma_t)dt + y(t, \Sigma_t)dW_t^2$$

where $\Sigma_t = \sigma_t^2$, W_t^1, W_t^2 are two Brownian motions, with a correlation coefficient $\rho(t)$, and the terms $\mu^1(t, S_t)$ and $\mu^2(t, \Sigma_t)$ may account for some mean reversion either in the spot prices or in the spot price volatility.

Stochastic volatility is certainly necessary if we want a diffusion representation to be compatible with the extreme spikes as well as the fat tails displayed by distribution of realized power prices. However, stochastic volatility puts us in a situation of incomplete markets since we only have one instrument, the spot power (or rather its surrogate) to hedge the option. Hence the risk-neutral valuation techniques are problematic even if widely used.

ii) Because of extreme temperatures, and hence, an extreme power demand, the dynamics of electricity spot prices can be advantageously represented by a jump-diffusion process (see Geman 1994). The simplest jump process one can add to the diffusion component is represented by a Poisson process N_t with a random magnitude, and we thus obtain the Merton (1976) model

$$dS_t = \mu S_T dt + \sigma S_t dW_t + U S_t dN_t$$

where N_t is a Poisson process whose intensity λ characterizes the frequency of occurrence of the jumps, while U is a real-valued random variable, for instance normal, which represents the direction and magnitude of the jump. With one tradable risky asset to hedge the sources of randomness represented by W_t , N_t and its random multiplier U, we face an extreme situation of market incompleteness. Again, risk-neutral valuation techniques becomes problematic even if widely used.

Pricing Asian Options

Asian options have a payoff which depends on the average value of the underlying asset over the maturity of the options. We have two cases

- $h_T = (S_T - \bar{S}_T)^+$
- $h_T = (\bar{S}_T - K)^+$

The average can be defined as arithmetic or geometric and in continuous or discrete time:

- Arithmetic Average
  discrete time:
$$A_T^D = \frac{1}{N}\sum_{i=1}^{N} S_{t_i}$$
  continuous time:
$$A_T^C = \frac{1}{T}\int_0^T S_t dt$$
- Geometric Average
  discrete time:
$$G_T^D = (\prod_{i=1}^{N} S_{t_i})^{1/N}$$
  continuous time:
$$G_T^C = e^{\frac{1}{T}\int_0^T \ln S_t dt}$$

Let's do the geometric case with continuous sampling.

Fix strike
$$S_t = S_0 e^{(r-1/2\sigma^2)t+\sigma W_t^*} = S_0 e^{\nu t+\sigma W_t^*}$$
where $\nu = r - 1/2\sigma^2$. Hence
$$G_T^C = e^{\frac{1}{T}\int_0^T \ln S_t dt} = S_0 e^{\frac{1}{T}\int_0^T \nu t dt + \frac{\sigma}{T}\int_0^T W_t^* dt} =$$
$$S_0 e^{\frac{1}{T}\nu\frac{T^2}{2}+\frac{\sigma}{T}\int_0^T W_t^* dt} = S_0 e^{\frac{\nu T}{2}+\frac{\sigma}{T}\int_0^T W_t^* dt}$$
But using integration by part formula
$$\int_0^T W_t^* dt = TW_T^* - \int_0^T t dW^* = \int_0^T (T-t) dW_t^* = \mathcal{I}_T(T-t)$$

Define
$$\mathcal{I}_T(X) = \int_0^T X_s dW_s$$
The stochastic integral $\mathcal{I}$ is a martingale if $E[\int_0^T X_t^2 dt] < \infty$ and

- $$E[\mathcal{I}_T(X)|\mathcal{F}_s] = \mathcal{I}_s(X)$$

  and

  $$E[\mathcal{I}_T(X)] = E[\mathcal{I}_0(X)] = 0$$

- Itô Isometry:

  $$E[\mathcal{I}_T^2(X)] = E[\int_0^T X^2(s)ds]$$

  if X_s is a <u>deterministic</u> function.

But then:

$$var[\mathcal{I}_T(T-t)] = E[\mathcal{I}_T^2(T-t)] = E[\int_0^t (T-t)^2 dt] = E[\int_0^T (T^2 - 2Tt + t^2)dt] =$$

$$T^3 - \frac{2T^3}{2} + \frac{T^3}{3} = \frac{T^3}{3}$$

Hence

$$\int_0^t W_s ds \sim N\left(0, \frac{t^3}{3}\right)$$

and the geometric average of a log-normal process S_t is a lognormal process and

$$\ln \frac{G_T^C}{G_0^C} \sim N\left(\frac{\nu T}{2}, \frac{\sigma}{T} var\mathcal{I}_T(T-t)\right)$$

$$\ln \frac{G_t^C}{G_0^C} \sim N\left(\nu \frac{t}{2}, \sigma^2 \frac{t}{3}\right)$$

$$G_t^C = G_0^C e^{\frac{\nu t}{2} + \frac{\sigma}{\sqrt{3}} W_t^G} = G_0^C e^{(r-q-\frac{\sigma^2}{6})t + \frac{\sigma}{\sqrt{3}} W_t^G}$$

Imposing

$$r - q - \frac{\sigma^2}{6} = \frac{1}{2}(r - 1/2\sigma^2)$$

we find

$$q = r - \frac{1}{2}(r - 1/2\sigma^2) - \frac{\sigma^2}{6} = \frac{1}{2}(r + \frac{\sigma^2}{6})$$

Then to price a continuous geometric average, fix-strike Asian Option we can apply the standard Black-Scholes formula and just replace the right dividend and volatility.

$$C(0) = S(0)e^{-qT}N(d_1) - Ke^{-rT}N(d_2)$$

with

$$d_1 = \frac{\ln(S(0)/K) + (r - q + 1/2\sigma_G^2)T}{\sigma_G\sqrt{T}} = \frac{\ln(S(0)/K) + (r - 1/2\sigma^2)T/2}{\sigma\sqrt{T/3}}$$

and

$$d_2 = \frac{\ln(S(0)/K) + (r - 1/2\sigma_G^2)T}{\sigma_G\sqrt{T}} = \frac{\ln(S(0)/K) + (r - 1/2\sigma^2)T/2 + \sigma^2 T/3}{\sigma\sqrt{T/3}}$$

Multi dimensional Brownian motion

In the following sections we will analyze option contract depending on more then one underlying asset. To avoid perfect correlation between the assets, the underlying noise process should be modeled by means of a multidimensional Brownian motion. Let's the price processes for two assets X_t and Y_t be given, under the risk neutral measure P^*, by

$$dX_t = X_t(\mu_X dt + \sigma_X dW_t^X)$$
$$dY_t = Y_t(\mu_Y dt + \sigma_Y dW_t^Y)$$

such that

$$E[dW_t^X dW_t^X] = dt$$
$$E[dW_t^Y dW_t^Y] = dt$$
$$E[dW_t^X dW_y^Y] = \rho dt$$

and

$$E[\frac{dX}{X}\frac{dX}{X}] = \sigma_X^2 dt$$
$$E[\frac{dY}{Y}\frac{dY}{Y}] = \sigma_Y^2 dt$$
$$E[\frac{dX}{X}\frac{dY}{Y}] = \sigma_X \sigma_Y \rho dt$$

We want to rewrite X and Y is terms of a two-dimensional Brownian motion

$$\bar{W} = (W_1, W_2)$$

with

$$E[dW^1 dW^1] = dt$$
$$E[dW^2 dW^2] = dt$$
$$E[dW^1 dW^2] = 0$$

So

$$dX_t = X_t(\mu_X dt + \sigma_{11} dW_t^1 + \sigma_{12} dW_t^2)$$
$$dY_t = Y_t(\mu_Y dt + \sigma_{21} dW_t^1 + \sigma_{22} dW_t^2)$$

In this case

$$E[\frac{dX}{X}\frac{dX}{X}] = E[(\sigma_{11} dW_t^1 + \sigma_{12} dW_t^2)(\sigma_{11} dW_t^1 + \sigma_{12} dW_t^2)] =$$

$$\sigma_{11}^2 E[(dW_t^1)^2] + \sigma_{12}^2 E[(dW_t^2)^2] = (\sigma_{11}^2 + \sigma_{12}^2)dt$$

from where

$$\sigma_{11}^2 + \sigma_{12}^2 = \sigma_X^2$$

$$E[\frac{dY}{Y}\frac{dY}{Y}] = E[(\sigma_{21}dW_t^1 + \sigma_{22}dW_t^2)(\sigma_{21}dW_t^1 + \sigma_{22}dW_t^2)] =$$
$$\sigma_{21}^2 E[(dW_t^1)^2] + \sigma_{22}^2 E[(dW_t^2)^2] = (\sigma_{21}^2 + \sigma_{22}^2)dt$$

from where

$$\sigma_{21}^2 + \sigma_{22}^2 = \sigma_Y^2$$

$$E[\frac{dX}{X}, \frac{dY}{Y}] = E[(\sigma_{11}dW_t^1 + \sigma_{12}dW_t^2)(\sigma_{21}dW_t^1 + \sigma_{22}dW_t^2)] =$$
$$\sigma_{11}\sigma_{21}E[(dW_t^1)^2] + \sigma_{12}\sigma_{22}E[(dW_t^2)^2] = (\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22})dt$$

from where

$$(\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22}) = \rho\sigma_X\sigma_Y$$

We need to solve the set of three equations in four variables.

One simple choice is

$$\sigma_{12} = 0 \qquad\qquad \sigma_{11} = \sigma_X$$

Then we get

$$\sigma_{21} = \rho\sigma_Y \qquad\qquad \sigma_{22} = \sigma_Y\sqrt{1-\rho^2}$$

We can then write

$$dX_t = X_t(\mu_X dt + \bar{\sigma}_X \cdot d\bar{W}_t)$$
$$dY_t = Y_t(\mu_Y dt + \bar{\sigma}_Y \cdot d\bar{W}_t)$$

where $\bar{\sigma}_X = (\sigma_X, 0)$, $\bar{\sigma}_Y = (\sigma_Y\rho, \sigma_Y\sqrt{1-\rho^2})$ and $\bar{W}_t$ is a bi-dimensional Brownian motion.

Let's find the risk neutral measure. We have to find the vector of market price of risk $\theta = (\theta_1, \theta_2)$

$$r - \mu_X = \theta_1\sigma_X$$
$$r - \mu_Y = \rho\sigma_Y\theta_1 + \sigma_Y\sqrt{1-\rho^2}\theta_2$$

which gives

$$\theta_1 = \frac{r-\mu_X}{\sigma_X}$$
$$\theta_2 = \frac{(r-\mu_Y)\sigma_X + (\mu_X - r)\sigma_Y\rho}{\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

and from Girsanov

$$W_t^{*1} = W_t^1 - \theta_1 t$$
$$W_t^{*2} = W_t^2 - \theta_2 t$$

We can extend this formalism to d-dimentional brownian motions.

$$\frac{dS_t}{S_t} = \mu dt + \sigma \cdot dW_t$$

where $\mu \in R$ is a constant drift and $\sigma \in R^d$ denotes a constant volatility vector and W is a d-dimensional Brownian motion:

$$\sigma \cdot dW_t = \sum_{i=1}^{d} \sigma_i dW_t^i$$

$$|\sigma|^2 = \sum_{i=1}^{d} \sigma_i^2$$

Then

$$S_t = S_0 e^{(\mu - 1/2|\sigma|^2)t + \sigma \cdot W_t}$$

In this case one finds that the vector of market price of risk is given by solution of the equation

$$\mu - r + \sigma \cdot \theta = 0$$

The solution of this equation is in general not unique unless the number of traded assets, including the domestic saving account, equal $d + 1$. For any solution of this equation a martingale measure can be defined by

$$\frac{dP^*}{dP} = e^{\theta \cdot W_T - \frac{1}{2}|\theta|^2 T}$$

and, using Girsanov,

$$W_t^* = W_t - \theta t$$

where W_t^* follows a d-dimensional Brownian under P^* and

$$S_t = S_0 e^{(r - 1/2|\sigma|^2)t + \sigma \cdot W_t^*}$$

Exchange Options and Spread Options

The payoff of this contract is:

$$h = (Y_T - X_T - S)^+$$

Let the price processes for two assets X_t and Y_t be given, under the risk neutral measure P^*, by

$$dX_t = X_t(rdt + \bar{\sigma}_X \cdot d\bar{W}_t^*)$$
$$dY_t = Y_t(rdt + \bar{\sigma}_Y \cdot d\bar{W}_t^*)$$

where $\bar{\sigma}_X = (\sigma_X, 0)$, $\sigma_Y = (\sigma_Y \rho, \sigma_Y \sqrt{1-\rho^2})$ and $\bar{W}_t$ is a bi-dimensional Brownian motion.

The case $S = 0$ is much easier and we focus on this.

Solving we find

$$X_t = X_0 e^{(r-1/2|\bar{\sigma}_X|^2)t + \bar{\sigma}_X \cdot \bar{W}_t^*}$$
$$Y_t = Y_0 e^{(r-1/2|\bar{\sigma}_Y|^2)t + \bar{\sigma}_Y \cdot \bar{W}_t^*}$$

Hence

$$\frac{Y_t}{X_t} = \frac{Y_0}{X_0} e^{-1/2(|\bar{\sigma}_Y|^2 - |\bar{\sigma}_X|^2)t + (\bar{\sigma}_Y - \bar{\sigma}_X) \cdot \bar{W}_t^*}$$

Let's call $\bar{\sigma}_{Y/X} = (\bar{\sigma}_Y - \bar{\sigma}_X) = (\rho\sigma_Y - \sigma_X, \sigma_Y\sqrt{1-\rho^2})$.

$$\frac{Y_t}{X_t} = \frac{Y_0}{X_0} e^{-1/2(|\bar{\sigma}_Y|^2 - |\bar{\sigma}_X|^2)t + \bar{\sigma}_{Y/X} \cdot \bar{W}_t^*}$$

Note that

$$|\bar{\sigma}_X|^2 = \sigma_X^2$$
$$|\bar{\sigma}_Y|^2 = \sigma_Y^2$$
$$|\bar{\sigma}_{Y/X}|^2 = \sigma_X^2 + \sigma_Y^2 - 2\rho\sigma_X\sigma_Y$$

Define $\beta(t) = e^{rt}$ and choose X as a numeraire. This defines the equivalent measure Q as

$$\frac{dQ}{dP^*} = \frac{X_T\beta(0)}{X_0\beta(T)} = e^{-1/2|\bar{\sigma}_X|^2 + \bar{\sigma}_X \cdot \bar{W}_t^*}$$

This shows that the market price of risk $\bar{\theta} = \bar{\sigma}_X = (\sigma_X, 0)$

Under Q

$$\bar{W}_t = \bar{W}_t^* - \bar{\theta}t = (W_1^* - \sigma_X t, W_2^*)$$

is a standard Brownian motion.

Hence under Q

$$\frac{Y_t}{X_t} = \frac{Y_0}{X_0} e^{-1/2(|\bar{\sigma}_Y|^2 - |\bar{\sigma}_X|^2)t + (\rho\sigma_Y - \sigma_X)(W_{1t} + \sigma_X t) + \sqrt{1-\rho^2}\sigma_Y W_{2t}} =$$

$$\frac{Y_0}{X_0} e^{-1/2(|\bar{\sigma}_Y|^2 + |\bar{\sigma}_X|^2 - 2\rho\sigma_X\sigma_Y)t + (\rho\sigma_Y - \sigma_X)W_{1t} + \sqrt{1-\rho^2}\sigma_Y W_{2t}} =$$

$$\frac{Y_0}{X_0}\ e^{-1/2\ (\sigma_Y^2 + \sigma_X^2 - 2\rho\sigma_X\sigma_Y)t\ +\ \bar{\sigma}_{Y/X}\cdot\bar{W}_t} = \frac{Y_0}{X_0}\ e^{-1/2\ |\sigma_{Y/X}|^2 t\ +\ \bar{\sigma}_{Y/X}\cdot\bar{W}_t}$$

Notice that $\frac{Y_t}{X_t}$ is now a martingale under Q.

Let's now calculate the price of an exchange option whose payoff is

$$C_T = (Y_T - X_T)^+.$$

Define $\beta(t) = \exp(rt)$. Choosing X as a numeraire the price of the option can be written as

$$C_0 = \beta(0)E_{P^*}[(Y_T - X_T)^+/\beta(T)] = X_0 E_Q[(Y_T - X_T)^+/X_T] =$$

$$X_0 E_Q[(Y_T/X_T - 1)I_{\{Y_T > X_T\}}] = X_0 E_Q[Y_T/X_T I_{\{Y_T > X_T\}}] - X_0 Q\{Y_T > X_T\} =$$

$$X_0 E_Q[Y_0/X_0 e^{-1/2|\bar{\sigma}_{Y/X}|^2 T + \bar{\sigma}_{Y/X}\cdot\bar{W}_T} I_{\{Y_T > X_T\}}] - X_0 Q\{Y_T > X_T\}$$

To calculate first integral we have to change measure to $\hat{Q}$ defined by:

$$\frac{d\hat{Q}}{dQ} = \eta_T = e^{-1/2|\bar{\sigma}_{Y/X}|^2 T + \bar{\sigma}_{Y/X}\cdot\bar{W}_T}$$

Under $\hat{Q}$

$$\hat{\bar{W}}_t = \bar{W}_t - \bar{\sigma}_{Y/X} t$$

is a standard Brownian motion.

We eventually obtain

$$C_0 = Y_0\hat{Q}\{Y_T/X_T > 1\} - X_0 Q\{Y_T/X_T > 1\} =$$

$$Y_0\hat{Q}\{Y_0/X_0 e^{1/2|\bar{\sigma}^2_{Y/X}|T + \bar{\sigma}_{Y/X}\cdot\hat{\bar{W}}_T} > 1\} - X_0 Q\{Y_0/X_0 e^{-1/2|\bar{\sigma}^2_{Y/X}|T + \bar{\sigma}_{Y/X}\cdot\bar{W}_T} > 1\} =$$

$$Y_0\hat{Q}\{1/2|\bar{\sigma}^2_{Y/X}|T + \bar{\sigma}_{Y/X}\cdot\hat{\bar{W}}_T > \ln(X_0/Y_0)\} - X_0 Q\{-1/2|\bar{\sigma}^2_{Y/X}|T + \bar{\sigma}_{Y/X}\cdot\bar{W}_T > \ln(X_0/Y_0)\} =$$

$$= Y_0 N(d_1) - X_0 N(d_2)$$

with

$$d_1 = \frac{\ln(Y_0/X_0) + 1/2|\bar{\sigma}_{Y/X}|^2 T}{|\sigma_{Y/X}|\sqrt{T}}$$

$$d_2 = \frac{\ln(Y_0/X_0) - 1/2|\bar{\sigma}_{Y/X}|^2 T}{|\sigma_{Y/X}|\sqrt{T}}$$

Notice that, as expected, this is just the Black-Scholes formula, with $r = 0$, strike $K = 1$, and volatility $\bar{\sigma}_{Y/X}$.

With dividends the formula becomes

$$C_0 = Y_0 e^{-q_Y T} N(d_1) - X_0 e^{-q_X T} N(d_2)$$

with

$$d_1 = \frac{\ln(Y_0/X_0) + (q_X - q_Y + 1/2|\bar{\sigma}_{Y/X}|^2)T}{|\sigma_{Y/X}|\sqrt{T}}$$

$$d_2 = \frac{\ln(Y_0/X_0) + (q_X - q_Y - 1/2|\bar{\sigma}_{Y/X}|^2)T}{|\sigma_{Y/X}|\sqrt{T}}$$

Modeling Forward Rates

When building an energy model, there are two main approaches for arriving at forward prices. One can either model the spot prices and work out the forward prices from the spot; or one can model the forward curve dynamics and derive spot prices from the forward curve. This distinction is equivalent to the distinction between interest rate short models and Heath-Jarrow-Morton models. The following sections will look at various mean reverting models under these two main approaches.

The single factor Schwartz model

In his 1997 paper "Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging", Eduardo Schwartz considers different approaches for modeling the commodity spot price. First, he considers a single factor model based on a standard, mean reverting Ornstein-Uhlenbeck process. The second model takes a standard asset price process and introduces a mean reverting convenience yield. Finally, Schwartz takes the two factor process and introduces stochastic interest rates. What makes these models appealing are their analytical tractability. Unfortunately the convenience yield is not transparent, leading to a disparity between the model and the market. Another large drawback of this approach is the inability to fit market observable forward data to the model.

The single factor spot price model is built on a probability space, $(\Omega, \mathcal{F}, \mathcal{Q}_T)$, equipped with the natural filtration, $\{\mathcal{F}_t\}$, of the standard one-dimensional Brownian motion z. Throughout this model, interest rates are assumed to follow a deterministic process which allows futures and forwards to be treated equivalently.

The single factor model considered argues that the spot price dynamics follow an Ornstein-Uhlenbeck mean reverting, time homogeneous diffusion of the form,

$$\frac{dS(t)}{S(t)} = \alpha(\mu - \ln S)dt + \sigma dz(t). \tag{3}$$

Writing this equation in terms of the natural log, $X(t) = \ln S(t)$, and applying Itô's Lemma yields the dynamics of the log process

$$dX(t) = \alpha(\hat{\mu} - X(t))dt + \sigma dz(t), \tag{4}$$

where

$$\hat{\mu} = \mu - \frac{\sigma^2}{2\alpha}.$$

The remainder of this model will be expressed in the equivalent risk-neutral measure $\mathcal{Q}_T^*$. The dynamics of equation (4) are now expressed as

$$dX(t) = \alpha(\bar{\mu} - X(t))dt + \sigma dz^*(t), \tag{5}$$

where

$$\bar{\mu} = \hat{\mu} - \sigma\lambda.$$

Notice that equation (5) is simply the Vasicek model.

Repeating the calculations we have already done we can write

$$X(t) = e^{-\alpha t}X_0 + \bar{\mu}(1 - e^{-\alpha t}) + e^{-\alpha t}\int_0^t \sigma e^{\alpha s} dz^*(s). \tag{6}$$

The mean is given by

$$E^*[X(t)] = e^{-\alpha t}X_0 + \bar{\mu}(1 - e^{\alpha t}) \tag{7}$$

and the variance is given by

$$var[X(t)] = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}). \tag{8}$$

We can also compute the moment generating function of $X(t)$ and find

$$\mathcal{G}_\theta = E^*\left[e^{\theta X(t)}\right] = \exp\left(\theta(e^{-\alpha t}X_0 + \bar{\mu}(1 - e^{\alpha t})) + \frac{1}{2\alpha}\theta^2\sigma^2(1 - e^{-2\alpha t})\right). \tag{9}$$

That shows that X_t is a normal random variable.

In terms of the spot price, the process in equation (6) is expressed as

$$S(t) = \exp\left(e^{-\alpha t}X_0 + \bar{\mu}(1 - e^{-\alpha t}) + \int_0^t e^{-\alpha s}\sigma dz^*(s)\right). \tag{10}$$

Since we have assumed constant interest rates, it is straight forward to see that $F(t,T)$, the forward price at time $t \in [0,T]$ on the commodity $S(t)$ and maturity date T, is the expectation of the spot commodity under the risk neutral measure $\mathcal{Q}_T^*$.

$$F(t,T) = E^*[S(T)|\mathcal{F}_t] \quad t \in [0,T] \tag{11}$$

$$F(t,T) = \exp[e^{-\alpha(T-t)}\ln S + (1 - e^{-\alpha(T-t)})\bar{\mu} + \frac{\sigma^2}{4\alpha}(1 - e^{-2\alpha(T-t)})]. \tag{12}$$

This is represented in log form by

$$\ln F(t,T) = e^{-\alpha(T-t)}\ln S + (1 - e^{-\alpha(T-t)})\bar{\mu} + \frac{\sigma^2}{4\alpha}(1 - e^{-2\alpha(T-t)}). \tag{13}$$

This last equation may be used to model the forward curve and test the estimations against market data.

Plotting equation (13) with various mean reversion parameters reveals a potential shortcoming in the single factor Schwartz model. While this model is tractable, the lack of transparency in the market spot price process (referring to the actual spot, not the front month) creates some calibration difficulties. When attempting to fit the single factor Schwartz model to the market futures curve of WTI crude oil the model does not give a very realistic looking curve. It instead serves as a rough approximation of reality. In order to achieve a more realistic curve, different modeling approaches have been suggested.

The two factor model

The simplistic nature of the single factor model creates an unrealistic volatility structure. The negative exponential form implies that contracts with a long maturity will have zero (in the limiting sense) volatility. John Maynard Keynes would point out that this might not pose a problem since, "In the long run, we're all dead". However, the volatilities also decay at an unrealistically fast rate. These deficiencies lead to the introduction of a stochastic convenience yield that is correlated with the commodity price.

Forward curve models

A one factor model

The biggest disadvantage of spot price models is the fact that the spot is difficult to observe. Energies ia mainly purchased on the basis of a futures contract. It is also very difficult to fit the market observed futures curves to the endogenous futures price functions derived from spot models. These deficiencies lead to another group of models which investigate the dynamics of the entire futures curve. The spot price may then be interpreted as the futures curve price level at 0 maturity [1].

Empirical evidence suggests that the volatility of the forward price dynamics can be approximated by a negative exponential form. A good approximation for the overall volatility function can usually be found using the general function

$$\sigma(\alpha, t, T) = \sigma_1 e^{-\alpha(T-t)} + \sigma_2. \tag{14}$$

Unfortunately, this will lead to a non-Markovian pricing framework. In order to obtain a more tractable model, the negative exponential,

$$\sigma(\alpha, t, T) = \sigma e^{-\alpha(T-t)}, \tag{15}$$

suggested by the single factor Schwartz model will be used. One sees that equation (15) is a definite sacrifice. The market data implies that long-term contracts will have a non-zero, constant volatility. By contrast, equation (15) decays to zero in the long-run (seen by examining the limit). This gives the impression that longer maturity contracts are more stable than they actually are.

The expected drift of a forward rate is zero (remember forward rate behaves like asset that pay a dividend equal to the risk-free rate).

The observations discussed above lead to the following stochastic differential equation (SDE): [2]

$$\frac{dF(t,T)}{F(t,T)} = \sigma e^{-\alpha(T-t)} dz^*(t) \tag{16}$$

The mean reversion rate, α, describes how fast the spot volatility will decline to its long-term level (seen by taking the limit as maturity, T, tends toward infinity). This framework is called a "one factor model" since the only input needed to work out the forward curve is the single factor:

$$\sigma e^{-\alpha(T-t)}.$$

The log forward dynamics may be expressed as

$$d(\ln F(t,T)) = -\frac{1}{2}\sigma^2 e^{-2\alpha(T-t)} dt + \sigma e^{-\alpha(T-t)} dz^*(t). \tag{17}$$

[1] Throughout this section, the term futures and forward will be treated equivalently.

[2] Notice that the dynamics are defined in the risk-neutral measure $\mathcal{Q}_T^*$.

Integrating equation (17) yields

$$\ln F(t,T) = X_0 - \frac{1}{2}\int_0^t \sigma^2 e^{-2\alpha(T-u)}du + \int_0^t \sigma e^{-\alpha(T-u)}dz^*(u). \tag{18}$$

Which leads to the forward price as a function of the single factor volatility

$$F(t,T) = F(0,T)e^{(-\frac{1}{2}\int_0^t \sigma^2 e^{-2\alpha(T-u)}du + \int_0^t \sigma e^{-\alpha(T-u)}dz^*(u))}, \tag{19}$$

where X_0 is a constant such that $F(0,T) = e^{X_0}$. It can easily be verified that the forward price, $F(t,T)$, in this framework is indeed an $\{\mathcal{F}_t,\, \mathcal{Q}_T^*\}$-martingale. However, this will not necessarily be the case for any given volatility function, $\sigma(t,T)$.

By taking the limit as maturity approaches the present we arrive at the spot price,

$$S(t) = F(t,t) = F(0,t)e^{(-\frac{1}{2}\int_0^t \sigma^2 e^{-2\alpha(t-u)}du + \int_0^t \sigma e^{-\alpha(t-u)}dz^*(u))}. \tag{20}$$

and

$$\ln S(t) = \ln F(0,t) - \frac{1}{2}\int_0^t \sigma^2 e^{-2\alpha(t-u)}du + \int_0^t \sigma e^{-\alpha(t-u)}dz^*(u), \tag{21}$$

The dynamics of the spot price $S(t)$ are now calculated using Itô's Lemma;

$$d(\ln S(t)) = \left\{\frac{\partial \ln F(0,t)}{\partial t} - \int_0^t \sigma e^{-\alpha(t-u)}\frac{\partial(\sigma e^{-\alpha(t-u)})}{\partial t}du + \int_0^t \frac{\partial(\sigma e^{-\alpha(t-u)})}{\partial t}dz^*(u)\right\}dt + \sigma dz^*(t)$$

$$= \left\{\frac{\partial \ln F(0,t)}{\partial t} + \alpha\left[\int_0^t \sigma^2 e^{-2\alpha(t-u)}du - \int_0^t \sigma e^{-\alpha(t-u)}dz^*(u)\right]\right\}dt + \sigma dz^*(t). \tag{22}$$

From equation (21) we note that

$$\int_0^t \sigma e^{-\alpha(t-u)}dz^*(u) = \ln S(t) - \ln F(0,t) + \frac{1}{2}\int_0^t \sigma^2 e^{-2\alpha(t-u)}du \tag{23}$$

Equations (22) and (23) combine for the desired result.

$$d(\ln S(t)) = \{\frac{\partial \ln F(0,t)}{\partial t} + \alpha[\int_0^t \sigma^2 e^{-2\alpha(t-u)} - (\ln S(t) - \ln F(0,t) + \frac{1}{2}\int_0^t \sigma^2 e^{-2\alpha(t-u)}du)]\}dt + \sigma dz^*(t) \tag{24}$$

$$\frac{dS(t)}{S(t)} = \left\{\frac{\partial \ln F(0,t)}{\partial t} + \alpha[\ln F(0,t) - \ln S(t)] + \frac{\sigma^2}{4}(1 - e^{-2\alpha t})\right\}dt + \sigma dz^*(t) \tag{25}$$

From here it is straight forward to see how the one factor model is an extension of the Schwartz single factor model (which in hind sight should be described beforehand). Recall that the Schwartz single factor model is given by

$$\frac{dS(t)}{S(t)} = \alpha(\hat{\mu}(t) - \ln S(t))dt + \sigma dz^*(t) \tag{26}$$

The drift term in the Schwartz model, $\hat{\mu}$, in equation (3) is simply set as

$$\hat{\mu}(t) = \frac{1}{\alpha}\frac{\partial \ln F(0,t)}{\partial t} + \ln F(0,t) + \frac{1}{\alpha}\frac{\sigma^2}{4}(1 - e^{-2\alpha t}) \tag{27}$$

So far, the spot and forward curve prices and dynamics have all been derived from the forward return dynamics. As a final step, the forward curve evolution is fitted to the spot price so that the forward curve is consistent with the observed market data. This is seen by observing that equation (23) may be expressed as

$$\int_0^t \sigma e^{-\alpha(t-u)} dz^*(u) = \ln\left(\frac{S(t)}{F(0,T)}\right) + \frac{\sigma^2}{4\alpha}(1 - e^{-2\alpha t}),$$

and that the spot price $S(t)$ (in equation (20)) may be expressed as

$$F(t,T) = S(t) e^{\{-\frac{1}{2}\int_0^t \sigma^2 e^{-\alpha(T-t)} du + \int_0^t \sigma e^{-\alpha(T-t)} dz^*(u)\}}.$$

From equation (23) one notes that

$$\int_0^t \sigma e^{-\alpha(T-u)} dz^*(u) = e^{-\alpha(T-t)}\left[\ln\left(\frac{S(t)}{F(0,t)}\right) + \frac{\sigma^2}{4\alpha}(1 - e^{-2\alpha t})\right].$$

Combining these two equations, one arrives at a function for the forward price

$$F(t,T) = F(0,T)\left(\frac{S(t)}{F(0,t)}\right)^{\exp(-\alpha(T-t))} \cdot$$

$$\exp\left(-\frac{\sigma^2}{4\alpha} e^{-\alpha t}(e^{2\alpha t} - 1)(e^{-\alpha T} - e^{-\alpha t})\right). \tag{28}$$

The result shows that the forward prices can be obtained from the spot price (approximated by the current price of the front month contract), the volatility estimate, and mean reversion rate estimate.

These curves fit the market data much better than the single factor Schwartz curves found before.

The multi-factor forward curve model

The multi-factor forward curve model can be viewed as a generalization of the one factor model described in the previous section. The dynamics of the forward curve are defined by the SDE

$$\frac{dF(t,T)}{F(t,T)} = \sum_{i=1}^{n} \sigma_i(t,T) dz_i(t). \tag{29}$$

In this setup, $F(t,T)$ is the forward price at time t with maturity date T and $\sigma_i(t,T)$ is the volatility function associated with the Brownian motion $z_i^*(t)$. Furthermore, the n Brownian motions are assumed independent (i.e. $dz_i^*(t)dz_j^*(t) = 0$ for $i \neq j$) and interest rates are assumed to be deterministic so that futures and forward prices are the same.

As in the One Factor model, equation (29) is written in terms of the natural log and integrated to give $F(t,T)$. From equation (29), the dynamics of $\ln F(t,T)$ are expressed as

$$d(\ln F(t,T)) = \sum_{i=1}^{n} \left\{ -\frac{1}{2}\sigma_i^2(u,T)du + \sigma_i(u,T)dz_i^*(u) \right\}$$

Integrating both sides of this equation yields:

$$\ln F(t,T) - \ln F(0,T) = \sum_{i=1}^{n} \left\{ -\frac{1}{2}\int_0^t \sigma_i^2(u,T)du + \int_0^t \sigma_i(u,T)dz_i^*(u) \right\}.$$

Finally, the exponential is taken to arrive at the desired result,

$$F(t,T) = F(0,T)exp\left[\sum_{i=1}^{n} \left\{ -\frac{1}{2}\int_0^t \sigma_i^2(u,T)du + \int_0^t \sigma_i(u,T)dz_i^*(u) \right\} \right]. \tag{30}$$

As in the One Factor model, the spot price is expressed by setting the maturity T to t

$$S(t) = F(0,t)exp\left[\sum_{i=1}^{n} \left\{ -\frac{1}{2}\int_0^t \sigma_i^2(u,t)du + \int_0^t \sigma_i(u,t)dz_i^*(u) \right\} \right] \tag{31}$$

From equation (31) the natural logarithm of the spot price is computed,

$$\ln S(t) = \ln F(0,t) - \frac{1}{2}\sum_{i=1}^{n}\int_0^t \sigma_i^2(u,t)du + \sum_{i=1}^{n}\int_0^t \sigma_i(u,t)dz_i^*(u). \tag{32}$$

Since the explicit form of the functions $\sigma_i(t,T)$ have not been given, it is further assumed that they satisfy the Itô isometry such that

$$E^*\left[\left(\int_0^t \sigma_i(u,T)dz_i^*(u) \right)^2 \right] < \infty \quad \forall \quad t. \tag{33}$$

Where the volatility functions are square integrable martingales and hence, their expectations are 0. In this way equation (32) has mean

$$E^*[\ln S(t)] = \ln F(0,t) - \frac{1}{2}\sum_{i=1}^{n}\int_0^t \sigma_i^2(u,t)du, \tag{34}$$

and variance

$$Var(\ln S(t)) = E^*[(\ln S(t))^2] - E[\ln S(t)]^2$$

$$= E^* \left[\left\{ \ln F(0,t) - \frac{1}{2} \sum_{i=1}^{n} \int_0^t \sigma_i^2(u,t)du + \sum_{i=1}^{n} \int_0^t \sigma_i(u,t)dz_i^*(u) \right\}^2 \right]$$
$$- \left(\ln F(0,t) - \frac{1}{2} \sum_{i=1}^{n} \int_0^t \sigma_i^2(u,t)du \right)^2 \qquad (35)$$
$$= E^* \left[\left(\sum_{i=1}^{n} \int_0^t \sigma_i(u,t)dz_i^*(u) \right)^2 \right].$$

where we used the fact that the $z_i(t)$ are independent. Using Ito's isometry one sees that this yields

$$Var(S(t)) = \sum_{i=1}^{n} \int_0^t \sigma_i^2(u,t)du. \qquad (36)$$

The dynamics od $S$are finally given by

$$\frac{dS(t)}{S(t)} = \left[\frac{\partial F(0,t)}{\partial t} - \sum_{i=1}^{n} \int_0^t \sigma_i(u,t) \frac{\partial \sigma_i(u,t)}{\partial t} + \sum_{i=1}^{n} \int_0^t \frac{\partial \sigma_i(u,t)}{\partial t} dz_i^*(u) \right] dt$$
$$+ \left[\sum_{i=1}^{n} \sigma_i(t,t)dz_i^*(t) \right] \qquad (37)$$

Unlike the One Factor model in section (), the solution of the forward price SDE cannot necessarily be found as in equation (28). The choice of the volatility functions, $\sigma_i(t,T)$, completely determine the tractability of the multi-factor model. This choice allows one to also find equivalent models for the Schwartz single factor and two factor frameworks (Clewlow and Strickland 1999).

Weather Derivatives

Weather derivatives are usually structured as swaps, futures, and call/put options based on different underlying weather indices. Some commonly used indices are heating and cooling degree-days rain and snowfall. The degree-days indices are most often used. We start with some basic definitions and terminology.

Given a specific weather station, let T^i_{max} and T^i_{min} denote the maximal and minimal temperatures (in degrees Celsius) measured on day i. We define the temperature for day i as

$$T_i = \frac{T_{max} + T_{min}}{2}$$

The heating degree-days, HDD_i and the cooling degree-days, CDD_i, generated on that day are defined as

$$HDD_i = \max\{18 - T_i, 0\}$$

and

$$CDD_i = \max\{T_i - 18, 0\}$$

respectively.

It has become industry standard in the US to set this reference level at 65 Fahrenheit (18 C). The names heating and cooling degree days originate from the US energy sector. The reason is that if the temperature is below 18C people tend to use more energy to heat their homes, whereas if the temperature is above 18C people start turning their air conditioners on, for cooling.

Most temperature based weather derivatives are based on the accumulation of HDDs or CDDs during a ceratin period, usually one calender month or a winter/summer period. Typically the HDD season includes winter months from November to March and the CDD season is from May to September. April and October are often referred to as the shoulder months.

The CME offers trading with futures based on the CME Degree Day Index, which is the cumulative sum of daily HDDs or CDDs during a calendar month, as well as options on these futures. The CME Degree Day Index is currently specified for eleven US cities.

The notional value of one contract is \$100 times the Degree Day Index. The futures are cash-settled

A CME HDD or CDD call option is a contract which gives the owner the right, but not the obligation, to buy one HDD/CDD futures contract at a specific price, usually called the strike or exercise price. The HDD/CDD put option analogously gives the owner the right, but not the obligation, to sell one HDD/CDD futures contract. On the CME the options on futures are European style, which means that they can only be exercised at the expiration date.

Outside the CME there are a number of different contracts traded on the OTC market. One common type of contract is the option. There are two types of options, calls and puts. The buyer of a HDD call, for example, pays the seller a premium at the beginning of the contract. In return, if the number of HDDs for the contract period is greater than the predetermined strike level the buyer will recieve a payout. The size of the payout is determined by the strike and the tick size. The tick size is the amount of money that the holder of the call receives for each degree-day above the strike level for the period. Often the option has a cap on the maximum payout unlike, for example, traditional options on stocks.

A generic weather option can be formulated by specifying the following parameters:

- The contract type (call or put)
- The contract period (e.g. January 2001)
- The underlying index (HDD or CDD)
- An official weather station from which the temperature data are obtained
- The strike level
- The tick size
- The maximum payout (if there is any)

To find a formula for the payout of an option, let K denote the strike level and α the tick size. Let the contract period consist of n days. Then the number of HDDs and CDDs for that period are

$$H_n = \sum_{i=1}^{n} HDD_i$$

$$C_n = \sum_{i=1}^{n} CDD_i$$

respectively.

The payout of an uncapped HDD call is

$$X = \alpha \max(H_n - K, 0)$$

The payouts for similar contracts like HDD puts and CDD calls/puts are defined in the same way.

Swaps are contracts in which two parties exchange risks during a predetermined period of time. In most swaps, payments are made between the two parties, with one side paying a fixed price and the other side paying a variable price. In one type of weather swap that is often used, there is only one date when the cash-flows are swapped, as opposed to interest rate swaps, which usually have several swap dates. The swaps with only one period can therefore be thought of as forward contracts. Often the contract periods are single calendar months or a period such as January-March.

In the case of a standard HDD swap, the parties agree on a given strike of HDDs for the period, and the amount swapped is, for example, 10000 euro/HDD away from the strike. Usually there is also a maximum payout corresponding to 200 degree days.

The main difference between derivatives and insurance contracts is that the holder of an insurance contract has to prove that he has suffered a financial loss due to weather in order to be compensated. If he is not able to show this, the insurance company will not pay him any money. Payouts of weather derivatives are based only on the actual outcome of the weather, regardless of how it affects the holder of the derivative. One does not need to have any weather sensitive production, for example, to buy and benefit from a weather derivative. As any derivatives, these contracts can be bought for mere speculation.

Insurance contracts are usually designed to protect the holder from extreme weather events such as earthquakes and typhoons, and they do not work well with the uncertainties in normal weather. Weather derivatives, on the other hand, can be constructed to have payouts in any weather condition.

A customer may wish to buy a strip of CDD or HDD contracts spanning the entire cooling season of April through October. Each contract has a specified strike for each month. Each option is listed with a price for each range of strikes. In the OTC market, it is common for options to be written on a multimonth period with a single strike over the entire period. The strike is set relative to the normal climatological values. The normal value is a matter of debate. The market currently seems to be converging on the average over the past 10-15 years. Clearly there are many cases in which the 15- year average may not be ideal (global warming may be an issue).

Pricing weather options requires an historical temperature database and application of statistical methods for fitting distribution functions to data. Historical data is available from the National Oceanic and Atmospheric Administration (NOAA). The Midwestern Climate Center (MCC) has an online subscriber database that provides monthly-total CDDs and HDDs for all U.S. cities and provides a useful resource for those participating in the market. Note that the standard used to calculate CDDs and HDDs is different in the weather market than what is used by the atmospheric community such as the National Weather Service (NWS) and NOAA. NWS rounds the average daily temperature to the nearest degree for each day, whereas the weather market does not round temperature and keeps one decimal place in the daily CDDs or HDDs.

Compared with ordinary financial derivatives, temperature derivatives are very unique in terms of valuation. To start with, the underlying is a meteorological variable rather than a traded asset. The conventional risk-neutral valuation by no-arbitrage does not apply (in fact in practice this method is used!). In addition, being a meteorological variable, temperature follows a predictable trend, especially over a longer horizon. This is in sharp contrast with most of the financial variables. The predictability does not suggest arbitrage opportunities, though. Instead, it implies that any reasonable pricing model must incorporate this feature. The unique nature of the temperature variable brings about two key issues: accurate modeling of the underlying and the assessment of the market price of risk.

Similar to stochastic interest rate and stochastic volatility, temperature as a non-tradeable variable will also carry a market price for its risk. This market price of risk will in turn filter into the fair value of a temperature derivative. The academic literature only begins to make progress in valuing this new class of derivative securities (see, e.g., Cao and Wei [2003]). From the modeling perspective, the existing valuation methods can be loosely classified into three categories: 1) insurance or actuarial valuation, 2) historical burn analysis, and 3) valuation based on dynamic models.

Insurance or Actuarial Method As is well known, this methodology is widely used by companies specializing in property, automobile, health and life insurance, to name a few. Statistical analysis based on historical data is the backbone of this method. A probabilistic assessment is attached to the insured event and a fair premium is calculated accordingly. In the case of weather derivatives, this method is less applicable for most contracts since the underlying variables (e.g., temperature, wind and humidity) tend to follow a recurrent, predictable pattern. Nonetheless, if the contract is written on rare weather events such as extreme heat or coldness, then the actuarial method will be very useful. In fact, one may even argue that this is the only appropriate method in this case. For instance, using a diffusion process to model the temperature will be misguided if the main interest is in extreme events.

Historical Burn Analysis

This method is perhaps the simplest in terms of implementation, and as a result, is the most probable to cause large pricing errors. In a nutshell, the method of historical burn analysis evaluates the contract against historical data and takes the average of realized payoffs as the fair value estimate. For instance, suppose a call option is written on a citys CDDs for the month of July, and suppose we have 20 years of daily temperatures. To apply the historical burn analysis method, for each July of the past 20 years, we calculate the option payoff using the realized CDDs. The average of the 20 payoffs is the estimate for the call option value. Thus this methods key assumption is that, the past always reflects the future on average. To be more precise, the method assumes that the distribution of the past payoffs accurately depicts the future payoffs distribution.

If we have only 20 payoff observations in the above example, which can hardly capture the complete characteristics of the true distribution. It is tempting to argue that one should use as long a time series as possible to enhance accuracy. But this is a one-sided argument. Surely, using more data will cover more temperature variations. However, a derivative securitys payoff depends on the future temperature behavior, which may be quite different from history. This is especially so if the maturity of the derivative security is short. The commonly accepted sample length in the industry appears to be between 20 and 30 years. Furthermore, one could combine the burn analysis with temperature forecasts to arrive at a more representative price estimate.

Valuation Based on Dynamic Models of Temperature Behavior

In contrast to previous methods, a dynamic model directly simulates the future behavior of temperature. This is achieved by postulating a stochastic process for the temperature, which could be continuous or discrete. The continuous process usually takes the following mean-reversion form,

$$dY(t) = \beta[\theta(t) - Y(t)]dt + \sigma(t)dz(t)$$

where $Y(t)$ is the current temperature, $\theta(t)$ is the deterministic long-run level of the temperature, β is the speed at which the instantaneous temperature reverts to the long-run level $\theta(t)$, $\sigma(t)$ is the volatility which is season-dependent, and $z(t)$ is a Wiener process which models the temperatures random innovations. The functional forms for $\theta(t)$ and $\sigma(t)$ can be specified based on careful statistical analyses.

Once the process in is estimated, one can then value any contingent claim by taking expectation of the discounted future payoff, i.e.,

$$X = e^{-r(T-t)}E[g(t, Y_t)]$$

where X is the current value of the contingent claim, r is the riskfree interest rate, T is the maturity of the claim, and $g(t, Y_t)$ is the payoff at time.

Given the complex form of $\theta(t)$ and $\sigma(t)$ and the path-dependent nature of most payoffs, the formula usually does not have closed-form solutions. Monte Carlo simulations must be used. There is a main drawbacks of this continuous setup, that is it allows no place for the market price of risk. Instead, a risk-neutral valuation is imposed without any theoretical justification.

Modelling temperature

The goal is to find a stochastic process describing the temperature movements. When we later on want to price weather derivatives based on temperature it will be of great use to have an idea of how the temperature process behaves.

The mean temperature clearly show a strong seasonal variation in the temperature. It should be possible to model the seasonal dependence with, for example, some sine-function. This function would have the form

$$sin(\omega t + \phi)$$

Moreover, a closer look at the data series reveals a positive trend in the data. It is weak but it does exist. The mean temperature actually increases each year. There can be many reasons to this. One is the fact that we may have a global warming trend all over the world. Another is the so called urban heating effect, which means that temperatures tend to rise in areas nearby a big city, since the city is growing and warming its surroundings. We will assume, as a first approximation, that the warming trend is linear. Summing up, a deterministic model for the mean temperature at time t,

$$T_t^m = A + Bt + Csin(\omega t + \phi)$$

Unfortunately temperatures are not deterministic. Thus, to obtain a more realistic model we now have to add some sort of noise to the deterministic model. One choice is a standard Wiener process. The daily temperature differences are well fitted by a normal distribution, though the probability of getting small differences in the daily mean temperature will be slightly underestimated. A closer look at the data series reveals that the volatility of the temperature varies across the different months of the year, but is nearly constant within each month. Especially during the winter the volatility is much higher than during the rest of the year. Therefore, we make the assumption that σ_t t is a piecewise constant function, with a constant value during each month.

We also know that the temperature cannot, for example, rise day after day for a long time. This means that our model should not allow the temperature to deviate from its mean value for more than short periods of time. In other words, the stochastic process describing the temperature we are looking for should have a mean-reverting property. Putting all the assumptions together, we model temperature by a stochastic process solution of the following SDE

$$dT_t = a(T_t^m - T_t)dt + \sigma_t dW_t$$

where a determines the speed of the mean-reversion. The solution of such an equation is usually called an Ornstein-Uhlenbeck process.

The problem with Eq.(3.4) is that it is actually not reverting to T_t^m in the long run. To obtain a process that really reverts to the mean we have to add the term

$$\frac{dT_t^m}{dt} = B + \omega C cos(\omega t + \phi)$$

to the drift term so that

$$dT_t = \left[\frac{dT_t^m}{dt} + a(T_t^m - T_t)\right] dt + \sigma_t dW_t$$

Starting at $T = x$ the solution of this model gives

$$T_t = (x - T_s^m)e^{-a(t-s)} + T_t^m + \int_s^t e^{-a(t-u)} \sigma_u dW_u$$
$$T_t^m = A + Bt + C sin(\omega t + \phi)$$

Pricing weather derivatives

The market for weather derivatives is a typical example of an incomplete market, because the underlying variable, the temperature, is not tradable. Therefore we have to consider the market price of risk *lambda*, in order to obtain unique prices for such contracts. Since there is not yet a real market from which we can obtain prices, we assume for simplicity that the market price of risk is constant. Note that estimating λ is a major problem in weather derivatives. Furthermore, we assume that we are given a risk free asset with constant interest rate r and a contract that for each degree Celsius pays one unit of currency. Thus, under a martingale measure

Q, characterized by the market price of risk λ, our price process also denoted by T_t satisfies the following dynamics:

$$dT_t = \left[\frac{dT_t^m}{dt} + a(T_t^m - T_t) - \lambda\sigma_t\right]dt + \sigma_t dW_t$$

Since the price of a derivative is expressed as a discounted expected value under martingale measure Q, we start by computing the expected value and the variance of T_t under the measure Q. Indeed, as a Girsanov transformation only changes the drift term, the variance of T_t is the same under both measures. Therefore,

$$Var[T_t|\mathcal{F}_s] = \int_s^t e^{-2a(t-u)}\sigma_u^u dW_u$$

Moreover

$$E^P[T_t|\mathcal{F}_s] = (T_s - T_s^m)e^{-a(t-s)} + T_t^m$$

and thus

$$E^Q[T_t|\mathcal{F}_s] = (T_s - T_s^m)e^{-a(t-s)} + T_t^m - \int_s^t \lambda e^{-a(t-u)}\sigma_u dW_u$$

$$T_t = (x - T_s^m)e^{-a(t-s)} + T_t^m + \int_s^t \lambda e^{-a(t-s)}\sigma_u dW_u$$

Integrating we find

$$E^Q[T_t|\mathcal{F}_s] = (T_s - T_s^m)e^{-a(t-s)} + T_t^m - \frac{\lambda\sigma_i}{a}(1 - e^{-a(t-s)})$$

$$Var[T_t|\mathcal{F}_s] = \frac{\sigma_i^2}{2a}(1 - e^{-2a(t-s)})$$

For later use, we need to compute the covariance of the temperature between two different days. For $0 \le s \le t \le u$,

$$Cov[T_t, T_u|\mathcal{F}_s] = e^{-a(u-t)}Var[T_t|\mathcal{F}_t]$$

Suppose now that t_1 and t_n denote the first and last day of a month and start the process at some time s from the month before $[t_1, t_n]$. To compute the expected value and variance of T_t in this case, we split the integrals and into two integrals where σ is constant in each one of them, and equal to σ_i and σ_j respectively.

$$E^Q[T_t|\mathcal{F}_s] = (T_s - T_s^m)e^{-a(t-s)} + T_t^m - \int_s^{t_1} \lambda e^{-a(t-u)}\sigma_i dW_u + \int_{t_1}^t \lambda e^{-a(t-u)}\sigma_j dW_u$$

We then get

$$E^Q[T_t|\mathcal{F}_s] = (T_s - T_s^m)e^{-a(t-s)} + T_t^m - \frac{\lambda(\sigma_i - \sigma_j)}{a}e^{-a(t-t_1)} + \frac{\lambda\sigma_i}{a}e^{-a(t-s)} - \frac{\lambda\sigma_j}{a}$$

and the variance is

$$Var[T_t|\mathcal{F}_s] = \frac{(\sigma_i^2 - \sigma_j^2)}{2a}e^{-2a(t-t_1)} - \frac{\sigma_i^2}{2a}e^{-2a(t-s)} + \frac{\sigma_j^2}{2a}$$

Pricing a heating degree day option

Recall from that the payout of the HDD call option is of the form

$$X = \alpha \max(H_n - K, 0)$$

where, for simplicity we take $\alpha = 1$ unit of currency/HDD and

$$H_n = \sum_{i=1}^{n} \max(18 - T_{t_i}, 0)$$

The contract (4.10) is a type of an arithmetic average Asian option. In the case of a log-normally distributed underlying process, no exact analytic formula for the price of such an option is known. Here we have an underlying process which is normally distributed, but the maximum function complicates the task to find a pricing formula. Suppose that we want to find the price of a contract whose payout depends on the accumulation of HDDs during some period in the winter, for example the month of January in some cold country. In this case the probability that $\max(18 - T_{t_i}, 0) = 0$ should be extremely small on a winter day. Therefore, for such a contract we may write

$$H_n = 18n - \sum_{i=1}^{n} T_{t_i}$$

The distribution of this is easier to determine. We know that T_{t_i} all samples from an Ornstein-Uhlenbeck process, which is a Gaussian process. Since the sum in H_n is a linear combination of these elements, H_n is also Gaussian.

It only remains to compute the first and second moments of H_n.

$$E^Q[H_n|\mathcal{F}_t] = E^Q[18n - \sum_{i=1}^{n} T_{t_i}|\mathcal{F}_t = 18n - \sum_{i=1}^{n} E^Q[T_{t_i}|\mathcal{F}_t]$$

and

$$Var[H_n|\mathcal{F}_t] = \sum_{i=1}^{n} Var[T_{t_i}|\mathcal{F}_t] + 2\sum_{j\neq i}\sum_{i=1}^{n} Cov[T_{t_i}, T_{t_j}|\mathcal{F}_t]$$

Now, suppose that we have made the calculations above, and found that

$$E^Q[H_n|\mathcal{F}_t] = \mu_n$$

and

$$Var[H_n|\mathcal{F}_t] = \sigma_n^2$$

Thus

$$H_n \sim N(\mu_n, \sigma_n^2)$$

Then the price at $t \leq t_1$ of the claim is

$$c(t) = e^{-r(t_n-t)} E^Q[(H_n - K)^+|\mathcal{F}_t] = e^{-r(t_n-t)} \left[(\mu_n - K)N(-\alpha_n) + \frac{\sigma_n}{\sqrt{2\pi}} e^{-\alpha_n^2/2} \right]$$

where $\alpha_n = (K - \mu_n)/\sigma_n)$

Similarly for the put whose payoff is $(K - H_n)^+$ we would find

$$p(t) = e^{-r(t_n - t)} \left[(\mu_n - K)N(\alpha_n) - N\left(-\frac{\mu_n}{\sigma_n}\right) + \frac{\sigma_n}{\sqrt{2\pi}}(e^{-\alpha_n^2/2} - e^{-\frac{1}{2}\mu_n^2/\sigma_n^2}) \right]$$

The formulas above hold primarily for contracts during winter months, which typically is the period November-March. During the summer we cannot use these formulas without restrictions because the mean temperatures are very close to, or even higher than, 18o C. For such contracts we could use the method of Monte Carlo simulations. The reference level 18C originates from the US market, but it seems to be used also in Europe. Perhaps it could be more interesting to base the derivatives on some reference level which is closer to the expected mean temperature for the period.

Maximum payouts In practice many options often have a cap on the maximum payout. The reason is to reduce the risks that extreme weather conditions would cause. An option with a maximum payout could be constructed from two options without maximum payouts. If we enter a long position in one option and a short position in another option with a higher strike value, we get a payout function that would look capped.

In-period valuation

Often one would like to find the price of the option inside the contract period. Suppose we want to find the the price at a time t_i s.t. $t_1 \leq t_i \leq t_n$. We could then rewrite the variable H_n as $H_n = H_i + H_j$ where H_i is known at t_i and H_j is stochastic. The payout of the HDD call option can then be rewritten as

$$X = (H_n - K)^+ = (H_i + H_j - K)^+ = (H_j - \tilde{K})^+$$

where $\tilde{K} = K - H_i$ An in-period option can thus be valued as an out-of-period option with transformed strike as above.

Appendix A: Review of probability concepts

Set

A **set** can be thought of as a well-defined collection of objects considered as a whole. The objects of a set are called elements or members. The elements of a set can be anything: numbers, people, letters of the alphabet, other sets, and so on. Sets are conventionally denoted with capital letters, A, B, C, etc. Two sets A and B are said to be equal, written A = B, if they have the same members.

A set can also have zero members. Such a set is called the empty set (or the null set) and is denoted by the symbol ϕ.

If every member of the set A is also a member of the set B, then A is said to be a subset of B, written $A \subset B$, also pronounced A is contained in B.

Union Let A and B be any two sets. The set which consists of all the points which are in A or B or both is defined the union and is written $A \cup B$.

Intersection Let A and B be any two sets. The set which consists of all the points which are both in A and B is defined the intersection and is written $A \cap B$ or AB.

Let $A_1, A_2, \cdots A_n \subset U$:

Disjoint if $A_i \cap A_j = \phi$ for any i, j.

A set U is called **open** if, intuitively speaking, you can "wiggle" or "change" any point x in U by a small amount in any direction and still be inside U. In other words, if x is surrounded only by elements of U; it can't be on the edge of U.

As a typical example, consider the open interval (0,1) consisting of all real numbers x with $0 < x < 1$. If you "wiggle" such an x a little bit (but not too much), then the wiggled version will still be a number between 0 and 1. Therefore, the interval (0,1) is open. However, the interval (0,1] consisting of all numbers x with $0 < x \leq 1$ is not open; if you take $x = 1$ and wiggle a tiny bit in the positive direction, you will be outside of (0,1].

A **closed** set is a set whose complement is open.

If $A \subset U$ then the complement of A in U, denoted by A^C, or $\bar{A}$ is:

$$A^C = U - A$$

a Σ-algebra (or Σ-field) over a set X is a family of subsets of X that is closed under countable set operations; Σ-algebras are mainly used in order to define measures on X.

0.1 Function

A **function** is a relation, such that each element of a set (the domain) is associated with a unique element of another (possibly the same) set (the codomain).

Probability Systems

We need to acquire an understanding of the different parts of a probability system and how they fit together. In order to make some sense of it all, we shall find it useful to think of a probability system as a physical experiment with a random outcome. To be more concrete, we shall use a specific example to guide us through the various definitions and what they signify.

Suppose that we toss a coin three times and record the results in order. This is a very simple experiment, but note that we should not necessarily assume that the coin toss is fair, with an equally likely outcome for heads or tails. There can, in principle, be many different probabilities associated with the same physical experiment. This will have an impact on how we price derivatives.

Sample Space

The basic entity in a probability system is the sample space, usually denoted Ω, which is a set containing all the possible outcomes of the experiment. If we denote heads by H and tails by T, then there are 8 different possible outcomes of the coin-tossing experiment, and they define the sample space as follows:

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

Definition:

The sample space $\Omega = \{\omega_i\}_{i=1}^N$ is the set of all possible outcomes of the experiment.

Note that we are assuming that the sample space is finite. This is applicable to the discrete time formalism that we will develop in the binomial model but will have to be modified for the continuous time formalism.

Event Space

We are eventually going to want to talk about the probability of a specific event occurring. Is the sample space, simply as given, adequate to allow us to discuss such a concept? Unfortunately not.

This is because we want to ask more than just,

What is the probability that the outcome of the coin toss is a specific element of the sample space.

We also want to ask,

What is the probability that such-and-such specific events occur.

In order to be able to answer this, we need the concept of the set of all the events that we are interested in. This is called the event space, usually denoted Σ.

What conditions should an event space satisfy? The most basic event is Ω itself, that is, the event that one of the possible outcomes occurs. This event has probability one, that is, it always happens. It would thus make sense to require the event space to contain Ω.

Likewise, we shall assume that the null event Φ, which occurs with probability zero, is also in the event space.

Next, suppose that the events $A = \{HTT, THH\}$ and $B = \{HTH, HHH, HTT\}$ are elements of Σ. It is natural to be interested in the event that either A or B occurs. This is the union of the events, $A \cup B = \{HTT, THH, HTH, HHH\}$. We would like Σ to be closed under the union of two of its elements.

Finally, if the event $C = \{HHH, HTH, HTT\}$ is an element of Σ, then the probability of it occurring is one minus the probability that the complementary event $\Omega - C = \{HHT, TTT, TTH, THT, THH\}$ occurs. Hence if an event is in Ω, we would also like its complement to be in Ω. We can summarise the definition of the event space as follows.

Definition: The event space Σ is a set of subsets of the sample space Ω, satisfying the following conditions:

1. $\Omega \in \Sigma$

2. if $A, B \in \Sigma$, then $A \cup B \in \Sigma$

3. if $A \in \Sigma$, then $\Omega - A \in \Sigma$

Note that for our purposes, we can take Σ to be the power set (the set of all subsets) of Ω. The power set of our example system is perhaps just slightly too large to comfortably write out. It contains $2^8 = 256$ elements.

The system consisting of the sample space and the event space (Ω, Σ) might appropriately be called a possibility system, as opposed to a probability system because all that it tells us are the possible outcomes of our experiment. It contains no information about how probable each event is. The so-called probability measure is an additional ingredient, that must be specified in addition to the pair (Ω, Σ).

0.2 Probability Measure

Now suppose that we want to assign a probability to each event in Σ. We can do this by means of a probability measure $P : \Sigma \to [0, 1]$. For any event $A \in \Sigma$, P[A] is the probability that the event A occurs. For example, if the coin is fair, then the probability of any event XY Z occurring (where X, Y , Z can be either H or T) is clearly 1/8.

Now, what conditions should we place on a probability measure? We have already constrained its values to lie between zero and one. Since the event Ω always occurs, its probability is one. Finally, if we have two disjoint sets, then the probability of their union occurring should be equal to the sum of the probabilities of the disjoint sets. For example,

$$P[\{HHH, TTT\}] = Prob[\{HHH\}] + Prob[\{TTT\}] = \frac{1}{4}$$

Definition: A probability measure P is a function $P \to [0, 1]$ satisfying

1. $P[A] > 0$ for every $A \in \Sigma$
2. $P[\Omega] = 1$
3. if $A, B \in \Omega$ and $A \cap B = \Phi$, then $P[A \cup B] = P[A] + P[B]$.

Taken together, the sample space, event space and probability measure form a so-called probability system, denoted $\mathcal{P} = (\Omega, \Sigma, P)$.

We can in principle consider various probability measures on the same sample and event spaces (Ω, Σ). This turns out to be very useful in financial analysis.

In our coin tossing example, we have already considered the probability measure P that we obtain if the coin that we are tossing is fair. However, we could also define a probability measure $Q : \Sigma \to [0, 1]$ that is based on an unfair coin. Suppose that for the unfair coin we get heads with probability 1/3, and tails with probability 2/3.

Then the probability measure is defined by the probabilities

$Q(\{HHH\}) = 1/27$

$Q(\{HHT\}) = Q(\{HTH\}) = Q(\{THH\}) = 2/27$

$Q(\{HTT\}) = Q(\{TTH\}) = Q(\{THT\}) = 4/27$

$Q(\{TTT\}) = 8/27.$

Both measures are, in principle, valid to consider, so that when we are talking about probabilities related to the coin tossing, we must specify whether we are in the probability system $\mathcal{P} = (\Omega, \Sigma, P)$ or in the probability system $\mathcal{Q} = (\Omega, \Sigma, Q)$ or possibly in some other system based on another weighting of the coins.

Random Variables

A random variable is a real-valued function X defined on the sample space Ω. Thus $X : \Omega \to \mathcal{R}$ assignes to each element ω_i of Ω an element of $\mathcal{R}$, that is, a real number. Even though the function is itself deterministic, that is, if we give X a definite input then we get a definite output, its argument ω_i is the random outcome of our physical experiment and hence $X(\omega_i)$ is also random. For example, X could be the function that counts the numbers of heads,

$X(\{HHH\}) = 3$

$X(\{HHT\}) = X(\{HTH\}) = X(\{THH\}) = 2$

$X(\{HTT\}) = X(\{TTH\}) = X(\{THT\}) = 1$

$X(\{TTT\}) = 0.$

Any function X is defined independently of any probability measure. Thus we could change probability measures from P to Q, and the values of X would be unaffected. However, what would be affected by a change of probability measure is the probability that X would take on some given value. In particular the expectation of X, which is the probability weighted sum over the sample space of the possible values of the random variable, will depend on the probability measure that we are using. This is obvious from the formula for the expectation,

$$E^P[X] = \sum_{i=1}^{n} P(\{\omega_i\})X(\omega_i)$$

which clearly depends in a crucial way on the probability measure. The notation $E^P[X]$ is used to denote the expectation of the random variable X with respect to the probability system P.

Going back to our example, suppose that we are in the fair-coin probability system P. Then the expectation of the random variable X that returns the number of heads is $E^P[X] = 1.5$ while in the unfair weighted-coin system Q, the expectation is $E^Q[X] = 1$

Recap

Definition : Probability Space A probability space is the triplet $(\Omega, \Sigma, P[\cdot])$ where Ω is a sample space, Σ is the σ-algebra of events and $P[\cdot]$ is a probability function with domain Σ.

Definition : Probability Function A probability function is a function with domain Σ (the σ-algebra of events) and counterdomain the interval $[0, 1]$ which satisfy the following axioms:

- $P[A] \geq 0$ for every A in Σ
- $P[\Omega] = 1$
- If $A_1, A_2, ...$ is a sequence of mutually exclusive events in Σ than

$$P[\bigcup_{i=1}^{\infty} A_i] = \sum_{i=1}^{\infty} P[A_i]$$

Properties of $P[\cdot]$

Theorem: $P[\phi] = 0$

Theorem: If A is an event in Σ, than $P[\bar{A}] = 1 - P[A]$

Theorem 4: If A_1 and A_2 are events in Σ, than $P[A_1] = P[A_1 \cap A_2] + P[A_1 \cap \bar{A}_2]$

Theorem (Law of Addition): If A_1 and A_2 are events in Σ, than

$P[A_1 \cup A_2] = P[A_1] + P[A_2] - P[A_1 \cap A_2]$

More generally for n events $A_1, A_2, \cdots, A_n$

$P[A_1 \cup \cdots \cup A_n] = \sum_{j=1}^{n} P[A_j] - \sum\sum_{i<j} P[A_i \cap A_j] + \sum\sum\sum_{i<j<k} P[A_i \cap A_j \cap A_k] + \cdots + (-1)^{n+1} P[A_1 \cap A_2 \cdots \cap A_n]$

If the events are mutually exclusive than $P[A_1 \cup \cdots \cup A_n] = \sum_{j=1}^{n} P[A_j]$

Theorem: If A_1 and A_2 are events is Σ and $A_1 \subset A_2$, than $P[A_1] \leq P[A_2]$

Conditional Probability and Independence

Let A and B be two events in Σ of the given probability space $(\Omega, \Sigma, P[\cdot])$.

The conditional probability of event A given B, denoted $P[A|B]$, is defined as

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

if $P[B] > 0$ and is undefined is $P[B] = 0$.

In the following $P(AB)$ is sometimes used as a short notation for $P(A \cap B)$.

Properties:

Assume $B \in \Sigma$ and $P[B] > 0$

Theorem: $P[\phi|B] = 0$

Theorem: If $A_1, A_2, \cdots, A_n$ are mutually exclusive than $P[A_1 \cup \cdots \cup A_n|B] = \sum_{i=1}^{n} P[A_i|B]$

Theorem : If A is an event in Σ, than $P[\bar{A}|B] = 1 - P[A|B]$

Theorem: If A_1 and A_2 are events in Σ, than $P[A_1|B] = P[A_1A_2|B] + P[A_1\bar{A}_2|B]$

Theorem: If A_1 and A_2 are events in Σ, than $P[A_1 \cup A_2|B] = P[A_1|B] + P[A_2|B] - P[A_1A_2|B]$

Theorem: If A_1 and A_2 are events is Σ and $A_1 \subset A_2$, than $P[A_1|B] \leq P[A_2|B]$

Theorem: If $A_1, A_2, \cdots A_n$ are events in Σ than $P[A_1 \cup \cdots \cup A_n|B] \leq \sum_{j=1}^{n} P[A_i|B]$

Theorem (Law of total probability): For a given probability space $(\Omega, \Sigma, P[\cdot])$, if $B_1, B_2, \cdots B_n$ is a collection of exhaustive and mutually disjoint events in Σ and $P[B_k] > 0$ for $k = 1, 2, \cdots n$, than for every A in Σ $P[A] = \sum_{j=1}^{n} P[A|B_j]P[B_j]$

Corollary: $P[A] = P[A|B]P[B] + P[A|\bar{B}]P[\bar{B}]$

Theorem 16 (Bayes Formula): For a given probability space $(\Omega, \Sigma, P[\cdot])$, if $B_1, B_2, \cdots B_n$ is a collection of exhaustive and mutually disjoint events in Σ and $P[B_k] > 0$ for
$k = 1, 2, \cdots n$, than for every A in Σ s.t. $P[A] > 0$

$P[B_k|A] = \frac{P[A|B_k]P[B_k]}{\sum_{j=1}^{n} P[A|B_j]P[B_j]}$

Corollary: $P[B|A] = \frac{P[A|B]P[B]}{P[A]}$

Theorem (Law of Multiplication): For a given probability space $(\Omega, \Sigma, P[\cdot])$, let $A_1, A_2, \cdots A_n$ be events in Σ for which $P[A_1 \cdots A_n] > 0$, than

$$P[A_1 A_2 \cdots A_n] = P[A_1]P[A_2|A_1]P[A_3|A_1 A_2] \cdots \qquad P[A_n|A_1 \cdots A_{n-1}]$$

Corollary: $P[AB] = P[A|B]P[B]$

Definition: Independent Events For a given probability space $(\Omega, \Sigma, P[\cdot])$, let A and B be two events in Σ. Events A and B are independent (or statistically independent or stocastically independent) if and only if any of the following conditions is satisfied:

- $P[A|B] = P[A]$ if $P[B] > 0$
- $P[B|A] = P[B]$ if $P[A] > 0$
- $P[AB] = P[A]P[B]$

Definition: Independence of several events For a given probability space $(\Omega, \Sigma, P[\cdot])$, let $A_1, A_2 \cdots A_n$ be n events in Σ. Events $A_1, A_2 \cdots A_n$ are defined to be independent if and only if

- $P[A_i A_j] = P[A_i]P[A_j]$ for $i \neq j$
- $P[A_i A_j A_k] = P[A_i]P[A_j]P[A_k]$ for $i \neq j, k \neq j, i \neq k$.

  $\vdots$
- $P[\cap_{i=1}^n A_i] = \prod_{i=1}^n P[A_i]$

Random Variables and Distribution

Definition: Random Variable. For a given probability space $(\Omega, \Sigma, P[\cdot])$, a random variable denoted by $X(\cdot)$ is a function with domain Ω and counterdomain the real line. The function $X(\cdot)$ makes some real number correspond to each outcome of the experiment.

Example Consider the experiment of tossing a coin. $\Omega = \{\text{tail, head}\}$. We can define a random variable X so that $X(\omega) = 1$ if $\omega =$ head and $X(\omega) = 0$ if $\omega =$ tail.

Definition: distribution function. The distribution function a random variable X, denoted $F_X(\cdot)$, is that function with domain the real line and counterdomain the interval $[0, 1]$ which satisfies $F_X(x) = P[X \leq x] = P[\omega : X(\omega) \leq x]$ for every real number x.

Properties

- $F_X(-\infty) = 0$ and $F_X(+\infty) = 1$
- $F_X(\cdot)$ is a monotone non decreasing function, i.e. $F_X(a) \leq F_X(b)$ if $a < b$

- $F_X(\cdot)$ is continuous from the right, i.e. $\lim_{0<h\to 0} F_X(x+h) = F_X(x)$

Definition: Discrete Random Variable A random variable will be defined discrete if the range of X is countable. If a random variable is discrete than the cumulative distribution function will be defined to be discrete.

Definition: Discrete Density Function If X is a discrete random variable with values $x_1, x_2, ...x_n, ...$ the function $f_X(x) = P[X = x_j]$ if $x = x_j, j = 1, ...n, ...$ and zero eitherwise is defined the discrete density function of X.

Definition: Continuous Random Variable A random variable will be called continuous if there exists a function $f_X(\cdot)$ such that $F_X(x) = \int_{-\infty}^{x} f_X(u)du$ for every real number x.

Definition: Probability Density Function If X is a continuous random variable the function $f_X(\cdot)$ in $F_X(x) = \int_{-\infty}^{x} f_X(u)du$ is called the probability density function (or continuous density function).

Any function $f(\cdot)$ with domain the real line and counterdomain $[0, \infty)$ is defined to be a probability density function iif:

- $f(x) \geq 0$ for all x
- $\int_{-\infty}^{\infty} f(x)dx = 1$

Expectations and moments

Definitions: Mean Let X be a random variable, the mean of X denoted by μ_X or $E[X]$ is defined by:

- $E[X] = \sum_j x_j f_X(x_j)$ id X is discrete with values $x_1, x_2, \cdots x_j, \cdots$.
- $E[X] = \int_{-\infty}^{\infty} x f_X(x)dx$ if X is continuous.

Definitions: Variance Let X be a random variable, the variance of X denoted as σ_X^2 or $var[X]$ is defined by

- $var[X] = \sum_j (x_j - \mu)^2 f_X(x_j)$ id X is discrete with values $x_1, x_2, \cdots x_j, \cdots$.
- $var[X] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)dx$ if X is continuous.

Theorem $var[X] = E[(X - E[X]^2)] = E[X^2] - (E[X])^2$

Definitions: Standard Deviation Let X be a random variable, the standard deviation of X, denoted by σ_X is defined as $\sqrt{var[X]}$

Definitions: Moments Let X be a random variable, the r^{th} moments of X here denoted as μ_r is defined by $\mu_X^r = E[X^r]$

Definitions: Central Moments Let X be a random variable, the r^{th} central moments of X here denoted as ν_X^r is defined by $\nu_X^r = E[(X - \mu_X)^r]$

Hence the mean is the first moment and the variance the second central moment.

ν_X^3 is called skewness; symmetrical distribution have $\nu_X^3 = 0$.

μ_X^3/σ^3 is called coefficient of skeweness.

ν_X^4 is called kurtosis and measure the flatness of a density near its center.

$\nu_X^4/\sigma^4 - 3$ is called coefficient of kurtosis. A Normal density has a coefficient of kurtosis equal to zero.

Definitions: Moment generating function Let X be a random variable with density $f_X(\cdot)$. The moment generating function is defined as

$$m(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x)dx$$

If a moment generating function exists than

$$\frac{d^r}{dt^r} m(t) = \int_{-\infty}^{\infty} x^r e^{tx} f_X(x)dx$$

Letting $t \to 0$ we have

$$\frac{d^r}{dt^r} m(0) = E[X^r] = \mu_X^r$$

A density function determines a set of moments (when they exist). But in general a sequence of moments does not determine a unique distribution function. However it can be proved that if a moment generating function exists than it uniquely determines the distribution function.

Definitions: Expected Value of a function of a random variable Let X be a random variable and $g(\cdot)$ a function with domain and counterdomain the real line. The expected value of $g(\cdot)$ denoted $E[g(X)]$, is defined

- $E[g(X)] = \sum_j g(x_j) f_X(x_j)$ id X is discrete with values $x_1, x_2, \cdots x_j, \cdots$.
- $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)dx$ if X is continuous.

Chebyshev Inequality

Suppose E(X^2) is finite, than

$$P[|X - \mu_X| \geq r\sigma_X] = P[(X - \mu_X)^2 \geq r^2\sigma_X^2] \leq \frac{1}{r^2}$$

The proof is not difficult. Let $g(X)$ be a nonnegative function of the random variable X with domain the real line; than for any k

$$P[g(X) \geq k] \leq \frac{E[g(X)]}{k}$$

Proof:

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

$$\int_{x:g(x)\geq k} g(x)f_X(x)dx + \int_{x:g(x)\leq k} g(x)f_X(x)dx$$
$$\geq \int_{x:g(x)\leq k} g(x)f_X(x)dx \geq \int_{x:g(x)\leq k} kf_X(x)dx$$
$$= kP[g(X) \geq k]$$

Now let $g(X) = (x - \mu_X)^2$ and $k = r^2\sigma_X^2$ and the Chebyshev inequality follows.

Remark:

$$P[|X - \mu_X| \leq r\sigma_X] \geq 1 - \frac{1}{r^2}$$

so that

$$P[\mu_X - r\sigma_x < X < \mu_X + r\sigma_x] \geq 1 - \frac{1}{r^2}$$

that is, the probability that X falls within $r\sigma_X$ units of μ_X is greater than or equal to $1 - \frac{1}{r^2}$. For $r = 2$ one gets:

$$P[\mu_X - 2\sigma_x < X < \mu_X + 2\sigma_x] \geq \frac{3}{4}$$

or, for any random variable with finite variance at least three-fourth of the mass of X falls within two standard deviation from its mean. The Chebyshev inequality gives a bound, which does not depend on the distribution of X, for the probability of particular events in terms of a random variable and its mean and variance.

Jensen Inequality Let X be a random variable with mean $E[x]$ and let $g(X)$ be a convex function; than $E[g(X)] \geq g([E[X])$. For example $g(x) = x^2$ is convex; hence $E[X^2] \geq (E[X])^2$, which guaranties that the variance of X is non-negative.

Important distributions

0.3 Discrete Distributions

- Binomial
- Poisson

0.4 Continuous Distributions

- Uniform
- Normal
- Exponential
- Gamma

Uniform distribution

$x \sim U(a, b)$

Density

$$f_X(x) = \frac{1}{b-a} I_{\{x \in (a,b)\}}$$

Normal distribution

$x \sim N(\mu, \sigma^2)$

Density

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Discrete random variables

- Bernulli distribution

$$f_X(x) = p^x(1-p)^{(1-x)} I_{\{0,1\}}(x)$$

  where $0 \leq p \leq 1$

- Binomal distribution

$$f_X(x) = \binom{n}{x} p^x(1-p)^{(n-x)} I_{\{0,1,2,\cdots,n\}}(x)$$

  where $0 \leq p \leq 1$ and n is a positive integer. The binomial coefficient $\binom{n}{x} = \frac{n!}{(n-x)!x!}$

- Poisson

$$f_X(x) = \frac{e^{-\lambda}\lambda^{-x}}{x!} I_{\{0,1,2,\cdots\}}(x)$$

  where $\lambda > 0$

Show that

- Bernulli distribution

$$E[x] = p \quad var[X] = pq, \quad m_x(t) = pe^t + q$$

  where q = 1-p

- Binomal distribution

$$E[x] = np \quad var[X] = npq, \quad m_x(t) = (pe^t + q)^n$$

- Poisson

$$E[x] = \lambda \quad var[X] = \lambda, \quad m_x(t) = e^{\lambda(e^t-1)}$$

Uniform random variables

$$f_X(x) = \frac{1}{b-a} I_{\{a,x,b\}}$$

Show that

$$E[x] = \frac{a+b}{2} \quad var[X] = \frac{(b-a)^2}{12}, \quad m_x(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}$$

Normal random variables

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Let's show that

$$I = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1$$

Change variable $y = \frac{(x-\mu)}{\sigma}$

$$I = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$

We do not know how to calculate I but we can calculate I^2

$$I^2 = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

$$I^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{\frac{-(y^2+z^2)}{2}} dydz$$

In polar coordinate $y = rcos\theta, \quad z = rsin\theta, \quad dydz = rd\theta dr$

$$I^2 = \frac{1}{2\pi} \int_0^{\infty} \int_0^{2\pi} e^{\frac{-r^2}{2}} rd\theta dr =$$

$$\frac{1}{2\pi} 2\pi \int_0^{\infty} e^{\frac{-r^2}{2}} rdr =$$

$$- \int_0^{\infty} de^{\frac{-r^2}{2}} = 1$$

Hence $I^2 = 1$ and also $I = 1$.

Exercise 1: Say X is $N(\mu, \sigma^2)$. Derive the moment generating function of X.

$$m_X(t) = E[e^{tX}] = e^{t\mu} E[e^{t(X-\mu)}] =$$

$$e^{t\mu} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{x-\mu} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx =$$

$$e^{t\mu} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}[(x-\mu)^2 - 2\sigma^2 t(x-\mu)]} dx =$$

$$e^{t\mu} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}[(x-\mu-\sigma^2 t)^2 - \sigma^4 t^2]} dx =$$

$$e^{t\mu + \frac{1}{2}\sigma^2 t^2} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}[(x-\mu-\sigma^2 t)^2]} dx = e^{t\mu + \frac{1}{2}\sigma^2 t^2}$$

From the moment generating function we can derive

$$m_X(t) = e^{t\mu + \frac{1}{2}\sigma^2 t^2}$$

we can derive

$$E[X] = m'_X(0) = \mu$$
$$var[X] = m''_X(0) = E[X^2] - E[X]^2 = \sigma^2$$

Exercise 2: Say X is $N(\mu, \sigma^2)$, where

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

Derive the p.d.f. of $W = e^X$.

<u>Solution:</u> We find the probability distribution function of W by first calculating its (cumulative) distribution function, that is,

$$F_W(w) = P(W \leq w) = P(e^X \leq w).$$

Since the exponential function is an increasing function, we can take its inverse on both sides of the last inequality to obtain

$$F_W(w) = P(X \leq \log w) = \int_{-\infty}^{\log w} \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx, \qquad 0 < w.$$

Using Leibniz's rule for differentiating an integral

$$\frac{d}{dz}\int_{a(z)}^{b(z)} f(x,z)\ dx = \int_{a(z)}^{b(z)} \frac{\partial f(x,z)}{\partial z}\ dx + f(b(z),z)\frac{db(z)}{dz} - f(a(z),z)\frac{da(z)}{dz}$$

Therefore, the p.d.f for W is

$$f_W(w) = \frac{dF_W}{dw}(w) = \frac{1}{\sqrt{2\pi}\sigma w} exp\left[-\frac{(\log w - \mu)^2}{2\sigma^2}\right], \qquad 0 < w.$$

This distribution is called a *lognormal* distribution, for obvious reasons.

Exercise 3: Say $\log W \sim N(\mu, \sigma^2)$. Show that

$$E[W] = e^{\mu + \frac{1}{2}\sigma^2}$$

and

$$var[W] = e^{2\mu + 2\sigma^2} - e^{2\mu + \sigma^2}$$

<u>Solution:</u>

Take $X \sim N(\mu, \sigma)$. Then $W = e^X$. Use

$$E[W] = E[e^X] = m_X(1)$$

and

$$varW = E[e^{2X}] - E[e^X]^2 = m_X(2) - m_X(1)^2$$

Exercise 4: Let $X \sim N(0,1)$. Derive the density function of $Z = X^2$, its moment generating function of and the first two moments.

Solution:

If X have a standard normal distribution $N(0,1)$ the distribution of their squares $Z = X^2$ is:

$$\begin{aligned} F_Z(z) = P(X^2 \le z) &= P(-\sqrt{z} \le X \le \sqrt{z}) \\ &= \int_{-\sqrt{z}}^{\sqrt{z}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &= 2\int_0^{\sqrt{z}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx. \end{aligned}$$

Using Leibniz's rule for differentiating an integral

$$\frac{d}{dz}\int_{a(z)}^{b(z)} f(x,z)\, dx = \int_{a(z)}^{b(z)} \frac{\partial f(x,z)}{\partial z}\, dx + f(b(z),z)\frac{db(z)}{dz} - f(a(z),z)\frac{db(z)}{dz}$$

the p.d.f for Z is

$$f_Z(z) = \frac{dF_Z}{dz} = \frac{1}{z\sqrt{2\pi}} e^{-z/2}.$$

This is the χ^2-distribution with parameter 1, or $\chi^2(1)$.

The moment generating function is

$$m_Z(t) = E[e^{tZ}] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{tx^2} e^{-\frac{1}{2}x^2} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2(1-2t)} dx$$

Change variable $y = x(1-2t)^{\frac{1}{2}}$ and

$$m_Z(t) = (1-2t)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy = (1-2t)^{-\frac{1}{2}}$$

Hence $E[Z] = m'_Z(0) = 1$ and $var[Z] = m''_Z(0) = 2$.

Exercise 5:

Let X_1 and X_2 be independent with normal distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$. Find the p.d.f. of $Y = X_1 - X_2$.

Solution: The moment generating function for Y is

$$\begin{aligned} m_Y(t) &= E(e^{t(X_1 - X_2)}) \\ &= E(e^{tX_1} e^{-tX_2}) \\ &= E(e^{tX_1})E(e^{-tX_2}) \\ &= m_{X_1}(t) m_{-X_2}(t), \end{aligned}$$

since X_1 and X_2 are independent. Using the fact that X_1 and X_2 are normally distributed, we have that

$$m_{X_1}(t) = \exp\left(\mu_1 t + \frac{1}{2}\sigma_1^2 t^2\right)$$

and

$$m_{-X_2}(t) = m_{X_2}(-t) = \exp\left(-\mu_2 t + \frac{1}{2}\sigma_2^2 t^2\right).$$

Therefore

$$\begin{aligned} m_Y(t) &= \exp\left(\mu_1 t + \frac{1}{2}\sigma_1^2 t^2\right)\exp\left(-\mu_2 t + \frac{1}{2}\sigma_2^2 t^2\right) \\ &= \exp\left((\mu_1 - \mu_2)t + \frac{(\sigma_1^2 + \sigma_2^2)t^2}{2}\right). \end{aligned}$$

We recognise this as being the moment generating function of a normally distributed random variable with mean $\mu_1 - \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$. Thus the p.d.f. for Y is $N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$.

Exercise: Show that if X_1 $N(\mu_1, \sigma_1^2)$ and X_2 $N(\mu_2, \sigma_2^2)$ then $Y = X_1 + X_2$ $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Bivariate discrete distributions:

If X and Y are random variables, then we can define a joint probability function $f(x,y) = P(X = x, Y = y)$ where $f(x,y)$ is non-negative and $\sum\sum f(x,y) = 1$. We then define the *marginal* distributions

$$f_X(x) = \sum_y f(x,y) \qquad f_Y(y) = \sum_x f(x,y)$$

These are the probability functions for X and Y separately.

The *conditional probability of X given Y* is:

$$P(X \,|\, Y) = \frac{P(XY)}{P(Y)}$$

Bivariate continuous distributions

Here we have a probability density $f(x)$, which it is important to recognize is not by itself a probability. Instead, the probability that X lies in the interval $(x, x+dx)$ is $f(x)dx$, or for a finite interval (a, b) it is $\int_a^b f(x)dx$.

In the same way, we can have continuous bivariate (or mutlivariate) probability distributions. For that, we have a probability density function $f(x, y)$ and again this isn't a probability on its own. Instead, the joint probability that X lies in the interval $(x, x+dx)$ and Y lies in the interval $(y, y+dy)$ is $f(x, y)dxdy$. Finite probabilities are then given by double integrals: $P[(X, Y) \in A] = \int\int f(x, y)dxdy$.

Any function $f(x, y)$ can be a bivariate probability function providing that it is non-negative and its integral over the entire $x - y$ plane is unity.

We can then define marginal densities:

$$f_X(x) = \int_{\infty}^{\infty} f(x, y)dy \qquad f_Y(y) = \int_{\infty}^{\infty} f(x, y)dx$$

The conditional densities are:

$$f(x \,|\, y) = \frac{f(x, y)}{f_Y(y)}$$

$$f(y \,|\, x) = \frac{f(x, y)}{f_X(x)}$$

The condition for the independence of X and Y is that

- The joint density factorizes

$$f_{X,Y}(x, y) = f_X(x) f_Y(y)$$

- the marginal densities are equal to the conditional densities:

$$f(x \,|\, y) = f_X(x) \qquad f(y \,|\, x) = f_Y(y)$$

Covariance and Correlation The covariance between two r.v. X and Y is defined as

$$Cov(X, Y) = E[X - E[X]]E[Y - E[Y]] = E[XY] - E[X]E[Y]$$

Exercise: prove equality above.

The correlation coefficient is defined as

$$\rho_{X,Y} = \frac{Cov(X, Y)}{\sqrt{var(X)var(Y)}} = \frac{Cov(X, Y)}{\sigma_(X)\sigma_(Y)}$$

Note that

- $cov[X,X] = var[X]$ and $\rho_{X,X} = 1$
- If $cov[X,Y] = 0$ we say the two r.v. are uncorrelated.
- if X and Y are independent they are also uncorrelated but the reverse in not normally true.

Bivariate normal distribution

The bivariate normal distribution for two random variables X,Y, is given by

$$\begin{aligned} f(x,y) &= \frac{1}{2\pi\sigma_X\sigma_Y(1-\rho^2)^{1/2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right.\right. \\ & \left.\left. -2\rho\frac{x-\mu_X}{\sigma_X}\frac{y-\mu_Y}{\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right\} \end{aligned}$$

where ρ is the correlation coefficient between X and Y. If $\rho = 0$

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right] \exp\left[-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right] = f_X(x)f_Y(y)$$

Multivariate normal distribution

The multivariate normal distribution for n random variables $\mathbf{X} = (X_1, X_2, \cdots X_n)$ is given by

$$f(x,y) = \frac{1}{(2\pi)^{n/2}(det\Sigma)^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\mu)\Sigma^{-1}(\mathbf{x}-\mu)'\right\}$$

where $\mathbf{x}, \mu \in \mathcal{R}^n$, Σ is the variance-covariance matrix, Σ^{-1} its inverse and $det\Sigma$ its determinant.

Exercise 6:

Let $f_{X,Y}(x,y)$ be the normal bivariate distribution. Calculate

a) $\int_{-\infty}^{\infty} f(x,y)dxdy$

b) $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$

c) The moment generating function of the bivariate normal distribution $m_{X,Y}(t_1,t_2)$

d) The conditional distribution $f_{X|Y}(x|y)$

Solution

(a) The bivariate normal distribution for two random variables X,Y, is given by

$$
\begin{aligned}
f(x,y) \;=\; & \frac{1}{2\pi\sigma_X\sigma_Y(1-\rho^2)^{1/2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right.\right. \\
& \left.\left. -2\rho\frac{x-\mu_X}{\sigma_X}\frac{y-\mu_Y}{\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right\} \qquad (38)
\end{aligned}
$$

To compute the integral of $f(x,y)$ over the entire plane, let us first do the change of variables

$$u = \frac{x-\mu_X}{\sigma_X} \qquad \text{and} \qquad v = \frac{y-\mu_Y}{\sigma_Y}.$$

Therefore,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = \frac{1}{2\pi(1-\rho^2)^{1/2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-[1/2(1-\rho^2)](u^2-2\rho uv+v^2)}dudv.$$

Now we complete the squares for the variable u in the exponent as follows

$$
\begin{aligned}
\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2) &= \frac{1}{2(1-\rho^2)}(u^2-2\rho uv+\rho^2v^2-\rho^2v^2+v^2) \\
&= \frac{1}{2(1-\rho^2)}[(u-\rho v)^2+(1-\rho^2)v^2] \\
&= \frac{1}{2}\left[\left(\frac{u-\rho v}{\sqrt{1-\rho^2}}\right)^2+v^2\right].
\end{aligned}
$$

We are then led to the following change of variables

$$w = \frac{u-\rho v}{\sqrt{1-\rho^2}},$$

which reduces the integral to

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-w^2/2}dw\int_{-\infty}^{\infty} e^{-v^2/2}dv = 1.$$

(b) We now need to calculate the marginal distribution for the random variable X. This amounts to performing an integration only over the variable y, so we begin with the change of variables

$$v = \frac{y-\mu_Y}{\sigma_Y}.$$

This leads to

$$
\begin{aligned}
f_X(x) \;&=\; \int_{-\infty}^{\infty} f(x,y)dy \;=\; \frac{1}{2\pi\sigma_X(1-\rho^2)^{1/2}} \\
&\times \int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2-2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)v+v^2\right]\right\}dv \\
&= \frac{\exp\left\{-\frac{(x-\mu_X)^2}{2(1-\rho^2)\sigma_X^2}\right\}}{2\pi\sigma_X(1-\rho^2)^{1/2}}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[v^2-2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)v\right]\right\}dv.
\end{aligned}
$$

we complete the squares in the exponent of the integrand as follows

$$\begin{aligned} v^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)v &= v^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)v + \rho^2\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \rho^2\left(\frac{x-\mu_X}{\sigma_X}\right)^2 \\ &= \left(v - \rho\frac{x-\mu_X}{\sigma_X}\right)^2 - \rho^2\left(\frac{x-\mu_X}{\sigma_X}\right)^2. \end{aligned}$$

Back to the integral and with the change of variables

$$t = \frac{1}{(1-\rho^2)^{1/2}}\left(v - \rho\frac{x-\mu_X}{\sigma_X}\right)$$

we obtain

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy \\ &= \frac{\exp\left\{\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right\}}{2\pi\sigma_X}\int_{-\infty}^{\infty} e^{-t^2/2}dt \\ &= \frac{1}{\sqrt{2\pi}\sigma_X}\exp\left\{\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right\}. \end{aligned}$$

(c) The joint moment generating function is defined as

$$\begin{aligned} m_{X,Y}(t_1,t_2) = m(t_1,t_2) &= E[e^{t_1X+t_2Y}] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1X+t_2Y} f(x,y)dxdy. \end{aligned}$$

We begin with the same change of variables as before, namely,

$$u = \frac{x-\mu_X}{\sigma_X} \qquad \text{and} \qquad v = \frac{y-\mu_Y}{\sigma_Y}.$$

This leads us to

$$\begin{aligned} m_{X,Y}(t_1,t_2) &= \frac{e^{t_1\mu_X+t_2\mu_Y}}{2\pi(1-\rho^2)^{1/2}} \\ &\times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1\sigma_X u + t_2\sigma_Y v - [1/2(1-\rho^2)](u^2-2\rho uv+v^2)}dudv. \end{aligned}$$

The total exponent in the integrand is

$$-\frac{1}{2(1-\rho^2)}[u^2 - 2(1-\rho^2)t_1\sigma_X u - 2\rho uv + v^2 - 2(1-\rho^2)t_2\sigma_Y v]$$

Completing the squares for the variable u we obtain

$$\begin{aligned} -\frac{1}{2(1-\rho^2)}\{[u &- (1-\rho^2)t_1\sigma_X - \rho v]^2 + v^2 - 2(1-\rho^2)t_2\sigma_Y v \\ &- (1-\rho^2)^2t_1^2\sigma_X^2 - 2(1-\rho^2)t_1\sigma_X\rho v - \rho^2v^2\}, \end{aligned}$$

and completing the squares in the variable v we get

$$\begin{aligned}-\frac{1}{2(1-\rho^2)}\{[u &- (1-\rho^2)t_1\sigma_X - \rho v]^2 + (1-\rho^2)(v - t_2\sigma_Y - t_1\sigma_X\rho)^2 \\ &- (1-\rho^2)(t_1^2\sigma_X^2 + t_2^2\sigma_Y^2 + 2t_2\sigma_Y t_1\sigma_X\rho)\}.\end{aligned}$$

With the change of variables

$$w = \frac{u - (1-\rho^2)t_1\sigma_X - \rho v}{\sqrt{1-\rho^2}} \quad \text{and} \quad z = v - t_2\sigma_Y - t_1\sigma_X\rho,$$

the exponent becomes

$$-\frac{1}{2}w^2 - \frac{1}{2}z^2 + \frac{1}{2}(t_1^2\sigma_X^2 + t_2^2\sigma_Y^2 + 2t_2\sigma_Y t_1\sigma_X\rho).$$

Returning to the integral, we obtain

$$m_{X,Y}(t_1, t_2) = \frac{e^{t_1\mu_X + t_2\mu_Y \frac{1}{2}(t_1^2\sigma_X^2 + t_2^2\sigma_Y^2 + 2t_2\sigma_Y t_1\sigma_X\rho)}}{2\pi} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}w^2 - \frac{1}{2}z^2} dz dw.$$

Therefore

$$m(t_1, t_2) = e^{t_1\mu_X + t_2\mu_Y \frac{1}{2}(t_1^2\sigma_X^2 + t_2^2\sigma_Y^2 + 2t_2\sigma_Y t_1\sigma_X\rho)}$$

(d) We obtain the conditional distributions from the joint and marginal distributions. The conditional distribution of X for a fixed value of Y is given by

$$\begin{aligned}f_{X|Y}(x|y) &= \frac{f(x,y)}{f_Y(y)} \\ &= \frac{1}{\sqrt{2\pi}\sigma_X(1-\rho^2)^{1/2}} \exp\left\{-\frac{1}{2\sigma_X^2(1-\rho^2)}\left[x - \mu_X - \frac{\rho\sigma_X}{\sigma_Y}(y - \mu_Y)\right]^2\right\},\end{aligned}$$

where for $f_Y(y)$ we have used the formula obtained in **(b)** with x and y interchanged. As we can see, this is a normal distribution with mean $\mu_X + \frac{\rho\sigma_X}{\sigma_Y}(y - \mu_Y)$ and variance $\sigma_X^2(1-\rho^2)$.

Joint density of functions of random variables

Assume you know the joint density function of two r.v. X_1, X_2 and you want to calculate the joint density of $Y_1 = g_1(X_1, X_2), Y_2 = g(X_1, X_2)$. One can show that

$$f_{Y_1,Y_2} = |J| f_{X_1,X_2}(g_1^{-1}(y_1, y_2), g_2^{-1}(y_1, y_2) I_D(y_1, y_2)$$

where J is the Jacobian i.e. the determinant of the matrix M whose elements are $M_{i,j} = \frac{\partial x_i}{\partial y_j}$ and $|J|$ is its absolute value.

Box-Muller approach

This is a method for generating normally distributed random variables starting from uniformly distributed random variables. Let $X_1 \sim U(0,1), X_2 \sim U(0,1)$

$$y_1 = \sqrt{-2\log x_1} cos(2\pi x_2)$$

$$y_2 = \sqrt{-2\log x_1} sin(2\pi x_2)$$

Then

$$x_1 = \exp\left(-\frac{y_1^2 + y_2^2}{2}\right)$$
$$x_2 = \frac{1}{2\pi} arctg(y_2/y_1)$$

Now

$$\frac{\partial x_1}{\partial y_1} = -y_1 \exp\left(-\frac{y_1^2 + y_2^2}{2}\right)$$
$$\frac{\partial x_1}{\partial y_2} = -y_2 \exp\left(-\frac{y_1^2 + y_2^2}{2}\right)$$
$$\frac{\partial x_2}{\partial y_1} = -\frac{y_2}{2\pi}\frac{1}{y_1^2 + y_2^2}$$
$$\frac{\partial x_2}{\partial y_2} = -\frac{y_1}{2\pi}\frac{1}{y_1^2 + y_2^2}$$

and

$$|J| = \frac{1}{2\pi} \exp\left(-\frac{y_1^2 + y_2^2}{2}\right)$$
$$D = \{-\infty < y_1 < \infty, -\infty < y_2 < \infty\}$$

$$f_{Y_1,Y_2} = \frac{1}{2\pi} \exp\left(-\frac{y_1^2 + y_2^2}{2}\right) I_D(y_1, y_2)$$

Hence Y_1, Y_2 are independent normally distributed random variables. Now

$$Z_1 = Y_1$$

and

$$Z_2 = \rho Y_1 + \sqrt{1-\rho^2} Y_2$$

are normally distributed and ρ correlated.

Limit theorems

The strong law of large numbers Let $X_1, X_2, \cdots X_n$ be a sequence if i.i.d r.v. each having a finite mean $\mu = E[X_i]$. Then

$$P\left\{\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu\right\} = 1$$

Central Limit Theory Let $X_1, X_2, \cdots X_n$ be a sequence if i.i.d r.v. each having a finite mean μ and variance σ^2. Then

$$\lim_{n\to\infty} P\left\{\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \leq a\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{\frac{-x^2}{2}} dx = N(a)$$

Stochastic process

A stochastic process is a sequence of r.v. $X = (X_t(\omega), t \in T, \omega \in \Omega)$ defined on some probability space Ω. T can be a finite set, or countably infinite set (discrete-time process) or an interval (a, b) or (a, ∞) (continuous-time process). The index t is usually refereed to as time. A stochastic process $X_t(\omega)$ is a function of two variables:

- at fixed time t, $X_t(\omega), \omega \in \Omega$ it is a random variable
- for a given random outcome ω, $X_t(\omega), t \in T$ it is a function of time, called a realization or a trajectory, or sample path of the process X.

Random Walk

A random walk is a formalization of the intuitive idea of taking successive steps, each in a random direction. A random walk is a simple stochastic process sometimes called a "drunkard's walk".

To generate a random walk

$$W(t + \Delta t) = W(t) + \sigma\epsilon\sqrt{\Delta t}$$

where $\epsilon \sim N(0, 1)$.

Wiener process or Brownian motion

As the step size Δt in the random walk tends to 0 (and the number of steps increased comparatively) the random walk converges to Brownian motion in an appropriate sense.

A stochastic process $W = (W(t) : t \geq 0)$ is a standard Brownian if

- $W(0) = 0$
- W has independent increments: $W(t+u)-W(t)$ is independent of $(W(s) : s \leq t)$ for $u \geq 0$.
- W has stationary increments: the law of $W(t + u) - W(t)$ depends only on u.
- W has Gaussian increments: $W(t + u) - W(t) \sim N(0, u)$.

Appendix B

A useful formula for option pricing

If V is a log normal variable, i.e. $V = e^X$ with $X \sim N(m, s^2)$

then

$$E[(V-K)^+] = E[V]N(d_1) - KN(d_2)$$

where

$$d_1 = \frac{\ln(E[V]/K) + s^2/2}{s}$$

and

$$d_2 = \frac{\ln(E[V]/K) - s^2/2}{s}$$

Before proving it let's use it and derive the Black-Scholes equation.

We need to calculate

$$C(0) = e^{-rT} E^*[(S_T - K)^+]$$

where $S_T = e^{X_T}$ and $X_T \sim N((r - 1/2\sigma^2)T, \sigma^2 T)$. Also under P^*

$$E^*[S(T)] = S(0)e^{rT}$$

hence replacing in formula above s with $\sigma\sqrt{T}$ we obtain

$$C(0) = S(0)N(d_1) - KB^{-1}(T)N(d_2)$$

with

$$d_1 = \frac{\ln(S(0)e^{rT}/K) + s^2/2}{s} = \frac{\ln(S(0)/K) + (r + 1/2\sigma^2)T}{\sigma\sqrt{T}}$$

and

$$d_2 = \frac{\ln(S(0)e^{rT}/K) - s^2/2}{s} = \frac{\ln(S(0)/K) + (r - 1/2\sigma^2)T}{\sigma\sqrt{T}}$$

Proof of the formula

If $V = e^X$ with $X \sim N(m, s^2)$ then using moment generating formula

$$E[V] = e^{m + \frac{1}{2}s^2}$$

Define

$$Q = \frac{\ln V - m}{s}$$

then

$$V = e^{Qs+m}.$$

Note that $Q \sim N(0,1)$ so

$$n(Q) = \frac{1}{\sqrt{2\pi}} e^{-Q^2/2}$$

and

$$I = E[(V-K)^+] = E[(e^{Qs+m}-K)^+].$$

Note that

$$e^{Qs+m} > K \to Q > \frac{\ln K - m}{s}$$

Hence

$$I = \int_{\frac{\ln K-m}{s}}^{\infty} (e^{Qs+m}-K)n(Q)dQ =$$

$$\int_{\frac{\ln K-m}{s}}^{\infty} e^{Qs+m}n(Q)dQ - K\int_{\frac{\ln K-m}{s}}^{\infty} n(Q)dQ$$

but

$$e^{Qs+m}n(Q) = \frac{1}{\sqrt{2\pi}}e^{-Q^2/2+Qs+m} = \frac{1}{\sqrt{2\pi}}e^{-1/2(Q^2-2Qs-2m)} =$$

$$\frac{1}{\sqrt{2\pi}}e^{-1/2[(Q-s)^2-2m-s^2]} = e^{m+s^2/2}n(Q-s)$$

So

$$I = e^{m+s^2/2}\int_{\frac{\ln K-m}{s}}^{\infty} n(Q-s)dQ - K\int_{\frac{\ln K-m}{s}}^{\infty} n(Q)dQ =$$

$$e^{m+s^2/2}\int_{\frac{\ln K-m}{s}-s}^{\infty} n(Z)dZ - K\int_{\frac{\ln K-m}{s}}^{\infty} n(Q)dQ$$

Remember that

$$\int_a^{\infty} n(u)du = 1 - \int_{-\infty}^{a} n(u)du$$

Also

$$\int_{-\infty}^{a} n(u)du = N(a) = 1 - N(-a)$$

Then

$$I = e^{m+s^2/2}[1-N(\frac{\ln K-m}{s}-s)] - K[1-N(\frac{\ln K-m}{s}) =$$

$$e^{m+s^2/2}N(-\frac{\ln K-m}{s}+s)] - KN(-\frac{\ln K-m}{s})$$

Now replacing

$$m = \ln E[V] - 1/2s^2$$

we obtain

$$E[(V-K)^+] = E[V]N(\frac{-\ln K + \ln E[V] - 1/2s^2 + s^2}{s})$$

$$-KN(\frac{-\ln K + \ln E[V] - 1/2s^2}{s}) =$$

$$E[V]N(\frac{\ln(E[V]/K) + 1/2s^2}{s}) - KN(\frac{\ln(E[V]/K) - 1/2s^2}{s}) =$$

$$E[V]N(d_1) - KN(d_2)$$