Models of Finance

Dominique HENRIET

Ecole Centrale Marseille

Master AMSE

CHAPTER 1

Introduction

The aim of this course is to give some general concepts that found the main models of finance. This in order to first, better understand the characteristics of financial products and second to to present models that allow to describe the main features of financial markets.The purpose of a model is to give a representation of the reality that allows to understand phenomena and to make predictions.

We begin by some naïve questions. Because simple interrogations are often necessary to understand clearly what we talk about.

1.1. Questions

The first question is quite simple : what are financial products designed for?

1.1.1. Real position. The original purpose of financial assets is to transfer wealth through time or, alternatively or simultaneously, to share risk to mitigate it.

The idea is very simple. The economic activity of an individual can lead, to him, to a particular income stream (a so-called real position). This flow can eg be distributed over time in some special and uneven way, or be conditioned by some random events. For example, a seller of glouttes on an international market, knows his income two months after delivery, at a price that depends on the price of some foreign currency at that date. He would like to insure his income immediately, so as to be sure of the amount of money he will get in two months. He his ready to exchange a random cash flow against a sure one. To do so he buys, on a financial market or over the counter (OTC), an asset that does the job : take the risk. Obviously this is not free. But the individual feels adantageous to buy this asset.

Demand for financial assets is hence, at least initially, triggered by a need of risk coverage or transfer wealth over time. But this coverage, or transfer, has itself a price, that can be fixed, for instance on a market on which supply and demand are confronted.

We can hence summarize :

- A financial product first allows to exchange wealth through time : I would like to have one euro on the 24th december (because I'd like to buy a gift for my grand-son). What does it cost today? Here I don't want a bottle of whisky, I want one "euro on the 24th december". The "euro on the

24th december" is a particular good which is different of the "euro on the 1st november" (which is useful to buy grave flowers). These two goods have not the same price today.
- A financial product also allows to exchange risk. I know that I will have a lot of euros if it rains next week (because I sell umbrellas), and nothing if not (I don't sell sun hats). I am ready to pay something today to smoothe my wealth and have some money even when the sun shines.

1.1.2. Market makers. These definitions insist on the real initial position of the demander. But some other agents intervene on this market although they have no initial real position. Their only role is to make the market liquid. To be sure that any demand (justified by a real position), meets an offer. They are called "market makers", in the sense that they supply the demand at a price they define. The price of this supply is obviously strategic.

In some sense, banks, insurance and other institutions are "financial intermediaries" : they offer counterparts to demands (and supplies) originated by real positions.

One purpose of models is to provide methods to "price" financial assets in order the market makers can offer a counterpart to agents.

One first goal of this course is to describe model that allows to "price" assets, that is to find formulas that can give the price of a given asset.

But, more than that, we will also describe models that explain how prices endogeneously arise.

1.1.3. Models and collateral questions. We are going to study several types of models. We can split them in two families :

- the arbitrage models : in these models we use a a particular condition of price equilibrium, the arbitrage free condition. Under this condition we can derive a very important property on the price structure. This is also called the "Risk Neutral Valuation" , RNV, of financial products.
- the behaviour models : in the behaviour models, we try to endogeneize prices. Here we will focus on several hypothesis on individual expectations and on individual information that form supply and demand.

The different models can be static, dynamic, discrete and continuous.

Other questions are, on the way, addressed in this course. One, for instance, is the question of market efficiency, and, information revelation through prices. The question is, as soon as some people are better informed, is that true that this information is immediately transfered to price so that, initially uninformed people are, by so, finally informed.

In the same way we will try to study models that could explain speculative bubbles.

For the first lesson we just describe some characteristics of the market and give two very simple models that can help us to understand complex ones.

1.2. Assets

Financial instruments can be decomposed in three classes :

- debt assets : the issuer of the paper promises a given future cash flow to the holder.
- equity assets : the paper represents a share of a company and gives right to potential (and not known in advance) dividends.
- currencies which are traded on the foreign exchange.

On the top of these basic "papers" one can imagine derivative assets.

- A derivative asset simply gives conditional cash flows that depend on the value of the underlying asset. These derivatives can be traded on a market or on an "Over The Counter" (OTC) basis.

The following table gathers the main different assets :

Asset Class	Instrument type		
	Securities	Exchange-trade derivatives	OTC derivatives
Debt (LT>1year)	Bonds	Bond options or futures	Interest rate swap, options
Debt (ST≤1year)	Bills, Commercial paper	Short term interest rate futures	forward rate agreements
Equity	Stock	Stock options, equity futures	stock options, exotic
FX (foreign exchange)	-	futures	swaps

To these assets, it important to add the markets of commodities (gold, oil, sugar, wheat...) which were the first ones to involve futures.

1.2.1. Equity assets : Stocks. A share is a title of ownership of a share of the capital of a company. At issuance, the value of these securities is determined by society. Then, the securities are traded on a market and the price is set according to supply and demand.

In addition, the holder receives a share of company profits as dividends. hence, from a pure financial viewpoint, a stock is a financial paper that gives right to receive, in the future, cash flows corresponding to the firm profits. Mathematically this amounts to a promise of random future cash flows.

Definition 1. (Notation) A stock is an asset which gives right to dividends $\widetilde{d}(t)$ or ($d(t,\omega)$, where ω is the state of nature) at time t. Stocks are traded on an Exchange platform. Supply an demand determine prices.

- Euronext is the European exchange created in September 2000 by the merger of managing markets stock exchanges of Paris, Brussels (BXS) and Amsterdam (AEX), Stock Exchanges of Lisbon and Porto (BVLP) have been included in this group in 2001, LIFFE (London International Financial Futures and Options Exchange) in 2002. In June 2006, Euronext announced a merger agreement with the NYSE (New York Stock Exchange), merger became effective in April 2007. Nyse-Euronext is a most important trading platform in the world.

- Indexes

An index is generally used to follow the evolution of prices of stocks. An index is generally the cumulated value of all the stocks of a given set of companies.

Most global stock markets offer indexes. The main French stock market indexes are:

- The CAC 40 is the main index of the Paris market. It is calculated from the prices of 40 stocks selected from hundred companies generating a high volume of trade on Euronext Paris.

- The CAC Next 20 includes the next twenty companies.

- The SBF 120 is calculated from the stocks of the CAC 40 and CAC Next 20 and 60 stocks of the first and second market.

- The SBF 250 calculated from 250 companies of all sectors. It is composed of the CAC 40 and CAC Next 20 and the CAC Mid 100 and CAC Small 90.

Obviously let us quote the well known international indexes :

- The Dow Jones Industrial Average. It is the oldest index in the world. It is composed of 30 major industrial stocks of the New York Stock Exchange (blue chips). For historical reasons, its value is the arithmetic average (not weighted by capitalization) assets that compose it.

- The SP 500 lists the 500 largest companies on Wall Street. Its value is more representative than the Dow Jones.

- The Nasdaq Composite Index is an index calculated from all the values of NASDAQ (second equity markets of the United States after NYSE). It contains more than 300 assets, but are not limited to technology.

- The Nikkei 225 index of leading the Tokyo Stock Exchange. Its value is calculated as the Dow-Jones.

- The Footsie (FTSE 100) is the main index of the London Stock Exchange.

- The DAX is the main index of the Frankfurt Stock Exchange consists of 30 blue chip stocks.

1.2.2. Debt : Bonds. A bond is a negotiable instrument issued by a corporation or a public authority for a loan. A Bond is defined by the name of the issuer, interest rate, maturity date, the currency in which it is issued, a periodicity of coupon, the date of issue. Repayment terms and the method of payment of the lenders are contractually fixed, the remuneration may be fixed or variable (indexed to the rate of interest and not on the outcome of the company).

From a mathematical point of view, a fixed rate bond is an asset that promise sure future cash flows.

Definition 2. (Notation) a Bond is associated to a loan issued (by a borrower) at date t_0 for a length (maturity) T. It is defined by the sequence of future "sure" cash flows (for the holder) $d(t)$ (coupons) and, when T is finite, a final payment $h(t_0 + T)$, (in this case $t > T + t_0 \Rightarrow d(t) = 0$) .

- "ultimately" or "In fine" bonds are defined by a face value (nominal or principal) $N$and a nominal rate r_0 per period such that $d(t) = N * r_0$ for $t < t_0 + T$ and $h(t_0 + T) = N(1 + r_0)$.

- Constant annuities are bonds such that $h(t_0 + T) = d(t)$

- A perpetuity is a bond for which $T = +\infty$.

- The nominal zero-coupon at date t_0 with maturity T is bond is such that $d(t) = 0$ and $h(t_0 + T) = N = 1$

Bonds are exchanged on the market. Bonds heve different names according to their maturity :

- short term (bills): maturities between one to five year; (instruments with maturities less than one year are called Money Market Instruments)
- medium term (notes): maturities between six to twelve years;
- long term (bonds): maturities greater than twelve years.

Variable rate bonds, are on contrary a promise of "random" cash flows contingent to some interest rate.

Nominal interest rate, nominal value, and conditions of repayment hence allows the calculation of coupon that the borrower agrees to pay annually (usually in the euro area), quarterly (especially English and American bonds) or with a shorter periodicity. Bond prices are expressed in percentage of the nominal value. Thus, a bond par value 10,000 euros (price cut) is not priced 9900 euros, but 99%. Even if, theoretically, a fixed rate bond is sure, there is a risk of default of the issuer. So, since the bond market is being internationalized, investors need ratings to measure the risk of issuer default.

States issue bonds in order to raise funds on the markets. The main European bonds are OAT (France), Bonos (Spain), Olo (Belgium), Btp (Italy) Guilt (UK), Bund (Germany).

1.2.3. Swap. A swap is a contract in which two counterparties agree to exchange two sets of cash flows (usually debt or currency) between two dates. Unlike trade in financial assets, trade in financial flows are instruments dealed "Over The Counter" OTC without affecting the balance sheet. The synthetic product resulting from the exchange reflects the characteristics sought by the investor. It amounts to accessing a synthetic product not available on the regular market. For example, two parties may agree to pay at predetermined times the difference between a fixed rate and a variable rate.

1.2.4. Forwards and futures. A futures contract is a commitment to buy or sell at a certain maturity T an amount of an asset (securities or commodities) negotiated at the date of commitment, but payable at maturity T .

DEFINITION 3. The forward contract negociated at t_0 for a term T sets at date t_0 a forward price $f(t_0, T)$ for a given quantity of a given good. Hence if $p(t)$ is the spot price at date t of this good, the seller of the forward contract will earn $f(t_0, T) - p(T)$ at T payed by the buyer.

Then we see that a forward contract involves a final payment from one party to the other which amounts to the difference between spot price and forward price.

Futures contracts are traded on regulated markets as opposed to forwards, which are traded "over the counter". The underlying Futures are either physical assets (futures commodity having existed since antiquity) or financial assets (such futures appeared in the early 1970s). Regulated markets (such as the Chicago Mercantile Exchange Group, NYSE Euronext LIFFE) offer standardized futures contracts with respect to the amounts, timing and quality of the underlying assets. They have clearinghouses that serve as intermediaries, so that the buyer has in front of him these clearinghouses as a seller, and vice versa for the seller. At any date the difference $f(t-1,T)-f(t,T)$ is payed by the buyer to the seller. Daily payments between winners and losers are made exclusively through the clearinghouse. Counterparty risk being transferred, the guarantee of futures contracts is thus much more important against the risk of default by the losing party than for contracts OTC (Forwards for which the loss is disbursed at maturity). Futures on physical assets are futures contracts whose underlying assets are commodities : agricultural products (beef, pork, dairy, wheat, mas, soybean, sweet, wood, etc.). , metals (gold, silver, copper, aluminum, zinc, palladium, etc..) and matters relating to energy (gas, oil, coal, electricity). Futures on these physical assets, which enable producers and traders to hedge against the risk of price change, most often give rise to the delivery of the underlying commodity unlike financial assets.

1.2.5. Options.

Definition 4. An option is a contract between two parties and allow one of the parties to ensure, on payment of a premium, the right (but not the obligation) to buy or sell to the other party a particular asset at a predetermined price at the end of a certain period (called European options) or during a certain period (American options). The underlying asset can be a financial asset (stock, bond, treasury bond, futures, currencies, indices, etc..) or a physical asset (agricultural or mineral). The value of the option is the amount of the premium that the option buyer is willing to pay the seller. An option is said to be negotiable if it can be traded on a regulated market. Otherwise, one speaks of OTC trading.

1.2.6. Structured Products. A structured product is a product designed by a bank to meet the needs of its customers. This is often a complex combination of conventional and derivatives. This will, for example, a fixed-rate investment with participation rising or falling prices of a basket of shares. They can take various legal forms and are traded OTC. As a structured product can not be traded on a market, its price is determined using mathematical models reproduces the behavior of the product over time and different market trends. These are often products with high margins.

1.2.7. Foreign exchange. In the foreign exchange market, currencies (money) are exchanged.

1.3. Functioning of trading

1.3.1. Market makers. Some agents have no initial real position. They intervene on the market only as "counterparty" to be sure that "real position" agents will find a counterparty to trade with.

Hence a so called market maker is a company, or an individual, that quotes both a buy and a sell price in a financial instrument. For a given asset a market maker sets hence two prices :

- Ask : the (floor) price they are ready to sell the asset.
- Bid : the price (ceiling) they are ready to buy.

Ask price is larger than bid price. One says that there is a "bid-ask spread". A market maker hence presents simultaneously a supply and a demand curve. Both are horizontal and the demand curve is below the supply one.

There are several factors that contribute to the difference between the bid and ask prices. The most evident factor is a security's liquidity. This refers to the volume or amount of stocks that are traded on a daily basis. Some stocks are traded regularly, while others are only traded a few times a day. The stocks and indexes that have large trading volumes will have narrower bid-ask spreads than those that are infrequently traded. When a stock has a low trading volume, it is considered illiquid because it is not easily converted to cash.

Another explanation of the bid-ask spread Bid ask spread is the asymmetry of information among traders. This is detailed with an example of microstructure models (Glosten-Milgrom). (see chapter on behaviour models).

1.3.2. Orders. When you want to buy or sell an asset you submit an order. There are roughly speaking two kinds of orders : "market orders" (no limit) and limit orders. Market or no limit means that you are ready to sell or buy, whatever the price, a specified quantity of a given asset. Limit order means that you set a ceiling (cap) price for buying or a floor price for selling. There exists also "stop orders" that are inverted limit orders : sell (buy) when the price falls (rises) below (above) some threshold.

- Main trading systems
  - Order-driven : traders place orders before prices are set (either by market makers or by a centralized mechanism or auction).
    - Trading can be continuous or in batches at discrete intervals. In many continuous systems the order submission is against a limit order book where orders have accumulated. Batch auction to open continuous trading (e.g. Paris Bourse, Deutsche Börse, Tokyo Stock Exchange).
  - Quote-driven : market makers set bid and ask prices (i.e. the price at which they are willing to buy and sell the asset) and traders submit orders.
  - Many trading mechanisms feature both systems.

- Order book

Orders are ranked in the following way : highest buy orders are put on the top of the list of buy orders, and are ranked decreasingly. Lowest sell orders are put on the top of sell orders and are ranked increasingly.

Call (v_i, b_i) the ith buy order (ready to buy a quantity b_i for a price less than v_i) and (c_j, s_j) the jth sell order . We have $v_1 \geq v_2 \geq ...v_i \geq v_{i+1}$and $c_1 \leq c_2 \leq ... \leq c_j \leq c_{j+1}$ ($v_1 = +\infty$ in case of no limit buy order, and $c_1 = -\infty$ in case of no limit sell order)

Practically these orders are gathered in the "order book" and ranked in the previous way.

- Trading session

There are three periods : pre-opening and fixing, trading session, closing and fixing. During pre-opening orders are gathered in the "order book".

1.3.3. Order book and opening fixing. An order book gathers sell and buy orders, when the market is closed the order book accumulates orders : high price buy orders and low price sell orders are placed on the top.

Consider for instance this (fictious) order book (real order books involve prices with 2 or even 3 decimals). On the left are buy orders, on the right sell orders.

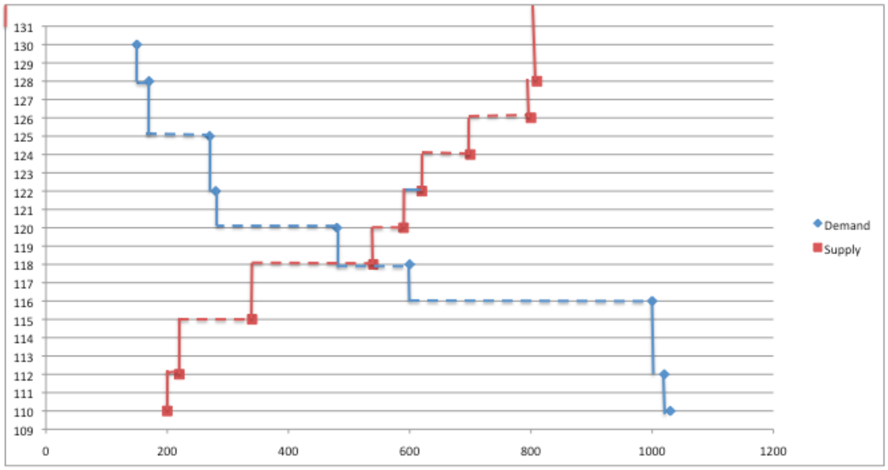
In the following order book the second buy order is a limit price order : ready to buy 50 units at a (cap) price 130.

Cumulative	Quantity	Price	Price	Quantity	Cumulative
100	100	no limit	no limit	150	150
150	50	130	110	50	200
170	20	128	112	20	220
270	100	125	115	120	340
280	10	122	118	200	540
480	200	120	120	50	590
600	120	118	122	30	620
1000	400	116	124	80	700
1020	20	112	126	100	800
1030	10	110	128	10	810

The equilibrium price is 118 : at this price buyers are ready to buy 600 units and sellers 540. That means that only 540 units will be effectively sold at 118. After the fixing, the order book is as follows. Top orders are executed and remain at 118, 60 to buy.

Cumulative	Quantity	Price	Price	Quantity	Cumulative
60	60	118	120	50	50
460	400	116	122	30	80
480	20	112	124	80	160
490	10	110	126	100	260
			128	10	270

On can draw supply/demand curves very easily.



1.3.4. Intraday and closing. Then orders are executed (if there is a counterpart) as they come in the order book.

1.4. Two first models

1.4.1. The miracle of complete and perfect markets : the Risk Neutral Valuation. The main idea of "risk neutral valuation" is that market prices of assets give information on the "perception" of the risk. Give an example.

Assume there are two bonds. The first one is issued by a government : its maturity is one year, it nominal is R_1 =100 euros and its price is $B_1 = 90$. That means that its yield is $r = \frac{100}{90} - 1 = 11,1\%$ (which is very high but very convenient for simple computations).

$$B_1 = 90, 1 + r_1 = \frac{100}{90}$$

The second one is a corporate bond issued by a corpotate, its repay is also 100 euro in one year, but there is a risk of default. In case of default the repay will be only 80. The repay $\widetilde{R}_2$ is hence random. Its nominal yield must hence be larger than the risk-free bond because investors demand a better return to compensate this risk of default.

Obviuously B_2 is larger than $\frac{90}{100}80 = 0,8 \times 90 = 72$: if I buy this bond I am sure to get at least 80 that I can also obtain by buying 0,8 bonds of type 1. But obviuously the price cannot be lager than 90: since 90 allows to get surely 100 at date 1. We can claim :

$$72 \leq B_2 \leq 90$$

Intuitively, B_2 close to 72 means that there is a high "perceived" probability of default. Conversely, if it is closed to 90, this perceived probability is quite low. We can hence write :

$$B_2 = \widehat{\Pi} \times 72 + \left(1 - \widehat{\Pi}\right) 90$$

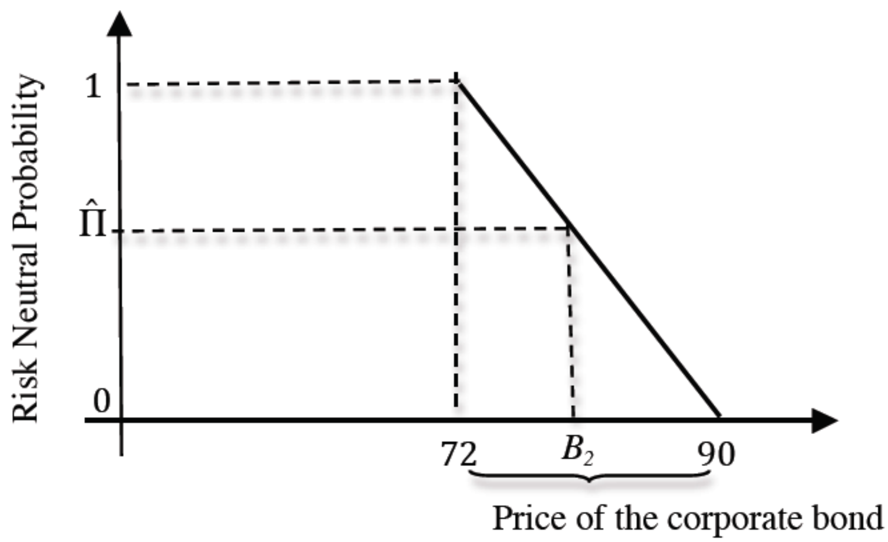
Or equivalently :

$$\widehat{\Pi} = \frac{90 - B_2}{18}$$

Where $\widehat{\Pi}$ is the "perceived probability of default".

We can also write :

$$B_2 = \frac{90}{100}\left[\widehat{\Pi} \times 80 + \left(1 - \widehat{\Pi}\right) 100\right] = \frac{1}{1 + r_1}\mathbb{E}_{\widehat{\Pi}}\left[\widetilde{R}_2\right]$$



Notice that this perceived probability is not necessary the "true" probability (see below). It simply results from the perception of risk by the market. This perceived probability is called : the Risk Neutral Probability.

PROPOSITION 5. *The price of each bond is equal the discounted expected repay under the Risk Neutral Probability (RNP). The discount factor is* $\frac{1}{1+r_1} = \frac{90}{100}$*, and the RNP is* $\widehat{\Pi} = \frac{90-B_2}{18}$

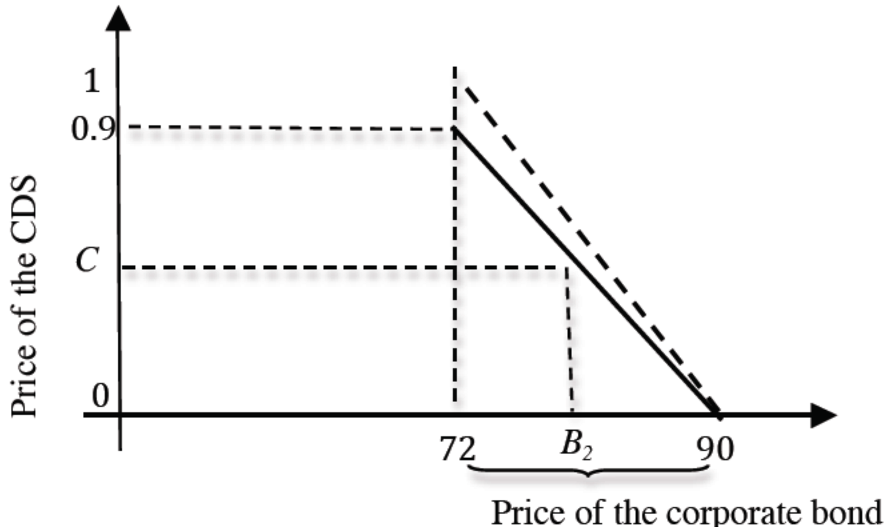
Assume that on top of these two bonds, there is a new financial product : a CDS (credit default swap). This product promises 1 euro in case of default of the corporate. In some sense, this is an insurance product because it indemifies loss due to default. Consider the following portfolio strategy : buy the second bond and immunate risk by buying 20 CDS. This strategy gives exactly 100 euros in one year, and hence is completely identical with the bond 1. The price of this portfolio must then be equal to $B1$.

$$B_2 + 20C = B_1$$

This gives the price of the CDS :

$$C = \frac{B_1 - B_2}{20} = \frac{90}{100}\left[\widehat{\Pi} \times 1 + \left(1 - \widehat{\Pi}\right) \times 0\right]$$

The important point is that the price of the CDS (which is an insurance product) can be computed without knowing the true probability of default, as soon as we know the prices of the underlying asset and the risk free rate (government bond yield)!



We get hence the interesting result :

Proposition 6. *the prices of the three assets are equal to their expected present values computed with the "Risk Neutral Probability".*

This result is the very basis of all the "pricer" models developped in stochastic finance.

Proposition 7. *Models based on "Risk Neutral Valuation" rely on the property that there must exist a probability distribution (Risk Neutral Probability) such that the price of any existing or composite asset is equal to the expected present value of its future cash flow.*

Mathematically, things appear quite simple. We are in a world where there are only two states of nature : default of the corporate, no default of the corporate. The repay in one year can be described by a (line) two components vector, each component being the repay in the corresponding state of nature. For the risk-free bond it is $d_1 = (100, 100)$ whereas it is $d_2 = (80, 100)$ for the second bond and $d_{CDS} = (1, 0)$ for the CDS. Clearly these 3 vectors are not independent since one of them is a linear combination of the two others :

$$d_1 = d_2 + 20 d_{CDS}$$

Let A the matrix of the repay at date 1 :

$$A = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} 100 & 100 \\ 80 & 100 \end{bmatrix}$$

A composite asset (or a portfolio) is a (column) vector θ, ${}^t\theta = (\theta_1, \theta_2)$ of quantities of the assets. A portfolio θ gives the repay :

$$^t\theta A$$

We say that the market is complete in the sense that any (line) vector of repay $d = (a, b)$ can be obtained with a portfolio. Indeed, A being regular we can write:

$$^t\theta A = d \Rightarrow^t \theta = dA^{-1}$$

The price of the portfolio giving d is hence :

$$P =^t \theta B = dA^{-1}B = \left(A^{-1}B\right)_1 a + \left(A^{-1}B\right)_2 b$$

But there are two interesting vectors of repay : $d^{(1,0)} = (1, 0)$ (which is precisely the CDS) and $d^{(0,1)} = (0, 1)$ (which is a kind of a call option). The first one is equal to $\frac{d_1 - d_2}{20}$, that is the portfolio $\left(\frac{1}{20}, -\frac{1}{20}\right)$, and the second to $\frac{1}{20}\left(d_2 - \frac{80}{100}d_1\right)$, that is the portfolio $\left(-\frac{4}{100}, \frac{1}{20}\right)$.

The “arbitrage free “ hypothesis implies that the two prices of these two particular portfolios must be positive (since these two portfolios give positive returns at date 1).

This means that the price of any portfolio is a positive linear combination of the repays.

This implies that $\left(A^{-1}B\right)_1 = \frac{B_1 - B_2}{20}$ and $\left(A^{-1}B\right)_2 = \frac{1}{20}\left(B_2 - \frac{80}{100}B_1\right)$ are positive. But their sum is equal to $\frac{B_1}{100} = \frac{1}{1+r_1}$. So that we find the RNP :

$$\widehat{\Pi} = \frac{100}{B_1}\left(A^{-1}B\right)_1 = (1 + r_1)\left(A^{-1}B\right)_1 = \frac{100}{90}\left(\frac{90 - B_2}{20}\right) = \frac{90 - B_2}{18}$$

RNP versus objective probability

Is the RNP the “true” probability of default? Assume for instance that $B_2 = 81$ so that the RNP is $\widehat{\Pi} = \frac{1}{2}$. And consider two portfolios : the first one with 10 corporate bonds and the second one with 9 government bonds. The values of both portfolios is 810. But one is risky and the second is not. Intuitively the first one must have a higher “objective” expected return (a risk premium) :

$$10 \times [\Pi \times 80 + (1 - \Pi) \times 100] \geq 9 \times [\Pi \times 100 + (1 - \Pi) \times 100]$$

That is :

$$\Pi \leq \frac{1}{2} = \widehat{\Pi}$$

The following model will make things more precise by relating the true probability and the RNP

1.4.2. Supply, demand and risk aversion : where the risk neutral probability comes from. Consider a very simple “market” where there are only two individuals A and B. A is an entrepreneur. His activity provides him a random net income (profit). There are two “states of nature” (events) : the good state where the income is large W_H and the bad state where it is low $W_L < W_H$. B is a lady who has a constant income w. In this economy the total income is hence $W_H + w$ or $W_L + w$. A has a real initial position : he is exposed to a risk and would like to mitigate it, that is to insure his income. What means insure? That means that he would like to have a final income X such that X_H (final income when W_H) is not very different than X_L (final income when W_L). He can try to trade with B by selling him some share of his firm. If he sells a share α at a price S we will have the following final situation

- A will have $X_L = \alpha S + (1-\alpha)W_L = W_L + \alpha\,(S - W_L)$ and $X_H = \alpha S + (1-\alpha)W_H = W_H + \alpha\,(S - W_H)$
- And B $Y_L = w - \alpha\,(S - W_L)$ and $Y_H = w - \alpha\,(S - W_H)$

In somme sense, by doing so, A decreases the risk he is exposed to : the gap between high and low income is $X_H - X_L = (1-\alpha)(W_H - W_L) \leq W_H - W_L$

For a given price, S, we define the supply $\alpha^S(S)$ being the share that A is ready to offer, and the demand $\alpha^D(S)$, as the share, B is ready to buy at this price.

What are the equilibrium values S^* of S? More precisely, which values of S are such that the demand can be equal to the supply?

First of all we necessarily have $W_L \leq S^* \leq W_H$.

Indeed $S^* < W_L$ is surely not an equilibrium : for $\alpha \geq 0$ this price would give him an income always lower than in the initial position! His supply cannot be positive : at that price he would be ready to BUY supplementary shares. B at this price is also ready to buy since it gives her always a better income than her initial position.

In the same way, $S^* > W_H$ cannot be an equilibrium price because $\alpha \geq 0$ would give her an always lower income.

S	$S < W_L$	$W_L \leq S \leq W_H$	$W_H < S$
A's supply	0	≥ 0	> 0
B's demand	> 0	≥ 0	$= 0$

PROPOSITION 8. *if S^* is an equilibrium price, then* $W_L \leq S^* \leq W_H$, or equivalently there exists $0 \leq q_L \leq 1$, $0 \leq q_H \leq 1$ / $q_H + q_L = 1$ and $q_H W_H + q_L W_L = S^*$. Or equivalently :there existsts q_H $0 \leq q_H \leq 1$ /$S^* = W_L + q_H(W_H - W_L)$. *The weight can be interpreted as probabilility : this is the Risk Neutral Probability.*

But obviously, there are a lot of candidates for S^*. The equlibrium value of S depends on A and B "behaviour". At which price S A is ready to sell α and B ready to buy α?

At this time of the reasonning, there is an interesting remark. When you look at these two equations, $q_H + q_L = 1$ and $q_H W_H + q_L W_L = \pi$ where $0 \leq q_L \leq 1$, $0 \leq q_H \leq 1$, you have the idea to replace q_H and

q_l by the probabilities π_H and π_L of the two sates of nature. If you do so S is simply equal to the expected value of the corporate! Is that an equilibrium price?

There is an important property if $S = \mathbb{E}_\pi(W)$: whatever α, the expected income is not changed for both individuals :

$$\pi_H X_H + \pi_L X_L = \pi_H (\alpha S + (1-\alpha) W_H) + \pi_L (\alpha S + (1-\alpha) W_L) = \alpha \mathbb{E}_\pi(W) + (1-\alpha)\mathbb{E}_\pi(W) = \mathbb{E}_\pi(W)$$

$$\pi_H Y_H + \pi_L Y_L = \pi_H (w - \alpha S + \alpha W_H) + \pi_L (w - \alpha S + \alpha W_L) = w - \alpha S + \alpha (\pi_H W_H + \pi_L W_L) = w$$

Look at B. She starts with a sure income w if she buys α shares at price $S = \mathbb{E}_\pi(W)$ she will have a random income with an unchanged expected value. So, at this price, every $\alpha \neq 0$ increases the risk borne by B without increasing the average income. We will assume that B is strictly risk-averse, in the sense that she would accept to take a risk only if her expected income is strictly increased. If this is the case, she will not accept any trade at the price $S = \mathbb{E}_\pi(W) : \alpha^D(\mathbb{E}_\pi(W)) = 0$.

Look at A. must be enclined to sell. Indeed, if we assume that A is himself risk-averse, he will surely offer $\alpha = 1$ if the price is $\mathbb{E}_\pi(W)$: by doing so he would get the same average income and cancell the risk.

Hence $\mathbb{E}_\pi(W)$ is not an equilibrium price since, $0 = \alpha^D(\mathbb{E}_\pi(W)) \neq \alpha^S(\mathbb{E}_\pi(W)) = 1$

So, if we want B buy some share, the price must be such that her expected income is increased : we must have $S^* < \mathbb{E}_\pi(W)$ or equivalently $q_H < \pi_H$.

Proposition 9. *When traders are risk-averse, The Risk Neutral probability q is such that $q_H < \pi_H$ (and $q_L > \pi_L$).*

We feel that we need to model the attitude towards risk of A and B to model her demand in the risky asset.

- Risk aversion

How can we be more precise on risk aversion? The idea is very simple : one individual is risk-averse if he accepts to take risk only if his average income icreases. Equivalently, average income being unchanged, a risk-averse individual always prefers when there is no risk at all.
Take an individual with a sure income w_0. Assume he is proposed the choice between two options : can stay with this income or take a risk that can increase or decrease his income : with a probability $1/2$ he gets an income of $w_0 + +a + x$ and with probability $1/2$ he gets $w_0 + a - x$ where $a \geq 0$ and $x \neq 0$. The individual has hence the choice beween two "lotteries" :

	$1/2$	$1/2$
lottery S	w_0	w_0
lottery R_a	$w_0 + a + x$	$w_0 + a - x$

Definition 10. We say that the individual is risk averse if , for all x, he prefers the lottery S to the lottery R_0 (where $a = 0$).
In order to make R_a more attractive it is necessary to propose the player a lottery with a positive value of a. The value of a such that R_a is equivalent to the sure lottery is a kind of risk premium: the expected value must be increased to compensate risk.

One way to capture risk aversion is to assume that the individual values lotteries by computing the expected value of som concave function u of the payments :

$$\text{Utility of S} = \tfrac{1}{2}u(w_0) + \tfrac{1}{2}u(w_0) = u\left(w_0\right)$$

and

$$\text{Utility of } R_a = \tfrac{1}{2}u(w_0 + a - x) + \tfrac{1}{2}u(w_0 - x + a)$$

Proposition 11. *If u is continuous, increasing and concave then there exists $\Pi \geq 0$ such that $a \geq \Pi \Longleftrightarrow$ Utility of $S \leq$ Utility of R_a.*

Indeed : if u is concave then $\frac{1}{2}u(w_0+x+a)+\frac{1}{2}u(w_0-x+a) \leq u\left(\frac{1}{2}(w_0 + x + a) + \frac{1}{2}(w_0 - x + a)\right) = u(w_0+a)$
For $a = 0$, the utility of R is lower than that of S, for $a = x$, it is larger.

- Risk sharing

Selling shares to B, amounts to risk sharing between A and B. A decreases income risk whereas B increases it.
Just look at the values of utilities of the final wealth of our two individuals. For A, the utility if he sells α at the price S is :

$$U_A(S, \alpha) = \pi_H u_A\left(\alpha S + (1 - \alpha)W_H\right) + \pi_L u_A\left(\alpha S + (1 - \alpha)W_L\right)$$

For B :

$$U_B(S, \alpha) = \pi_H u\left(w - \alpha S + \alpha W_H\right) + \pi_L u_A\left(w - \alpha S + \alpha W_L\right)$$

What is the supply function of A? In other words, how many shares is he ready to sell if the price is π? The answer is, the value of α that maximizes U_A for the given value of π.
We have hence, computing the derivative of U_A with respect of α:

$$\pi_H u'_A(X_H)(S - W_H) + \pi_L u'_A(X_L)(S - W_L) = 0$$

The same reasonning for B :

$$\pi_H u'_B(Y_H)(W_H - S) + \pi_L u'_B(Y_L)(W_L - S) = 0$$

The final allocation is an equilibrium if the supply is equal to the demand :

Claim 12. X_H^*, X_L^*, Y_H^*, Y_l , S^* and α^* are at equilibrium if :

$$\begin{cases} \pi_H u'_A(X_H^*)(S^* - W_H) + \pi_L u'_A(X_L^*)(S^* - W_L) = 0 \\ \pi_H u'_B(Y_H^*)(W_H - S^*) + \pi_L u'_B(Y_L^*)(W_L - S^*) = 0 \\ X_H^* + Y_H^* = w + W_H \\ X_L^* + Y_L^* = w + W_L \\ X_H^* = \alpha^* S^* + (1 - \alpha^*) W_H \\ X_L^* = \alpha^* S^* + (1 - \alpha^*) W_L \end{cases}$$

The first equation is quite interesting : we can rewrite it :

$$S^* = \pi_H \frac{u'_A(X_H^*)}{\pi_H u'_A(X_H^*) + \pi_L u'_A(X_L^*)} W_H + \pi_L \frac{u'_A(X_L^*)}{\pi_H u'_A(X_H^*) + \pi_L u'_A(X_L^*)} W_L$$

The second one is very similar :

$$S^* = \pi_H \frac{u'_B(Y_H^*)}{\pi_H u'_B(Y_H^*) + \pi_L u'_B(Y_L^*)} W_H + \pi_L \frac{u'_B(Y_L^*)}{\pi_H u'_B(X_H^*) + \pi_L u'_B(X_L^*)} W_L$$

We immediately see that π is somewhere in between W_L and W_H but lower than $\mathbb{E}_p W = p_H W_H + p_L W_L$ as soon as $X_H \geq X_L$. Indeed in this case, by concavity of u, $u'(X_H) \leq u'(X_L)$.

Proposition 13. *At equilibrium the price is* $S^* = q_H W_H + q_L W_l$ *with :*

$$q_H = \pi_H \frac{u'_A(X_H^*)}{\pi_H u'_A(X_H^*) + \pi_L u'_A(X_L^*)} = \pi_H \frac{u'_B(Y_H^*)}{\pi_H u'_B(Y_H^*) + \pi_L u'_B(Y_L^*)}$$

$$q_L = \pi_L \frac{u'_A(X_L^*)}{\pi_H u'_A(X_H^*) + \pi_L u'_A(X_L^*)} = \pi_L \frac{u'_B(Y_L^*)}{\pi_H u'_B(X_H^*) + \pi_L u'_B(X_L^*)}$$

Exercise 14. Solve the above model with CARA utility functions : $u_i(s) = -\exp -\rho_i s$.

CHAPTER 2

Static model : arbitrage free condition

In this chapter we propose a generalization in a static case of the first small model we have described above. The main result is the following :

> The arbitrage free condition implies :
> There exists some positive weight such that the prices of existing assets and the price of any asset that can be built with existing assets (portfolios) is simply equal to the weighted sum of its cash flows! If the market is "complete", i.e. any "cash flow profile" can be obtained through some portfolio, this weighting vector is unique. This principle is the basis that we call "risk neutral valuation".

2.1. No arbitrage condition in a static model

2.1.1. Arbitrage. Consider the very simple model with two date : date 0 and date 1. At date one, there are several "states of naure", these states of nature capture the fact that, at date 0, future is not known, but the various possibilities are known. Let Ω the (finite) set of possible states of nature at date 1. We set $\mathrm{Card}\Omega = N$.

At date 0 several assets are available. The asset i gives rise to a payment, at date 1, $d_i(\omega)$ in the state ω. Let p_i the price at date 0 of this asset : paying p_i allows to get $d_i(\omega)$ at date 1 in the state ω.

We note $\theta = (\theta_1, \theta_2, ..., \theta_K)$ a portfolio, θ_i is the quantity of asset i owned by an investissor. The price of a portfolio is $\sum_{i=1}^{K} p_i\theta_i$ and its payment at date 1 $\sum_{i=1}^{K} \theta_i d_i(\omega)$ in the state ω.

We define an arbitrage porfolio in the following manner :

Definition 15. $\theta \neq 0$ is an arbitrage portfolio if $W_0(\theta) = -\sum_{i=1}^{K} p_i\theta_i > 0$ and $W_1(\theta, \omega) = \sum_{i=1}^{K} \theta_i d_i(\omega) \geq 0$ for all ω.

Let d_i the raw vector $[d_i(\omega_1), d_i(\omega_2), ..., d_i(\omega_N)]$ and D the matrix $\begin{bmatrix} d_1 \\ . \\ d_K \end{bmatrix}$. The row vector of payment at date 1 is $W_1(\theta) = ^t\theta D$ and at date 0 $W_0(\theta) = -^t\theta p$.

In the above definition, W_t , $t = 0$ or 1, is the flow of income at date t. If there exists an arbitrage portfolio, that means that one can make money at date 0 without any counterpart at date 1. One says that an arbitrage portfolio "is a free lunch". The no arbitrage condition implies that there does not exist arbitrage portfolios. If the market is at equilibrium, then there cannot be "free lunch". We will assume, hence, that there does not exist arbitrage portfolios.

A first very clear and obvious consequence of such an assumption is that if two portfolios give the same returns then they must have the same price. Indeed, assume that two portfolios θ and θ' are such that ${}^t\theta.D ={}^t \theta'.D$ then if we had ${}^t\theta' p <{}^t \theta p$ then $\theta' - \theta$ would be an arbitrage portfolio.

This have a very practical consequence :

Definition 16. "arbitrage pricing", that is computing the price of a given "new" asset, consists in finding a portfolio of the existing assets that gives the same return. On says that the portfolio replicates the asset.

But this is possible if one can find a portfolio that gives the same return. So, a very important case must be enhanced : the case of complete markets where any date 1 cash-flow profile can be obtained through existing assets :

Definition 17. The market is complete if any new asset (that is any profile of payments $W_1(\omega)$, $\omega \in \Omega$ can be replicated :

$$\forall W_1(\omega),\ \exists \theta \in \mathtt{R}^K, \sum_{i=1}^{K} \theta_i d_i(\omega) = W(\omega)$$

In matrix notations : $\forall W_1 \in \mathbb{R}^N,\ \exists \theta \in \mathbb{R}^K, {}^t \theta D = W_1$

In a complete market there must be as many assets as numbers of states of nature : $K \geq N$. Moreover the linear application $\theta \longmapsto W_1(\theta)$ mapping $\mathbb{R}^K$ in $\mathbb{R}^N$ must be surjective, so that one can extract a $N \times N$ regular matrix A from D. So that the portfolio associated to W is ${}^t\theta = WA^{-1}$.

We have the very important following proposition that characterizes markets where there is no arbitrage possibility.

Proposition 18. *There does not exist arbitrage portfolio, in other words the market is "arbitrage free", if and only if prices and payments are such that there exists a family q of positive numbers $q(\omega)$ such that :*

$$\forall i,\ p_i = \sum_{\omega \in \Omega} q(\omega) d_i(\omega)$$

The coefficients $q(\omega)$ are called state pseudo-prices. Moreover, if the market is complete then q is unique.

The proof of this proposition is quite simple when the market is complete. Indeed, if the market is complete there always exist q such that $\forall i,\ p_i = \sum_{\omega\in\Omega} q(\omega) d_i(\omega)$, that is $p = Dq$, or after extracting the regular matrix A, $p = Aq$. Obviously, in that case we must have $q = A^{-1}p$. But then, the ith component, $q(\omega_i)$ is simply the price of the asset giving 1 in state ω_i and 0 otherwise! If the market is arbitrage free this price cannot be negative.

When the market is not complete, the proof is more sophisticated and can be found is the appendix of the chapter.

2.1.2. Pricing. It is interesting to use the formula above to compute the price of a portfolio. Indeed, we have for a portfolio θ :

$$-W_0(\theta) =^t \theta p = \sum_i \theta_i \sum_\Omega q(\omega) d_i(\omega) = \sum_\Omega q(\omega) W_1(\theta, \omega)$$

That means that the value at date 0 of the portfolio is simply the weighted value of its payments computed with the weights q.

This formula is the one one uses when a financial product can be obtained by a combination of existing assets (one says replicated).

When there is a risk-free asset whose yield is r, the no arbitrage condition gives :

$$1 = \sum_\Omega q(\omega)(1+r)$$

Let then $\widehat{q}(\omega) = (1+r)q(\omega)$ we have :

PROPOSITION 19. *If there exists a risk-free asset, then for any portfolio we have :*

$$-W_0(\theta) = \theta.p = \frac{1}{1+r} \sum_\Omega \widehat{q}(\omega) W_1(\theta, \omega)$$

where

$$\sum_\Omega \widehat{q}(\omega) = 1$$

$\widehat{q}(\omega)$ *is called a "risk-neutral" probability of the event* ω *and the formula above reads : the value of the portfolio is the discounted expected value of its cash flow (under some risk-neutral probability). Note that* $\widehat{q}(\omega)$ *is not uniquely defined if the market is not complete.*

If the market is complete, then any income flow $W(\omega)$ at date 1 can be priced :

PROPOSITION 20. *In a arbitrage free complete market, there exists a unique family of positive weights $q(\omega)$ such that the value at date 0 of an income flow $W_1(\omega)$ at date 1 writes :*

$$P = \sum_{\Omega} q(\omega) W_1(\omega)$$

Tf ther is a risk-free asset with yield r : then the formula writes :

$$P = \frac{1}{1+r} \sum_{\Omega} \widehat{q}(\omega) W_1(\omega) = \frac{1}{1+r} E_{\widehat{q}}[W_1]$$

Where $\widehat{q}(\omega) = (1+r)\, q(\omega)$ is the Risk-Neutral Probability.

One says that the value of an asset is equal to the the expected discounted value of its cash flows computed with the risk-neutral probability.

2.1.3. Example : The model of Cox Ross and Rubinstein with 2 dates. In this model there are two assets and two states of nature $\Omega = \{u, d\}$. The risk free asset has a yield r. The other asset is risky. Its price at date 0 is S_0 and the cash obtained at date 1 is either low , $S_1(d)$ or high, $S_1(u)$.

The equations giving the weigths or state prices are hence (we write that present prices are the weighted sums of future values) :

$$1 = (1+r)q(u) + (1+r)q(d)$$

$$S_0 = q(u)S_1(u) + q(d)S_1(d)$$

There equations have positive solutions if and only if :

$$\frac{S_1(d)}{S_0} \leq 1 + r \leq \frac{S_1(u)}{S_0}$$

And the market is complete if $S_1(d) \neq S_1(u)$.

Finally the state prices are :

$$(1+r)q(u) = \frac{1 + r - \frac{S_1(d)}{S_0}}{\frac{S_1(u)}{S_0} - \frac{S_1(d)}{S_0}} = \frac{(1+r)\, S_0 - S_1(d)}{S_1(u) - S_1(d)}$$

$$(1+r)q(d) = \frac{\frac{S_1(u)}{S_0} - (1+r)}{\frac{S_1(u)}{S_0} - \frac{S_1(d)}{S_0}} = \frac{S_1(u) - (1+r)\, S_0}{S_1(u) - S_1(d)}$$

2.1.4. Optimal hedging. Assume a financial institution sells a financial product which promises $W_1(\omega)$ at date 1. Assume the price is P. What is the optimal strategy (hedging and pricing) such that the bank will be able to meet his commitment. As he receives P he can buy Δ risky assets and invest the remaining $P - \Delta S_0$ on the risk-free asset. This strategy gives :

$$\begin{cases} W_{\triangle P1}(d) = (1+r)\,(P - \triangle S_0) + \Delta S_1(d) \\ W_{\triangle P1}(u) = (1+r)\,(P - \triangle S_0) + \Delta S_1(u) \end{cases}$$

In order to be hedged, we must have $W_1(\omega) = W_{\triangle P1}(\omega)$ for both $\omega = u$ and $\omega = d$. This gives :

$$\begin{cases} \Delta = \frac{W_1(u) - W_1(d)}{S_1(u) - S_1(d)} \\ P = \frac{1}{1+r}\left(\frac{S_1(u)-(1+r)S_0}{S_1(u)-S_1(d)} W_1(d) + \frac{(1+r)S_0 - S_1(d)}{S_1(u)-S_1(d)} W_1(u)\right) \end{cases}$$

The above solution gives the price as expected (as the weighted sum of future values) and also the so called Δ hedging : the quantity of assets that must be bought to immunate risk. We can remark that we have also :

$$\Delta = \frac{\partial P}{\partial S_0}$$

2.2. Mathematical appendix

To derive the main result on no arbitrage theory we need some maths. The following proposition is known as "Convex separation theorme". It tells us that if we consider a point outside a closed convex set, one can separate it from this set by an affine hyperplane.

2.2.1. Convex separation tehorem.

PROPOSITION 21. *in an Hilbert space H (vect space with a norm coming from a scalar product, and complete), if C is a closed convex subset of H, x in H but not in C, then there exists h in H such that , $\forall y \in C, h.x < h.y$*

PROOF. let $P(x)$ the projection of x on C. $P(x) = \arg\min_{y \in C} \| x - y \|$.

It can be shown that this projection exists and is unique :

- it indeed exists :let $d = \inf_{y \in C} \| x - y \|$; obviously, $d > 0$.

Take a sequence y_n in C such that $d^2 \leq \| x - y_n \|^2 \leq d^2 + 1/n$

we have (by the equality of the median) : $\| x - \frac{(y_n + y_m)}{2} \|^2 + \frac{1}{4} \| y_n - y_m \|^2 = \frac{1}{2} \| y_n - x \|^2 + \frac{1}{2} \| y_m - x \|^2$

$\frac{(y_n + y_m)}{2} \in C$ (convex)

we have hence $\| x - \frac{(y_n + y_m)}{2} \|^2 \geq d^2$,and hence : $\frac{1}{4} \| y_n - y_m \|^2 \leq \frac{1}{2n} + \frac{1}{2m}$.

The sequence y_n is a Cauchy seq and then converges towards y which lies in C (closed) and is such that $\| x - y \| = d$

- Unique : let to the reader

We have :

$\forall y \in C \ (x - P(x)) \, .(y - P(x)) \leq 0,$

indeed $\| ty + (1-t)P(x) - x \|^2 = \| t\,(y - P(x)) + P(x) - x \|^2 = t^2 \| y - P(x) \|^2 + 2t\,(y - P(x))\,.\,(P(x) - x) + \| P(x) - x \|^2$

as $\| ty + (1-t)P(x) - x \|^2 \geq \| P(x) - x \|^2$ (since $P(x)$ is the closest y in C)

we have $t \| y - P(x) \|^2 + 2\,(y - P(x))\,.\,(P(x) - x) \geq 0$ for all t in [0,1] and y in C.

And hence $(y - P(x))\,.\,(P(x) - x) \geq 0$

take then $h = (P(x) - x)$we have $\forall y \in C \ h.(P(x) - y) \leq 0$, that is $h.(h + x - y) \leq 0$,that is $h.y \geq h.x + \| h \|^2$

that is $h.y > hx$ □

2.2.2. proof of proposition. It can be shown that the subset C of $\mathtt{R}^K$ of vector such that $v \in C \iff \exists \lambda \in \mathtt{R^+}^N, v_i = \sum_\Omega \lambda(\omega) d_i(\omega)$, is a closed (not trivial) convex (trivial) set.

Assume then that $p \notin C$.

That would mean that Proposition 14 does not hold.

If so the convex separation theorem says that there exists h in $\mathtt{R}^K$ such that $\forall v \in C, h.p < h.v$.

At this step it is important to remark that this implies $\forall v \in C, h.v \geq 0$. Indeed if it were not the case there would exist v_0 in C such that $h.v_0 < 0$, as $kv_0 \in C$ for all $k \geq 0$, $h.(kv_0) = k(h.v_0)$ would go to $-\infty$ for k going to $+\infty$ but that woud contradict $h.p < h.(kv_0)$.

Moreover, we have $\inf_C (h.v) = 0$ since $0 \in C$. This implies $h.p < 0$. So h is an arbitrage portfolio.

CHAPTER 3

Dynamics (finite discrete models)

In this chapter we generalize the previous model to an intertemporal framework. Here we will assume that the horizon is finite, T, and that time is discrete $t = 0, .1 \ldots T$. In the same way, uncertainty is modelled through finite sets of states of nature. In this framework we will derive the main result : the price of an asset is equal to the expected present discounted value of its payments. The expectation being computed with some "risk neutral" probability.

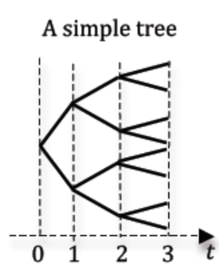
3.1. The tree of states of nature

We will model uncertainty as following. At date 0 there is one unique state of nature. At date 1 several states of nature are possible

$\Omega_1 = \left\{\omega_1^1, \cdots \omega_1^N\right\}$, for each state of nature ω_1^i of Ω_1, several sates of nature are possible at date 2, and so on up to T. Call Ω_t the set of possible states of nature at date t . We can define a "successor" relation between these states of nature.

Definition 22. At each date the set of states of nature is Ω_t. We say that ω_{t+1} in Ω_{t+1} is a successor of ω_t in Ω_t, $\omega_{t+1} > \omega_t$, if ω_{t+1} is one of the possible states at date $t+1$ if the realized state of nature at date t is ω_t.

In a tree structure, for all t (except $t = 0$) , each state of nature is the sucessor of one unique state of nature.



We have hence the following lemma :

Lemma 23. *for all t and t' , $t < t'$, for all state $\omega_{t'}$ there exists a unique state ω_t at t such that there is a successor path between ω_t and $\omega_{t'}$. (Past is perfectly known!).*

We can define partial information on this tree structure. The idea is that if one has some information, this information cannot be forgotten.

DEFINITION 24. Information and Filtration. An information structure $\mathcal{P}_t$ at date t is a partition of Ω_t. There is perfect information if $\mathcal{P}_t$ is composed by all the singletons $\left\{\omega_t^i\right\}$. There is no information if $\mathcal{P}_t = \{\Omega_t\}$. In the continuous case it will be more convenient to define information through σ-algebras. Here the corresponding σ-algebra $\mathscr{F}_t$ will be composed by all the sets obtained by complementations of the sets of $\mathcal{P}_t$ and the $\emptyset$.

We say moreover that $(\mathcal{P}_t)_{t=0\ldots T}$ is a filtration if, for all t, all A_t in $\mathcal{P}_t$, all ω_t and ω_t' in A_t the predecessors of ω_t and ω_t' belongs to the same set in the partition $\mathcal{P}_{t-1}$.

On this tree structure we can define a process simply as a mapping that takes values at the different states of nature.

DEFINITION 25. A process X is a mapping that takes real values (potentially mulidimensional) on the tree. We note $X_t(\omega_t)$ the value of the process at date t in the state ω_t.

EXAMPLE 26. In the Cox, Ross Rubinstein Model at each date, the price of a stock can be multiplied either by u or or by d . $S_t = \varepsilon.S_{t-1}$ with $\varepsilon = d$ or u. Here a state of nature at date t is given by a sequence on t digits $\omega_t = (\varepsilon_1, \varepsilon_2, \cdots \varepsilon_t)$ each digit being d or u. There are exactly 2^t states of nature at date t.

3.2. Stochastic process on a tree

We can add a probability structure on the tree. There are two differents ways to do so. The first one consists in defining, for each state ω_t the probability transition $\pi(./\omega_t)$ defined in Ω_{t+1} :

DEFINITION 27. For each ω_t in Ω_t the probability transition $\pi(./\omega_t) \geq 0$ gives the probability of each state in Ω_{t+1} knowing that the realized state at t is ω_t.

$\pi(\omega_{t+1}/\omega_t) = 0$ if ω_{t+1} is not a successor of ω_t.

$\sum_{\omega_{t+1} > \omega_t} \pi(\omega_{t+1}/\omega_t) = 1$

It is then very easy to compute the probability of any state $\omega_{t'}$ at any anterior date in any predecessor state.

LEMMA 28. *The probability $\pi(\omega_{t'}/\omega_t)$ where $t' > t$ and ω_t the unique state at t which is predecessor of $\omega_{t'}$ is :*

$\pi(\omega_{t'}/\omega_t) = \pi(\omega_{t'}/\omega_{t'-1}) \times \pi(\omega_{t'-1}/\omega_{t'-2}) \times \cdots \times \pi(\omega_{t+1}/\omega_t)$ *where* $\omega_{t'} > \omega_{t'-1} > \omega_{t'-2} > \cdots > \omega_{t+1} > \omega_t$ *is the unique backward path from* $\omega_{t'}$ *to* ω_t.

The "ex ante" (at date 0) probability of a state is hence :

$\pi(\omega_t) = \pi(\omega_t/e_0) = \pi(\omega_t/\omega_{t-1}) \times \cdots \times \pi(\omega_1/\omega_0)$ *where* $\omega_t > \omega_{t-1} > \cdots > \omega_1 > \omega_0$ *is the unique backward path from* ω_t *to* ω_0.

The other way is to give the probability of the terminal states $\pi(\omega_T)$. Then the probability of any state $\pi(\omega_t)$ is simply the probability of subset of terminal states that are successors of ω_t :

$$\pi(\omega_t) = \sum_{\omega_T > \omega_t} \pi(\omega_t)$$

3.3. No arbitrage condition on a dynamic model

In this paragraph we want to derive the no arbitrage condition. We suppose that there exists K assets. These assets are defined by their dividend process $d_i(\omega_t)$ which gives for each sate ω_t at date t the cash flow the owner earns. We simply define a portfolio strategy as a mapping that gives at each date and state the quantity of assets owned in the portfolio. These quantity can be negative in the sense that short selling is allowed. The no arbitrage condition will give a condition on prices.

DEFINITION 29. We note $\theta_i(\omega_t)$ the quantity of asset i owned if the state is ω_t. The price of asset i in the state ω_t is noted $p_i(\omega_t)$.

To a portfolio strategy is associated a cash flow and a value.

The cash flow is simply composed of dividends and the net product of transactions :

$$W_\theta(\omega_t) = \sum_i \left(\theta_i(\omega_{t-1})d_i(\omega_t) + p_i(\omega_t)\left(\theta_i(\omega_{t-1}) - \theta_i(\omega_t)\right)\right)$$

Where ω_{t-1} is the unique predecessor of ω_t .

At date 0 this gives $W_\theta(\omega_0) = -\sum_i p_i(\omega_0)\theta(\omega_0)$, which is simply the opposite of the initial cost of the portfolio.

The value of the portfolio is :

$$V_\theta(\omega_t) = \sum_i p_i(\omega_t)\theta_i(\omega_t)$$

We have obviously :

$$W_\theta(\omega_t) + V_\theta(\omega_t) = \sum_i \theta_i(\omega_{t-1})\left[d_i(\omega_t) + p_i(\omega_t)\right]$$

The no arbitrage condition imposes some restriction on the process of prices p. Indeed we have :

DEFINITION 30. An arbitrage portfolio is a portfolio such that $W_\theta(\omega_t) \geq 0$ for all t and ω_t with at least one strict inequality. An arbitrage portfolio makes money without taking any risk.

We have the following proposition:

PROPOSITION 31. *The market is arbitrage free if there exists a set of "pseudo state transition prices"* $q(\omega_t/\omega_{t-1}) \geq 0$ *such that for all asset* i *:*

$$p_i(\omega_t) = \sum_{\omega_{t+1} > \omega_t} q(\omega_{t+1}/\omega_t)\left[d_i(\omega_{t+1}) + p_i(\omega_{t+1})\right]$$

or for any portfolio :

$$V_\theta(\omega_t) = \sum_{\omega_{t+1} > \omega_t} q(\omega_{t+1}/\omega_t)\left[W_\theta(\omega_{t+1}) + V_\theta(\omega_{t+1})\right]$$

This proposition is quite simple to proove. It is sufficient to consider one state ω_t and its successors at a given date t . Then consider portfolios such that only $\theta_i(\omega_t)$ is not zero. . Apply then the result obtained in the static model.

The above equation tells us that the price of an asset is equal to a weighted sum of its liquidation value at the next date. The positive weights are common to all asets.

3.4. Risk neutral probability

Assume now that there is are risk free assets. At each date, a risk free asset gives at next date a constant cash flow $d(\omega_{t+1}) = 1$, and is replaced by a new risk free asset. In other words, there are T short term risk free assets of maturity one. The price at date t in the state ω_t of this asset (Bond) is hence :

$$B(t, \omega_t, 1) = \sum_{\omega_{t+1} > \omega_t} q(\omega_{t+1}/\omega_t) = \frac{1}{1 + r(\omega_t)}$$

Where $r(\omega_t)$ is simply the short term interest rate at date t in the state ω_t .

Just then define :

$$\widehat{q}(\omega_{t+1}/\omega_t) = q(\omega_{t+1}/\omega_t)\left(1 + r(\omega_t)\right)$$

This is a transition probability! Since it is positive and the sum on the successors of ω_t is one. This gives hence :

$$p_i(\omega_t) = \frac{1}{1 + r(\omega_t)} \sum_{\omega_{t+1} > \omega_t} \widehat{q}(\omega_{t+1}/\omega_t)\left[d_i(\omega_{t+1}) + p_i(\omega_{t+1})\right] = \frac{1}{1 + r(\omega_t)} \mathbb{E}_{\widehat{q}}\left[d_i(\omega_{t+1}) + p_i(\omega_{t+1})/\omega_t\right]$$

The price is simply the present value of the expected liquidation cash flow (computed with the risk neutral probability).

PROPOSITION 32. *when there exists a risk free asset, the price of any asset is the present value of its expected liquidation cash-flow, computed with some risk neutral probability. If the market is complete this probability is unique. If it is incomplete it is not, but gives the same value to all possible portfolios.*

DEFINITION 33. Consider $Q'(\omega_t) = \frac{1}{1+r(\omega_t)} \times \frac{1}{1+r(\omega_{t-1})} \cdots \frac{1}{1+r(\omega_0)}$. Where $\omega_t > \omega_{t-1} \cdots \omega_0$ is the unique path from ω_t to e_0 . It is called the discount factor at state ω_t. Let $\hat{p}_i(\omega_t) = Q'(\omega_{t-1})p_i(\omega_t)$ and $\hat{d}_i(\omega_t) = Q'(\omega_{t-1})d_i(\omega_t)$ the discounted price and divided of the asset i.

It is easy to compute the price of any asset at any state, function only of future dividends. Indeed we have :

$$\hat{p}_i(\omega_t) = \sum_{\omega_{t+1} > \omega_t} \widehat{q}(\omega_{t+1}/\omega_t) \left[\widehat{d}_i(\omega_{t+1}) + \hat{p}_i(\omega_{t+1})\right]$$

Define then as previously the probability of state ω_t knowing e_{t_0} by : $\widehat{q}(\omega_t/e_{t_0}) = \widehat{q}(\omega_t/\omega_{t-1})\widehat{q}(\omega_{t-1}/\omega_{t-2}) \cdots \widehat{q}(\omega_{t_0+1}/\omega_{t_0})$, with $\omega_t > \omega_{t-1} > \omega_{t-2} \cdots > \omega_{t_0+1} > \omega_{t_0}$ the unique path from ω_t back to ω_{t_0} . We have the following results :

PROPOSITION 34. *The (discounted) price at date t_0 of an asset in the state ω_{t_0} is equal to the expected discounted value of future dividends :*

$$\widehat{p}_i(\omega_{t_0}) = \sum_t \sum_{\omega_t} \widehat{q}(\omega_t/\omega_{t_0})\widehat{d}_i(\omega_t) \tag{3.4.1}$$

If an asset does not distibute dividends (except at final date T) then its discounted price is a martingale under the risk neutral probability:

$$\hat{p}(\omega_t) = \sum_{\omega_{t+1} > \omega_t} \widehat{q}(\omega_{t+1}/\omega_t) \left[\hat{p}(\omega_{t+1})\right]$$

3.5. The Cox Ross Rubinstein model.

The idea is to repete the 2 dates model. A state of nature at date t is a t digits sequence of *up* or *down* : $\omega_t = (\varepsilon_1, \varepsilon_2 ..., \varepsilon_t)$ with $\varepsilon_k = up$ or *down* .

Assume that $S_{t+1}((\omega_t, up)) = uS_t(\omega_t)$ and $S_{t+1}((\omega_t, down)) = dS_t(\omega_t)$ where u and d are two positive numbers with $u > d$. The (logarithm of the) price of the stock is an additive random walk. At each date the (log of the) price is incremented or decremented by $\ln u$ or $\ln d$.

Assume furthermore that the short term interest rate is constant. Then the transition prices are very simple whatever the state ω_t we have :

$$q((\omega_t, up)/\omega_t) = q_u = \frac{1}{1+r}\frac{1+r-d}{u-d}$$

$$q\left(\left(\omega_t, down\right)/\omega_t\right) = q_d = \frac{1}{1+r}\frac{u-(1+r)}{u-d}$$

And the risk neutral transition probability does not depend on ω_t:

$$\widehat{q}_u = \frac{1+r-d}{u-d}$$

$$\widehat{q}_d = \frac{u-(1+r)}{u-d}$$

This give a transition probability on the tree. And the allow to compute any discounted expected cash flow on that tree at any date.

3.6. Appendix 1 : filtration and tree

Above we have first defined a tree between dates 0 and T, and then a stochastic process on the tree. We can generalize in the following way.

3.6.1. Trajectories of a process. First, we can begin, more generaly, by defining all the possible "trajectories" of the process X. Assume for instance, $T = 3$ and that there are exactly 4 different trajectories named ω_3^i:

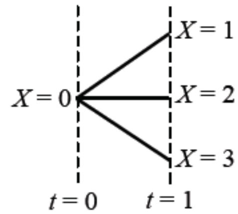
	$t=0$	$t=1$	$t=2$	$t=3$
ω^1	0	1	2	1
ω^2	0	2	3	3
ω^3	0	1	1	2
ω^4	0	3	2	2

(we have omitted the subscript 3).

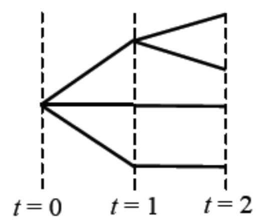
We call $\Omega = \left\{\omega^1, \omega^2, \omega^3, \omega^4\right\}$ the set of all possibilities. The process is simply a family of applications $\omega \to X_t(\omega)$.

Now we can build a tree which is associated to the progressive revelation of information on ω.

At $t = 0$, one observe $X = 0$, so that we have no information about the "realized" state in Ω. Now at $t = 1$ we acquire some information if $X_1 = 1$ we know that ω is either ω^1 or ω^3 , if $X_1 = 2$ then we are sure that $\omega = \omega^2$ and if $X_1 = 3$ then $\omega = \omega^4$. So we have a first partition of Ω :$\mathcal{P}_1 = \left\{\left\{\omega^1, \omega^3\right\}, \left\{\omega^2\right\}, \left\{\omega^4\right\}\right\}$. We can represent the 3 elements of $\mathcal{P}_1$ by drawing the first 3 branches of a tree :



Then at $t = 2$, one acquire information since the trajectories ω^1 and ω^3 are different! We have the partition $\mathcal{P}_2 = \left\{\left\{\omega^1\right\}, \left\{\omega^3\right\}, \left\{\omega^2\right\}, \left\{\omega^4\right\}\right\}$, so that the tree becomes :



We can define the σ-algebras associated to partitions. The σ-algebra associated to $\mathcal{P}_1$ is the set of subsets containing the three elements of the partition, their complementary and their unions :

$\mathscr{F}_1 = \left\{\left\{\omega^1, \omega^3\right\}, \left\{\omega^2\right\}, \left\{\omega^4\right\}, \left\{\omega^2, \omega^4\right\}, \left\{\omega^1, \omega^2, \omega^3\right\}, \left\{\omega^1, \omega^2, \omega^4\right\}, \left\{\omega^1, \omega^2, \omega^3, \omega^4\right\}, \emptyset\right\}$

it is the set of all parts of Ω where ω^1 and ω^3 are locked together (undistinguishable)!

At date 2 , $\mathscr{F}_2$ is simply the set of all parts of Ω because, observing X_2 allows to distinguish between ω_1 and ω_3. At $t = 0$ we have $\mathscr{F}_0 = \left\{\left\{\omega^1, \omega^2, \omega^3, \omega^4\right\}, \emptyset\right\}$

So the revelation of information gives rise to an increasing sequence : $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \mathscr{F}_2 = \mathscr{F}_3 = F_4$. We say we have a "filtration" on Ω, that is an increasing sequence of σ-algebras.

3.6.2. Generalization. More generally we can define first a probabilizable space Ω endowed with a σ-algebra $\mathscr{F}$. Then, a process is simply a family $X_t : \omega \to X_t(\omega) \in \mathbb{R}$, $t \in \{0, 1, 2, ..., T\}$. obviously X_t must be $\mathscr{F}$ measurable, that is that , for any date and for any borellian B of $\mathbb{R}$, $X_t^{-1}(B) \in \mathscr{F}$.

A filtration is an increasing sequence of σ-algebras : $\mathscr{F}_t \subset \mathscr{F}_{t+1} \subset \mathscr{F}$. We say that X is adapted to the filtration if, for any date and any borellian : $X_t^{-1}(B) \in \mathscr{F}_t$

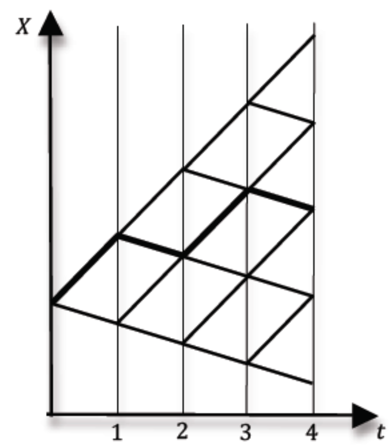
A very "natural filtration" is the one generated by X itself. $\mathscr{F}_t^X$ is exactly composed of all the $X_t^{-1}(B)$. (This is what we have done in the previous paragraph).

3.7. Appendix 2 : the example of a random walk

In a discrete random walk between 0 and T, the set Ω is the set of sequences $\omega = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_T)$ where the "digits" ε_t can take two values (for instance, a and b, $a > b$). There are 2^T such sequences. The random walk itself is a process X such that $X_{t+1}(\omega) = X_t(\omega) + \varepsilon_{t+1}(\omega)$ where $\varepsilon_{t+1}(\omega)$ is the $(t+1)$th digit of ω, X_0 being given.

One can represent such a random walk on a grid (see figure). A trajectory is a path from $t = 0$ to $t = T$ on the grid. To each ω corresponds a unique trajectory.

There are exactly 2^T trajectories. But at each date, there are only $t+1$ different values for X. Knowing the value of X only at date t does not give the whole information on the past trajectory. Several trajectories can lead to the same value of X.



The last ingredient of the random walk is the probability structure attached to Ω.

Remember that one way is to define transition probabilities on the tree associated to Ω: $\pi(\omega_{t+1}/\omega_t) = \pi((\omega_t, \varepsilon_{t+1})/\omega_t)$. For each ω_t this gives two numbers (with sum equal to one) $\pi((\omega_t, a)/\omega_t)$ and $\pi((\omega_t, b)/\omega_t)$.

Definition 35. We say that the process X has the Markov property if that transition probability $\pi((\omega_t, a)/\omega_t)$ (and hence $\pi((\omega_t, b)/\omega_t) = 1 - \pi((\omega_t, a)/\omega_t)$) is the same for all the ω_t leading to the same value of X_t that is for all the past trajectories leading to the same value of X_t. The future evolution of the process depends on the past uniquely through the present value.

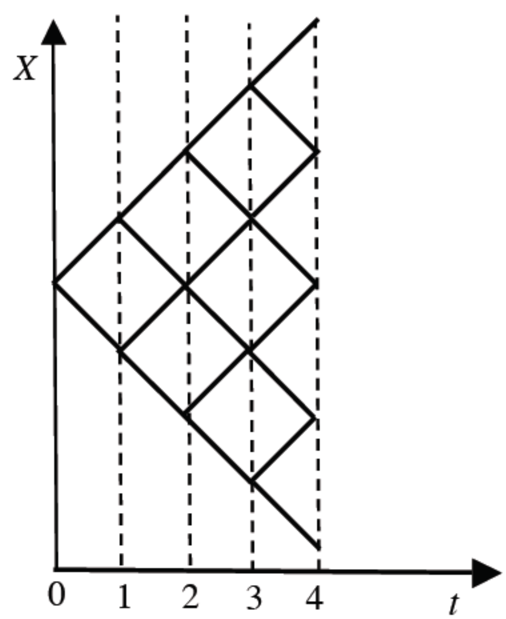
On the grid, the Markov property means that on can give "weights" to the two different vertices starting from each node of the grid (that is for each value of X_t) : $\pi_a(X_t)$ and $\pi_b(X_t)$.

When these transition probabilities are independent of X_t , say π_a and π_b, one says we have a stationnary random walk.

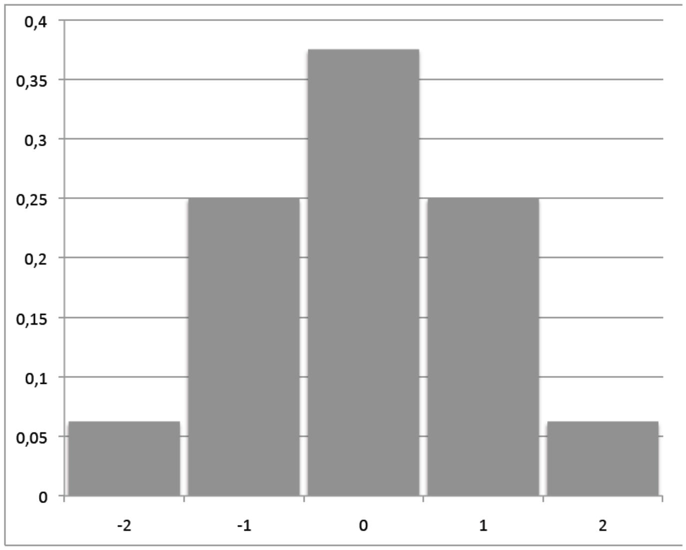
In that case the transition probabilities for the process itself is very simple. The probability that the process reaches y at date $t' \geq t$ if it is equal to x at date t depends only on $y - x$ and $t' - t$. Moreover $y - x$ can take only $t' - t + 1$ values : $ka + (t' - t - k)\, b$, $k = 0, 1, ..., t' - t$.

$$\Pr\left(X_{t'} = x + ka + (t' - t - k)\, b / X_t = x\right) = \begin{pmatrix} n \\ k \end{pmatrix} \pi_u^k \pi_d^{t'-t-k}$$

For instance, for $t' - t = 4$, $a = -b$ and $\pi_a = \pi_b = \frac{1}{2}$, the 5 probabilities are $\frac{1}{16}, \frac{4}{16}, \frac{6}{16}, \frac{4}{16}, \frac{1}{16}$



This give an histogram centrerd on X_0 =0, for $a = -b = 1$:



Now we can make another remark.

Consider the probability transition $\Pr(X_{t'} = y/X_t = x)$. And an intermediate date t_1 $(t < t_1 < t')$. We have :

$$\Pr(X_{t'} = y/X_t = x) = \sum_z \Pr(X_{t_1} = z/X_t = x)\Pr(X_{t'} = y/X_{t_1} = z)$$

That we can write :

$$\Pr(X_{t'} = y/X_t = x) = \mathbb{E}\left[\Pr(X_{t'} = y/X_{t_1})/X_t = x\right]$$

Define now, t' and y being fixed the process, $Y_t = \Pr(X_{t'} = y/X_t)$

The previous equation shows that Y is a martingale since this equation implies $Y_t = \mathbb{E}\left[Y_{t_1}/Y_t\right]$

3.8. Asymptotic random walk

The idea is to take a fixed interval $[0, T]$, subdivision $t_0 = 0, t_1, ...t_k, ..., t_n = T$ of $[0, T]$ with $t_k = k\frac{T}{n}$, and make n tend to infinity.

For n given, a stationnary random walk is defined as previously :

$$W^n_{t_{k+1}} = W^n_{t_k} + \varepsilon^n$$

with $\varepsilon^n = a_n$ or b_n

Take a symmetric random walk $a_n = -b_n = s_n$, and $\pi_a = \pi_b = \frac{1}{2}$

As all the increment are independent, the variance of X_T is equal to the sum of the variances of each step :

$$\text{var}W_T^n = \sum_{k=1}^{n} \text{var}\varepsilon_k = ns_n^2$$

If we want that varW_T be finite it is necessary that $s_n \sim \frac{1}{\sqrt{n}}$.

Taking $s_n = \sqrt{\frac{T}{n}}$ allows to something interesting : the variance of the process at T is equal to the time elapsed.

REMARK 36. In order to get a random walk with variance equal to T, each step must be such that $s_n = \sqrt{\frac{T}{n}}$, that we can write with an enormous abuse of notation $(dW)^2 = dt$

When we make n tend to infinity, a trajectory becomes a "fractal" object which is continuous but nowhere differentiable. The length of a trajectory is $n\sqrt{\left(\frac{T}{n}\right)^2 + \frac{T}{n}}$ that tends to infinity.

One can show that the "limit" process is exactly a standard Brownian motion. One can see this by applying the Central Limit Theorem and show that W_T is gaussian.

CHAPTER 4

Continuous models

4.1. The "continuous random walk" : Brownian motion

We want to replace the random walk, of chapter 3 by some time-continuous process. The idea is very simple. between to "consecutive" dates t and $t + dt$ instead of having a finite set of possibilities, we will have an "infinite" non countable set of "states". We would like to define a random continuous variable "drawn" at each date in each previous state. There is a way to do that : brownian motions.

4.1.1. Brownian motion. We want to define the underlying probabilizable space $(\Omega, \mathscr{F})$. As in the discrete case, we would like to define ω in Ω as a "trajectory" between 0 and T of some process (the Brownian Motion) W. $W(t, \omega)$ is the value of W at date$t$ if the trajectory is ω. We will note $W(t)$ the random variable $\omega \longrightarrow W(t, \omega)$. The process is defined as follows

Definition 37. A standard Brownian Motion (SBM) is a process $W(t, \omega)$ such that :

1. $W(0, \omega) = 0$ with probability 1

2. $t \rightarrow W(t, \omega)$ is continuous for all ω

3. for all subdivision $t_0 = 0$, $t_1, ..., t_k, t_n = T$, $W(t_{k+1}) - W(t_k)$ and $W(t_{h+1}) - W(t_h)$ are independent

4. $W(t_{k+1}) - W(t_k)$ is a normal variable with zero mean and variance equal to $t_{k+1} - t_k$.

This definition is consistent because, and it is very important, normal distribution is stable : if two variables are normal, then any linear combination is normal. So the definition above does not depend on the way you subdivise the interval.

A brownian motion is a generalization of a random walk : at each "step" W "moves" randomly (increases or decreases) according to a Gaussian law.

Let us examine the main properties of a SBM.

4.1.1.1. *Quadratic variations.* Let $W(t)$ a SBM : $W(t')-W(t)$ is a normal variable centered with variance $t'-t$.

Let a subdivision t_i.

We want to study :

$$\Delta_k = (W(t_{k+1}) - W(t_k))^2 - (t_{k+1} - t_k)$$

We have :

$$\sum \Delta_k = \sum (W(t_{k+1}) - W(t_k))^2 - T$$

Let us examine the moments of Δ_k.

First, the mean is zero since $W(t_{k+1}) - W(t_k)$ is precisely a normal variable with variance $(t_{k+1} - t_k)$,

$$\mathrm{E}\left[\Delta_k\right] = \mathrm{E}\left[(W(t_{k+1}) - W(t_k))^{\,2}\right] - (t_{k+1} - t_k)$$

$$= \mathrm{var}\left[W(t_{k+1}) - W(t_k)\right] - (t_{k+1} - t_k) = 0$$

What is its variance (remember that the fourth moment of a normal is $3\sigma^4$) :

$$\begin{aligned}
\mathrm{var}(\Delta_k) &= E((W(t_{k+1}) - W(t_k))^4 - 2(t_{k+1} - t_k)(W(t_{k+1}) - W(t_k))^2 + (t_{k+1} - t_k)^2) \\
&= 3\left(\mathrm{var}\left[W(t_{k+1}) - W(t_k)\right]\right)^2 - 2(t_{k+1} - t_k)\mathrm{var}\left[W(t_{k+1}) - W(t_k)\right] + (t_{k+1} - t_k)^2 \\
&= 2(t_{k+1} - t_k)^2
\end{aligned}$$

Take for example $t_{k+1} - t_k = 1/n$,

Then we have :

$$\lim \mathrm{var}(\sum \Delta_k) = 0$$

The variance of the quadratic variation is zero! That means that $\sum (W(t_{k+1}) - W(t_k))^2$ is not very far from being “very often” T.

In fact one can show that : $\sum (W(t_{k+1}) - W(t_k))^2$ is asymptotically almost surely equal to T.

That allows to note :

$$\left[dW(t)\right]^2 = dt$$

This is obviuously an abuse of notation, but very useful...

4.1.2. Filtration. As in the discrete case one can define the filtration$\mathscr{F}_t$ on Ω induced by a SBM. But this is a very complicated object. The idea is the following.

For each$t$, $W(t)$ is $\mathscr{F}_t$-measurable, for each t and for future subdivision $t < t_1 < t_2 < ... < t_n$, the Brownian motion increments $W(t_1) - W(t), W(t_2) - W(t_1), ..., W(t_n) - W(t_{n-1})$ are independent of $\mathscr{F}_t$.

Here is one way to construct $\mathscr{F}_t$. First fix t. Let $s \in [0, t]$ and $C \in \mathcal{B}(\mathbb{R})$ a borellian be given. Put the set $\{W(s) \in C\} = \{\omega, W(s, \omega) \in C\}$ in $\mathscr{F}_t$. Do this for all possible numbers $s \in [0; t]$ and $C \in \mathcal{B}(\mathbb{R})$. Then put in every other set required by the -algebra properties. This $\mathscr{F}_t$ contains exactly the information learned by observing the Brownian motion up to time t.

In the sequel, the underlying probability structure is the one of a given SBM with the associated filtration.

4.1.3. Stochastic integral.

4.1.3.1. *The simple case of a real function.* Let $f(t)$ a real function of t. Just define :

$$\sum_{0}^{n-1} f(t_k) \left(W(t_{k+1}) - W(t_k)\right)$$

This is a normal variable whose mean is zero and whose variance is :

$$\sum_{0}^{n-1} \left(f(t_k)\right)^2 (t_{k+1} - t_k)$$

What is the limit of the above sum? Obviously $\int_0^T \left(f(t)\right)^2 dt$

That alllows us to define the stochastic integral of f :

$$\int_0^T f(t)dW(t)$$

This is a normal variable whose mean is zero and variance $\int_0^T \left(f(t)\right)^2 dt$.

4.1.3.2. *Stochastic integral of a process f.* Ther intuition is the following : $f(t)$ and $W(t)$ must be measurable on the same filtration, or put differently f is adapted to the filtration of W.

In that case one can define the stochastic integral

$$\int_0^T f(t)dW(t)$$

- Stochastic differential

Instead of $X(T) = \int_0^T f(t)dW(t)$ one may writes :

$$dX(t) = f(t)dW(t)$$

This extends the notion of differential, but the right writing would be under the integral form.

4.1.4. Stochastic differential equation. We can define process as solutions of a :

DEFINITION 38. Stochastic Differential Equation :

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

This means that the random process X must be such that :

$$X(t) = X(0) + \int_0^t \mu(X(s), s)ds + \int_0^t \sigma(X(s), s)dB(s)$$

The second integral being a "stochastic integral".

The solution of such a "stochastic differentiel equation" is called a diffusion process.

A good example is the geometric brownian motion :

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

4.1.5. Martingale. In this context what is a martingale? Remember that in the discrete finite model, a random process is a martingale if the value at time t is equal to the mean of the values at the successors.

Here the definition is :

DEFINITION 39. X (adapted to the filtration) is a martingale if $\forall t, \forall t', t \leq t', X(t) = \mathbb{E}\left[X(t') / \mathscr{F}_t\right]$

Knowing W up to t, the value of X at t is the expected value of X at any posterior date.

Here it can be shown that :

PROPOSITION 40. *$X(t)$ is a martingale if there exists a process $\gamma(t)$ adapted such that $dX(t) = \gamma(t)dW(t)$.*

4.1.6. Ito lemma. Let f a function of a real variable. We would like to study the process $f(W(t))$ where B is a SBM. Or more generally $f(X(t))$ where $X(t)$ is a diffusion process.

Just write :

$$f(W(t_{k+1})) - f(W(t_k)) = f'(W(t_k))\left(W(t_{k+1}) - W(t_k)\right) + \frac{1}{2}f''(W(t_k))\left(W(t_{k+1}) - W(t_k)\right)^2 + R\left(\left(W(t_{k+1}) - W(t_k)\right)\right)$$

take the sum :

$$f(W(t)) - f(W(0)) = \sum f(W(t_{k+1})) - f(W(t_k))$$

The first and the second term give (at the limit) :

$$\int_0^T f(W(t))dW(t) + \frac{1}{2}\int_0^T f''(W(t))dt$$

one can show that the last term is almost surely 0 at the limit.

We can hence write (simple Ito lemma):

$$df(W(t)) = f'((W(t))dW(t) + \frac{1}{2}f''(W(t))dt$$

This formula is very easy to understand. Think of a particle moving on the real line according to the Brownian motion. Take for example $f(x) = x^2$, $f(W(t))$ is simply the square of the "distance" traveled by the particle. Obviously this distance is not zero on average. At each step the particle moves away by a distance equally distributed around a positive value equal to the time spent : $dX^2 = dt + 2W(t)dW(t)$.

More generally if we have a diffusion process :

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

And a real function $f(t, x)$.If we set $Y(t) = f(t, X(t))$. We can write :

$$dY(t) = \frac{\partial f}{\partial t}(t, X(t))dt + \frac{\partial f}{\partial x}(t, X(t))\left\{\mu(t, X(t))dt + \sigma(t, X(t))dW(t)\right\} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, X(t))\left(\sigma(t, X(t))\right)^2 dt$$

Or :

LEMMA 41. *(Ito Lemma) Let $X(t)$ a diffusion process such that $dX(t) = \mu(t)dt + \sigma(t)dW(t)$. Let $Y(t) = f(X(t), t)$, where f is a C2 real function. Then Y is a diffusion process and :*

$$dY(t) = \left\{\frac{\partial f}{\partial t}(t, X(t)) + \frac{\partial f}{\partial x}(t, X(t))\mu(t, X(t)) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, X(t))\left(\sigma(t, X(t))\right)^2\right\} dt$$

$$+\frac{\partial f}{\partial x}(t, X(t))\sigma(t, X(t))dW(t)$$

We define the Dynkin operator by

$$Df(t, x)) = \frac{\partial f}{\partial t}(t, x) + \frac{\partial f}{\partial x}(t, x)\mu(t, x) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, x)\left(\sigma(t, x)\right)^2$$

This is simply the mean of the variation of the process $Y(t) = f(t, X(t))$.

4.1.7. Brownian motion with drift and scaling. A Brownian motion with drift and scaling is the simple diffusion X solution of :

$$dX(t) = \mu dt + \sigma dW(t)$$

Where μ is a constant “drift” and σ a constant scaling (variance) factor.

Obviously :

$$X(t) = X(0) + \mu t + \sigma W(t)$$

We can also write :

$$dX(t) = \sigma \left(\frac{\mu}{\sigma} dt + dW(t) \right) = \sigma d\widetilde{W}(t)$$

$\widetilde{W} = \frac{\mu}{\sigma} t + W(t)$ is a drifted modification of W with a drift equal to $\theta = \frac{\mu}{\sigma}$

4.1.8. Geometric BM. Just look at a geometric brownian motion, diffusion solution of :

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

Take $Y(t) = \ln(X(t))$. By Ito formula we get :

$$dY(t) = \left(\mu - \frac{1}{2}\sigma^2 \right) dt + \sigma dW(t)$$

That is :

$$X(t) = X(0) \exp \left[\left(\mu - \frac{1}{2}\sigma^2 \right) t + \sigma W(t) \right]$$

4.1.9. Transition probabilities. It is interesting to write the transition probabilities. For a SBM the probability density of $W(t')$, knowing that $W(t) = x$ is :

$$f(t, t', x, y) = \frac{1}{\sqrt{2\pi (t' - t)}} \exp \left(-\frac{(y - x)^2}{2 (t' - t)} \right)$$

This is the density of probability to go from x at t to y at t' . In some sense the probability of all the trajectories from x to y. Take then an intermediate date t_1 . By definition of the SBM we have:

$$f(t, t', x, y) = \int_{-\infty}^{+\infty} f(t, t_1, x, z) f(t_1, t', z, y) dz$$

Now fix $t'$and y and define

$$\phi(t) = f(t, t', W(t), y)$$

Rewrite the previous equation :

$$\phi(t) = \int_{-\infty}^{+\infty} f(t, t_1, W(t), z) f(t_1, t', z, y) dz$$

That is :

$$\phi(t) = \mathbb{E}(\phi(t_1) / W(t)) = \mathbb{E}(\phi(t_1) / \phi(t))$$

$P(t)$ is a martingale.

By Ito lemma

$$d\phi(t) = \left[\frac{\partial f}{\partial t}(t, t', W(t), y) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, t', W(t), y)\right] dt + \frac{\partial f}{\partial x}(t, t', W(t), y)\, dW(t)$$

We check that the Dynkin term is 0 :

$$\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} = 0$$

Indeed :

$$\frac{\partial f}{\partial t} = \frac{1}{\sqrt{2\pi}\,(t'-t)^{3/2}}\left[\frac{1-(y-x)^2}{2(t'-t)}\right]\exp\left(-\frac{(y-x)^2}{2(t'-t)}\right)$$

and

$$\frac{\partial f}{\partial x} = \frac{y-x}{\sqrt{2\pi}\,(t'-t)^{3/2}}\exp\left(-\frac{(y-x)^2}{2(t'-t)}\right)$$

and

$$\frac{\partial^2 f}{\partial x^2} = \left[\frac{-1}{\sqrt{2\pi}\,(t'-t)^{3/2}} + \frac{(y-x)^2}{\sqrt{2\pi}\,(t'-t)^{5/2}}\right]\exp\left(-\frac{(y-x)^2}{2(t'-t)}\right)$$

This equation is a kind of heat equation :

Remark 42. If $f(t, t', x, y)$ is the transition probability of the SBM, we have

$$\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} = 0$$

4.1.10. Probability change : Girsanov Theorem (light). Take $\widetilde{W}=\theta t+W$ a standard Brownian with drift. The “trajectories” of $\widetilde{W}$ are qualitatively the same as those of W. But their probabilities are not the same. To be more precise let us compute the transition probability of $\widetilde{W}$:

$$f_\theta\left(0,t,0,y\right)=\frac{1}{\sqrt{2\pi\left(t'-t\right)}}\exp\left(-\frac{\left(y-\theta t\right)^2}{2\left(t'-t\right)}\right)$$

That can be rearranged in the following way :

$$f_\theta\left(0,t,0,y\right)=\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{y^2}{2t}\right)\exp\left(\frac{2\theta yt-\theta^2t^2}{2t}\right)$$

$$f_\theta\left(0,t,0,y\right)=\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{y^2}{2t}\right)\exp\left(-\frac{1}{2}\theta^2t+\theta y\right)$$

So :

$$f\left(0,t,0,y\right)=f_\theta\left(0,t,0,y\right)\exp\left(\frac{1}{2}\theta^2t-\theta y\right)$$

Roughly speaking, if you take a trajectory between 0 and t of $\widetilde{W}$, if you multiply its weight by $\exp\left(\frac{1}{2}\theta^2t-\theta\widetilde{W}\left(t\right)\right)=\exp\left(-\frac{\theta^2}{2}t-\theta W\left(t\right)\right)$, the new weight of this trajectory is the one of the SBM.

More precisely, let define Z by :

$$Z\left(t\right)=\exp\left(-\frac{1}{2}\theta^2t-\theta W\left(t\right)\right)$$

One can show that Z is a martingale, indeed (Ito):

$$dZ\left(t\right)=Z\left(t\right)\left(-\frac{1}{2}\theta^2dt-\theta dB\left(t\right)\right)+\frac{1}{2}Z\left(t\right)\theta^2dt$$

That is :

$$dZ\left(t\right)=-\theta Z\left(t\right)dB\left(t\right)$$

That implies in particular :

$$\mathbb{E}\left(Z\left(t\right)\right)=Z\left(0\right)=1$$

Take now an horizon T.

Let now define a new probability by :

$$\widetilde{\Pr}(A) = \int_A Z(T)\, d\Pr(\omega) = \mathbb{E}_{\Pr}\left[\mathbb{I}_A Z(T)\right]$$

If $A \in \mathscr{F}_t$ for $t \leq T$, we have, by definition of conditional expectation and because Z is a martingale :

$$\widetilde{\Pr}(A) = \mathbb{E}_{\Pr}\left[\mathbb{I}_A Z(T)\right] = \mathbb{E}_{\Pr}\left[\mathbb{I}_A \mathbb{E}_{\Pr}\left[Z(T)/\mathscr{F}_t\right]\right] = \mathbb{E}_{\Pr}\left[Z(t)\,\mathbb{I}_A\right]$$

Under this new probability measure, $\widetilde{W}$ is a SBM. Indeed, let us compute the new probability density of $\widetilde{W}(t)$, for instance take $A = \left\{\widetilde{W}(t) \in [a,b]\right\} = \{W(t) \in [a-\theta t, b-\theta t]\}$:

$$\widetilde{\Pr}(A) = \int_{a-\theta t}^{b-\theta t} \exp\left(-\frac{1}{2}\theta^2 t - \theta y\right) \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{y^2}{2t}\right) dy = \int_{a-\theta t}^{b-\theta t} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y+\theta t)^2}{2t}\right) dy = \int_a^b \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{y^2}{2t}\right) dy$$

4.1.11. First time passage. Let a SBM W and define the “maximum process” as $M(t) = \max(W(s), s \leq t)$

Take $m > 0$ and let

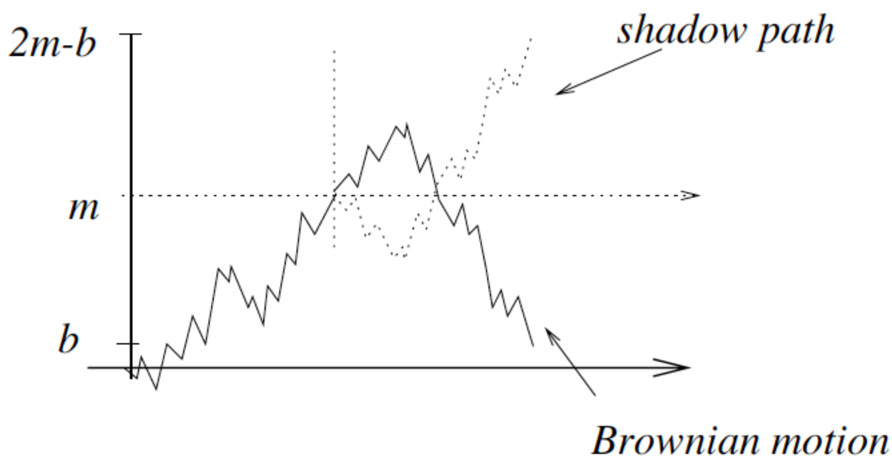
$$\tau = \inf(t, W(t) > m)$$

We have obviously :

$$\Pr[\tau < t] = \Pr[M(t) > m]$$

Take a trajectory of W such that $M(t) > m$ and $W(t) < b$ with $b < m$. At the first time τ (necessarily $< t$) where $W(\tau) = m$, take the symmetrical (shadow) trajectory.

This new trajectory has the same “probability” as the original one. This new trajectory is such that $W(t) > 2m - b$. Reciprocally to any trajectory such that $W(t) > 2m - b$, corresponds a unique trajectory such that $M(t) > m$ and $W(t) < b$.



So :

$$\Pr\left[W\left(t\right) < b, M\left(t\right) > m\right] = \Pr\left[W\left(t\right) > 2m - b\right]$$

We have :

$$\Pr\left[M(t) > m\right] = \int_{-\infty}^{m} \frac{\partial}{\partial b} \int_{2m-b}^{+\infty} \frac{1}{\sqrt{2\pi t}} \exp\left[-\frac{x^2}{2t}\right] dx$$

We have also :

$$\Pr\left[M(t) > m\right] = \Pr\left[W(t) < m, M(t) > m\right] + \Pr\left[W(t) > m\right] = 2\Pr\left[W(t) > m\right]$$

$$\Pr\left[M(t) > m\right] = \Pr\left[\tau < t\right] = 2\int_{m}^{+\infty} \frac{1}{\sqrt{2\pi t}} \exp\left[-\frac{x^2}{2t}\right] dx$$

Change variable $z = \frac{x}{\sqrt{t}}$:

$$\Pr\left[M(t) > m\right] = \Pr\left[\tau < t\right] = 2\int_{\frac{m}{\sqrt{t}}}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] dz$$

the density $f_\tau\left(t\right) = \frac{\partial}{\partial t}\Pr\left[\tau < t\right]$ is then :

$$f_\tau\left(t\right) = \frac{m}{\sqrt{2\pi}t^{3/2}} \exp\left[-\frac{m^2}{2t}\right]$$

Which is a Levy law.

4.2. Arbitrage free condition

As in the discrete case, one can show that the no arbitrage condition amounts to the existence of a risk neutral probability structure. We must find a probability structure on "the tree" such that the price of an asset is in this case equal to the discounted expected value of the future cash flows. But first consider the deterministic case.

4.2.1. Preamble Arbitrage free hypthesis in a deterministic continuous model : zero coupons bonds. Recall that a zero coupon bond has only a terminal payment at the termination date T. Its price at date t (prior to T) is noted $B(t,T)$ (with $B(T,T)) = 1$).

How can we express that, given all theses prices at all dates, the market is arbirage free?

Intuitively the no arbitrage condition means that the yields between t and $t{+}h$, per unit of time , $\frac{B(t+h,T)-B(t,T)}{B(t,T)h}$ does not depend on T. Indeed if it were not the case ons could make money by selling and buying two zero coupons with different maturities. The limit when h goes to zero is called the short term interest rate $r(t)$.

PROPOSITION 43. *If the market is arbitrage free , then there exists a deterministic short term interest rate $r(t)$ such that :*

$$\forall t \leq T \qquad \frac{\partial B}{\partial t}(t,T) = r(t)B(t,T)$$

That is , using $B(T,T) = 1$:

$$B(t,T) = \exp\left[-\int_t^T r(s)ds\right]$$

If one knows the dynamics of short term rate $r(t)$, then one knows all the zero coupon prices.

4.2.1.1. *The yield curve.* On can rewrite the price of the z-c bond introducing its yield as a function of maturity :

$$B(t,T) = \exp\left[-\int_t^T r(s)ds\right] = \exp\left[-R(t,T)(T-t)\right]$$

That is :

$$R(t,T) = \frac{\int_t^T r(s)ds}{T-t}$$

Writing it as a function of maturity τ:

$$\widehat{R}(t,\tau)=\frac{\int_t^{t+\tau}r(s)ds}{\tau}$$

Definition 44. $\tau\to\widehat{R}(t,\tau)$ is the "yield curve" at date $t:$ the yield per unit of time as a function of maturity.

4.2.1.2. *Pricing of more complicated assets.* Assume that all the ZC exists and consider one asset A that gives a continuous cash flow $d(t)$ per unit of time.

Obviously, the price of such an asset must be :

$$y(t_0)=y(t)B\left(t_0,t\right)+\int_{t_0}^{t}d(s)B\left(t_0,s\right)ds$$

It is quite easy to find this result another way:

With one euro at date t one has two possible strategies : buy the asset, earn the cash flow and resell it at date $t+dt$. or buy a zero coupon and get $1+r(t)dt$ per euro. Let $y(t)$ the value of asset A at date t. Arbitrage-free hypothesis implies :

$$\overset{\bullet}{y}+d=ry$$

It is quite easy to solve this differential equation :

$$y(t)=\exp\left(\int_{t_0}^{t}r(s)ds\right)\left[y(t_0)-\int_{t_0}^{t}d(s)\exp\left(-\int_{t_0}^{s}r(u)du\right)ds\right]$$

4.2.1.3. *Examples.* The above formula has two components : the first one takes into account the resell value at date t. The second is the present value of the cash flow.

Consider then the following particular cases :

(1) Finite bond. Imagine there is a time T such that $d(t)=0$ for $t\geq T$. This implies $y(T)=0$, so that:
$$y(t_0)=\int_{t_0}^{T}d(s)\exp\left(-\int_{t_0}^{s}r(u)du\right)ds$$
that is the discounted value of cash flow: the fundamental
(2) Infinite living asset with no dividend: pure bubble :
$$y(t)=y(0)\exp\left(\int_{0}^{t}r(s)ds\right)$$
here the value grows at a rate equal to r

(3) General case : fondamental and bubble part. Let $f(t) = \int_t^{\infty} d(s) \exp\left(-\int_t^s r(u)du\right) ds$, the "fondamental value" of the asset, we have : $\dot{f} = d + rf$, so that

$$y - f = r(y - f)$$

$y - f$ is a pure bubble

These examples shows that arbitrage free condition does not prevent bubbles (permanent bubbles) growing at the rate r.

4.2.1.4. *(kind of) derivatives.* Imagine now there is a "state variable" governed by the following equation :

$$\dot{X}(t) = f(t, X(t))$$

Let $s \to \widehat{X}(t, x, s)$ the trajectory, that is the value of X at date s when the value at t is x. To be simple we write, when there is no ambiguity :

$$x(s) = \widehat{X}(t, x, s)$$

We have an asset (kind of) "derivative" whose cash flow $d(t, x)$ depends on the value of the state variable, and that terminates at T giving $h(x)$ en T.

One assumes that one can define the instantaneous rate of return $r(t, x)$: all the zero coupons exist.

Let $V(t, x)$ the value of the asset , $v(s) = V(s, X(t, x, s))$

Hence :

$$\dot{v}(s) = \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} f(s, x(s))$$

Or :

$$\dot{v}(s) + d(s, x(s)) = r(s, x(s))v(s)$$

So :

$$V(t, x) = v(t) = h(x(T)) \exp\left(-\int_t^T r(s, x(s))ds\right)$$

$$+ \int_t^T d(s, x(s)) \exp\left(-\int_t^s r(u, x(u))du\right) ds$$

4.2.2. The stochastic case.

4.2.2.1. *Diffusion with linear drift.* First of all, it is useful to study linear drift diffusions, solutions of :

$$dY(t) = (a(t)Y(t) - b(t))\, dt + \sigma(t)dW(t)$$

Where a b and σ are adapted to the filtration of W.

The solution of

$$dY(t) = (a(t)Y(t) - b(t))\, dt + \sigma(t)dW(t)$$

is such that :

$$Y(t_0) = \mathbb{E}\left[Y(t)\exp\left[-\int_{t_0}^{t} a(u)du\right] + \int_{t_0}^{t}\left(b(s)\exp\left[-\int_{t_0}^{s} a(u)du\right]\right)ds/\mathscr{F}_{t_0}\right]$$

PROOF. take

$$X_0(t) = -\int_{t_0}^{t}\left(b(s)\exp\left[\int_{s}^{t} a(u)du\right]\right)ds$$

we have

$$dX_0 = (-b + aX_0)\, dt$$

Set

$$Y(t) = Z(t)\exp\left[\int_{t_0}^{t} a(s)ds\right] + X_0(t)$$

Then

$$(aY - b)\, dt + \sigma dW = dZ\exp\left[\int_{t_0}^{t} a(s)ds\right] + aZ\exp\left[\int_{t_0}^{t} a(s)ds\right]dt + (aX_0 - b)dt$$

Thatt is :

$$\sigma dW = dZ\exp\left[\int_{t_0}^{t} a(s)ds\right]$$

And hence :

$$Z(t) = Z(t_0) + \int_{t0}^{t}\sigma(s)\exp\left(-\int_{t_0}^{s} a(u)du\right)dW(s)$$

So :

$$Y = Y(t_0)\exp\left[\int_{t_0}^{t} a(u)du\right] - \int_{t_0}^{t}\left(b(s)\exp\left[\int_{s}^{t} a(u)du\right]\right)ds + \int_{t_0}^{t}\sigma(s)\exp\left(\int_{s}^{t} a(u)du\right)dW(s)$$

□

4.2.2.2. *Arbitrage free condition.* We are going to give here a very general result concerning arbitrage free condition. We assume that randomness is described by an underlying diffusion process that we take as a “state variable”.

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

- Assets

Assume there are $K+1$ assets labeled 0,1,2.., K.

The price p_i of asset i depends on t and on the value at that date taken by the state variable :

$$p_i(t) = P_i(t, X(t))$$

Each asset gives a cash flow $b_i(t, X(t))$

Asset number 0 is locally risk free asset :

$$\frac{dp_0}{dt} = r(t, X(t))p_0(t)$$

- Portfolios

Let $\Delta = (\Delta_1, \Delta_2, ..., \Delta_k)$ a portfolio : Δ_i is the quantity of asset i in the portfolio.

- Arbitrage free principle :if a portfolio is risk free, it must have the same return as the risk free asset :

Consider the value of a portfolio at date t when $X(t) = x$:

$$P_{\triangle}(t, x) = \sum \Delta_i P_i(t, x)$$

So that noting D the Dynkin operator, the value added is :

$$\left\{\sum \Delta_i DP_i(t, x) + b_i(t, x)\right\} dt + \left\{\sum \Delta_i \frac{\partial P_i}{\partial x}(t, x)\sigma(t, x)\right\} dW(t)$$

or put differently

$$dp_{\triangle} + < \triangle.b > = < \triangle.DP + b > dt + < \triangle.\frac{\partial P}{\partial x} > dW$$

Δ is risk free if the coefficient of dW is 0 :

$$< \triangle.\frac{\partial P}{\partial x} > = 0$$

In that case we have

$$\forall \Delta < \Delta.\frac{\partial P}{\partial x} > = 0 \Rightarrow < \Delta.DP + b > = r < \Delta, P >$$

$$\forall\Delta,\ <\Delta.\frac{\partial P}{\partial x}>=0 \Rightarrow <\Delta.\left(DP+b-rP\right)>=0$$

the two hyperplanes of $\mathbb{R}^K$, $\left(\frac{\partial P}{\partial x}\right)^{\perp}$ and $(DP+b-rP)^{\perp}$ are identical so that there exists $\lambda(t,x)$ such that for all i >0:

$$DP_i(t,x)+b_i(t,x)-r(t,x)P_i(t,x)=\lambda(t,x)\frac{\partial P_i}{\partial x}(t,x)$$

The market is arbitrage free if there exists a function $\lambda(t,x)$ such that for any asset i we have : (D is the Dynkin) :

$$DP_i\left(t,x\right)+b_i\left(t,x\right)-r\left(t,x\right)P_i\left(t,x\right)=\lambda\left(t,x\right)\frac{\partial P_i}{\partial x}\left(t,x\right)$$

This equation will be very powerful in some particular cases.

4.3. the Black and Scholes model.

4.3.1. Risk neutral dynamics. The state variable is the price of one of the assets $X=S=P_1$ distributing no dividend $b_1=0$:

$$dS\left(t\right)=\mu\left(t,S\left(t\right)\right)dt+\sigma\left(t,S\left(t\right)\right)dW(t)$$

We have : $DP_1(t,s)=\mu(t,s)$ and $\frac{\partial P_1}{\partial S}=1$.

So that the arbitrage free condition gives :

$$\mu\left(t,S\right)-r(t,S)S=\lambda\left(t,S\right)$$

And then we have for every other asset :

$$DP_i(t,S)+b_i(t,S)-r(t,S)P_i(t,S)=\left(\mu(t,S)-r(t,S)S\right)\frac{\partial P_i}{\partial S}(t,S)$$

That is :

$$\frac{\partial P_i}{\partial t}(t,S)+\frac{\partial P_i}{\partial S}(t,S)r(t,S)S+\frac{1}{2}\frac{\partial^2 P_i}{\partial S^2}(t,S)\sigma(t,S)^2=r(t,S)P_i(t,S)-b_i(t,S)$$

That one can write :

$$\hat{D}P_i(t,S)=r(t,S)P_i(t,S)-b_i(t,S)$$

Where $\widehat{D}$ is the Dynkin associated to the process :

$$dS\left(t\right)=r(t,S)Sdt+\sigma(t,S)dB$$

Which is the Risk Neutral dynamics!

PROPOSITION 45. *The probability structure associated to the process :*

$$dS\left(t\right)=r(t,S)Sdt+\sigma(t,S)dB$$

Is such that the price of both assets is equal to the discounted value of their cash flow (here the resell value). This is the RN dynamics.
Any asset has a value $P_i\left(t,S\right)$ such that :

$$\hat{D}P_i(t,S)=r(t,S)P_i(t,S)-b_i(t,S)$$

4.3.2. The geometric brownian case.

We can solve the equation in a particular case :

$$dS=\mu Sdt+\sigma SdB$$

where r and σ are constant.

The risk neutral dynamics is :

$$dS=rSdt+\sigma SdB$$

That is :

$$S(T)=\exp\left[\ln\left(S\left(t\right)\right)+\left(r-\frac{1}{2}\sigma^2\right)\left(T-t\right)+\sigma\left(B(T)-B(t)\right)\right]$$

The arbitrage free EDP is hence :

$$DP=rP-b$$

We need to completely solve this equation a final value at T. Assume for instance that $P\left(T,x\right)=h\left(x\right)$ where h is given.

The price of an asset distributing $b\left(t,S\right)$ and with the terminal value $h\left(S\right)$ is :

$$P(t,S)=\mathbb{E}\left[\int_t^T\exp\left(-r\left(u-t\right)\right)b\left(s,S\left(u\right)\right)du+\exp(-r\left(T-t\right))h(S(T))/S(t)=S\right]$$

Take for instance a call for which the terminal value is : $h(S(T))=\left(S(T)-K\right)^+$. The formula gives :

$$C(t,S) = \exp(-r\,(T-t))\mathrm{E}\left[(S(T)-K)^{+}/S(t)=S\right]$$

It is relatively easy to compute this value.

LEMMA 46. *Let* $I = \mathbb{E}\left[(\exp(u)-K)^{+}\right]$ *where* u *is* $\mathcal{N}\left(m,s^2\right)$*, then :'*

$$I = e^{\mu+\frac{s^2}{2}}N\left(s+\frac{\mu-\ln(K)}{s}\right)-KN\left(\frac{\mu-\ln(K)}{s}\right)$$

Where N *is the cumulative distribution function of a standard normal.*

PROOF. We have

$$I = \frac{1}{\sqrt{2\pi}s}\int_{ln(K)}^{+\infty}\left(e^{u}-K\right)e^{-\frac{(u-\mu)^2}{2s^2}}du$$

take $v = \frac{u-\mu}{s}$:

$$I = \frac{1}{\sqrt{2\pi}}\int_{\frac{\ln(K)-\mu}{s}}^{+\infty}\left(e^{\frac{2st+2\mu-v^2}{2}}-Ke^{-\frac{v^2}{2}}\right)dv$$

$$I = \frac{1}{\sqrt{2\pi}}\int_{\frac{\ln(K)-\mu}{s}}^{+\infty}\left(e^{-\frac{(v-s)^2-2\mu-s^2}{2}}-Ke^{-\frac{v^2}{2}}\right)dv$$

$$I = \frac{1}{\sqrt{2\pi}}e^{\mu+\frac{s^2}{2}}\int_{\frac{\ln(K)-\mu}{s}}^{+\infty}e^{-\frac{(v-s)^2}{2}}dv-\frac{1}{\sqrt{2\pi}}\int_{\frac{\ln(K)-\mu}{s}}^{+\infty}Ke^{-\frac{vt^2}{2}}dv$$

$$I = e^{\mu+\frac{s^2}{2}}N\left(s-\frac{\ln(K)-\mu}{s}\right)-KN\left(\frac{\mu-\ln(K)}{s}\right)$$

□

We have then the Black and Scholes formula by taking :

$$\mu = \ln(S)+\left(r-\frac{1}{2}\sigma^2\right)(T-t)$$

$$s = \sigma\sqrt{T-t}$$

in the expression :

$$C(t,S) = \exp(-r\,(T-t))\mathrm{E}\left[(S(T)-K)^{+}/S(t)=S\right]$$

4.4. Appendix : conditional expectation

Our starting point is a random experiment modeled by a probability space $(\Omega, \mathscr{F}, \Pr)$, so that Ω is the sample space, $\mathscr{F}$ is the σ -algebra of events, and Pr is the probability measure on $(\Omega, \mathscr{F})$. To study the conditional expected value of a real-value random variable X, the more general approach is to condition on a sub σ-algebra $\mathscr{G}$ of $\mathscr{F}$.

Before we get to the definition, we will assume that any expected values that we mention exist. That is, if we write $\mathbb{E}(X)$ then we are assuming that X is a real-valued random variable and that $\mathbb{E}(|X|) < \infty$.

suppose that X is a random variable and that is a sub σ-algebra of $\mathscr{F}$. The conditional expected value of X given $\mathscr{G}$ is the random variable $\mathbb{E}(X/\mathscr{G})$ defined by the following properties:

- $\mathbb{E}(X/\mathscr{G})$ is measurable with respect to $\mathscr{G}$.
- If $A \in \mathscr{G}$ then $\mathbb{E}(\mathbb{I}_A X) = \mathbb{E}\left[\mathbb{I}_A \mathbb{E}(X/\mathscr{G})\right]$

The basic idea is that $\mathbb{E}(X/\mathscr{G})$ is the expected value of X given the information in the σ -algebra $\mathscr{G}$.

One can define the random variable "conditional probability" (which is not a measure) as : $\Pr(A/\mathscr{G}) = \mathbb{E}(\mathbb{I}_A/\mathscr{G})$

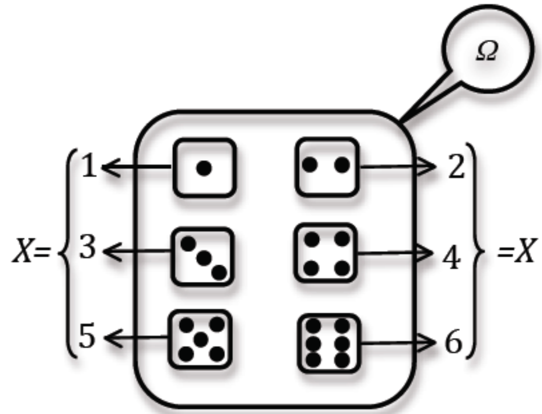
Main property : the next result is the main one: $\mathbb{E}(X/\mathscr{G})$ is closer to$X$ in the mean square sense than any other random variable that is measurable with respect to $\mathscr{G}$. Thus, if $\mathscr{G}$ represents the information that we have, then $\mathbb{E}(X/\mathscr{G})$ is the best we can do in estimating X.

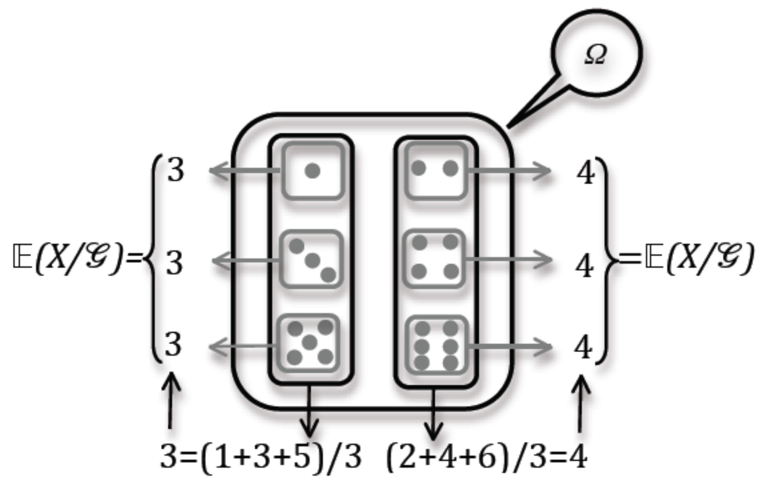
To make these results more concrete just look at the finite case wher Ω is composed of a dice drawing $\Omega = \{1,2,3,4,5,6\}$, $\mathscr{F} = 2^{\Omega}$ is the set of parts of Ω and X defined by $X(\omega) = \omega$.

We have :

$$\mathbb{E}[X] = \frac{1}{6}[1+2+3+4+5+6] = 3,5$$

Now assume that the sub σ-algebra $\mathscr{G}$ is the one associated to the partition $\{\{1,3,5\},\{2,4,6\}\}$. This σ-algebra has only 4 elements and corresponds to the information on the parity of the dice : odd or even. $\mathbb{E}[X/\mathscr{G}]$ must be $\mathscr{G}$measurable : it must take only two values, one for ω in $\{1,3,5\}$ and another for ω in $2,4,6\}$. The second condition tells us that if $A \in \mathscr{G}$ then $\mathbb{E}(\mathbb{I}_A X) = \mathbb{E}\left[\mathbb{I}_A \mathbb{E}(X/\mathscr{G})\right]$. One obtains that $\mathbb{E}(X/\mathscr{G})$ is a random variable such that for $\omega \in \{1,3,5\}$: $\frac{1+3+5}{6} = \frac{3}{6}\mathbb{E}(X/\mathscr{G})(\omega)$, i.e $\mathbb{E}(X/\mathscr{G})(\omega) = 3$, and for $\omega \in \{2,4,6\}$: $\frac{2+4+6}{6} = \frac{3}{6}E(X/\mathscr{G})(\omega)$, i.e $\mathbb{E}(X/\mathscr{G})(\omega) = 4$. One replaces X by its average in each subset of the σ-algebra.


Ω
1
2
X=
3
4
=X
5
6


Ω
3
4
𝔼(X/𝒢)=
3
4
=𝔼(X/𝒢)
3
4
3=(1+3+5)/3
(2+4+6)/3=4

CHAPTER 5

Microstructure and behaviour models

We depart here from the hypothesis of "always rational agents" and perfect markets.

5.1. The market efficiency hypothesis

The concept of efficiency for financial market corresponds to the idea that the current prices of financial assets perfecly reflect all the information available to investors. In other words, financial markets are efficient when an informed trader cannot make systematically more money than other who only observe current prices.

There are three different degrees of market efficiency. To explain this, take the example of one asset (distributing no dividends) whose price is p_t at date t. At date t the price p_{t+1} is "random".

The weakest notion of market efficiency says that price is a a "martingale" for some probability distribution :

$$p_t = \mathbb{E}\left(p_{t+1}/p_t\right)$$

The semi-strong efficiency says that the price at date t reflects all public information of date t:

$$p_t = \mathbb{E}\left(p_{t+1}/\mathcal{I}_t\right)$$

The strong efficiency assumption says that the price reflects all public and private information : the equilibrium process makes public all private information.

To precise these concepts we are going to study one very simple model proposed by Grosman and Stiglitz.

5.2. The Competitive REE

There is a single, risky asset with random liquidation value $\widetilde{\theta}$ and riskless asset (with unitary return). These are traded by risk averse agents and "noise traders." The utility derived by a trader i for the (random) profit $\pi_i = (\widetilde{\theta} - p)x_i$ of buying x_i units of the asset at price p is of the CARA type: $U(\pi_i) = -\exp -\rho_i \pi_i$, where$\rho_i > 0$ is the CARA coefficient that measures risk aversion. We call risk tolerance the inverse of risk aversion $t_i = \frac{1}{\rho_i}$. Initial wealth of each trader i is normalized to 0 (wlog). Trader i is endowed with a piece of private information about $\widetilde{\theta}$. Noise traders are assumed to trade for liquidity reasons submitting a random trade $\widetilde{u}$.

Suppose that a fraction of traders $\mu \in [0,1]$ receives a private signal $\widetilde{s}$ about $\widetilde{\theta}$, we call them Informed, subscript I, while the complementary fraction does not, Uninformed, subscript U . Both classes of traders condition their orders on the price p. Let $\rho_i = \rho_I > 0$ for Informed and $\rho_i = \rho_U \geq 0$, for Uninformed. $\widetilde{s}, \widetilde{\varepsilon}, \widetilde{u}$ are (pairwise independent) normally distributed:

$$\widetilde{s} \rightsquigarrow N(\overline{\theta}, \sigma_s^2)$$

$$\widetilde{\theta} = \widetilde{s} + \widetilde{\varepsilon} \quad , \widetilde{\varepsilon} \rightsquigarrow N(0, \sigma_\varepsilon^2)$$
$$\widetilde{u} \rightsquigarrow N(0, \sigma_u^2)$$

We call "precision" the inverse of the variance for $j = s, \varepsilon$ or u :

$$\tau_j = \frac{1}{\sigma_j^2}$$

If the price is p , what is the demand of a trader?

The quatity demanded maximizes the expected utility, the expectation being conditionnal to the information $\mathcal{J}$ detained :

$$X_i(p/J) = \arg\max \left(\mathbb{E} \left[U_i \left(\left(\widetilde{\theta} - p \right) x_i \right) / \mathcal{J} \right] \right)$$

But we know that $\widetilde{\theta}$ is normally distributed so that :

$$\mathbb{E} \left[U_i \left(\left(\widetilde{\theta} - p \right) x_i \right) / \mathcal{J} \right] = U_i \left(\mathbb{E} \left[\left(\widetilde{\theta} - p \right) x_i / \mathcal{J} \right] - \frac{1}{2} \rho_i \mathrm{var} \left(\left(\widetilde{\theta} - p \right) x_i / \mathcal{J} \right) \right)$$

So that i's maximization amounts to :

$$\max \left\{ \mathbb{E} \left[\left(\widetilde{\theta} - p \right) x_i / \mathcal{J} \right] - \frac{1}{2} \rho_i \mathrm{var} \left(\left(\widetilde{\theta} - p \right) x_i / \mathcal{J} \right) \right\}$$

which gives :

$$X_i(p/\mathcal{J}) = \frac{\mathbb{E} \left[\widetilde{\theta} / \mathcal{J} \right] - p}{\rho_i \mathrm{var} \left(\widetilde{\theta} / \mathcal{J} \right)}$$

This expression is quite intuitive : the demand is proportionnal to the spread between expected value and price. The coefficient of proportionality is large when risk aversion is low and/or risk (measured through variance) is low.

Before computing the equilibrium we must recall some simple formulas when random variables are normal.

Let $\widetilde{x}$,$\widetilde{y}$ a gaussian vector (i.e the variables are such that every linear combination is gaussian).

we have :

LEMMA. *$\widetilde{x}$,$\widetilde{y}$ a gaussian vector, we have :*

$$\mathbb{E}\left(\widetilde{x}/\widetilde{y}\right)-\mathbb{E}\left(\widetilde{x}\right)=\frac{cov\left(\widetilde{x},\widetilde{y}\right)}{var\ \left(\widetilde{y}\right)}\left[\widetilde{y}-\mathbb{E}\left(\widetilde{y}\right)\right]$$

$$var\ \left(\widetilde{x}/\widetilde{y}\right)=var\left(\widetilde{x}-\mathbb{E}\left(\widetilde{x}/\widetilde{y}\right)\right)=var\left(\widetilde{x}\right)-\frac{\left(cov\left(\widetilde{x},\widetilde{y}\right)\right)^{2}}{var\left(\widetilde{y}\right)}=\left(1-\frac{\left(cov\left(\widetilde{x},\widetilde{y}\right)\right)^{2}}{var\left(\widetilde{x}\right)var\left(\widetilde{y}\right)}\right)var\left(\widetilde{x}\right)$$

5.2.1. Naïve equilibrium. Now, consider a first "naive" equilibrium. Each trader optimize with his information. Uninformed has no information (noted $\mathcal{J}_U^0$) so that $\mathbb{E}\left[\widetilde{\theta}/\mathcal{J}_U^0\right]=\overline{s}$ and var $\left(\widetilde{\theta}/\mathcal{J}_U^0\right)=\sigma_s^2+\sigma_\varepsilon^2$:

$$X_U(p/\mathcal{J}_U^0)=\frac{\overline{\theta}-p}{\rho_U\left(\sigma_s^2+\sigma_\varepsilon^2\right)}=t_U\tau_\theta\left(\overline{\theta}-p\right)$$

(with $\frac{1}{\tau_\theta}==\frac{1}{\tau_s}+\frac{1}{\tau_\varepsilon}$)

Informed observe $\widetilde{s}$ so that $\mathbb{E}\left[\widetilde{\theta}/\mathcal{J}_I^0\right]=s$ and var $\left(\widetilde{\theta}/\mathcal{J}_I^0\right)=\sigma_\varepsilon^2$.

So :

$$X_I(p/\mathcal{J}_I^0)=\frac{s-p}{\rho_I\sigma_\varepsilon^2}=t_I\tau_\varepsilon\left(s-p\right)$$

The supply by noise trader is u. So that market clearing gives :

$$(1-\mu)t_U\tau_\theta\left(\overline{\theta}-p\right)+\mu t_I\tau_\varepsilon\left(s-p\right)=u$$

Which gives the price :

$$p=\frac{(1-\mu)t_U\tau_\theta\overline{\theta}+\mu t_I\tau_\varepsilon s-u}{(1-\mu)t_U\tau_\theta+\mu t_I\tau_\varepsilon}$$

The first remark that can be done is that when $\sigma_\varepsilon=0$, that is when the Informed traders are "perfectly" informed, then their demand curve is horizontal $p=s$. The only possible equilibrium is hence $p=s$ perfectly revealing their info.

In the general case, remark that the equilibrium price depends (linearly) on s and u. For instance, when $u=0$, p is larger than $\overline{\theta}$ when s is larger than $\overline{\theta}$, that is when Informed has "a good news about θ". This dependence implies that the price conveys information about s!

Indeed we have at equilibrium:

$$\widetilde{s} = \frac{\left[(1-\mu)t_U\tau_\theta + \mu t_I\tau_\varepsilon\right]\widetilde{p} - \left((1-\mu)\,\tau_\theta t_U\right)\overline{\theta} + \widetilde{u}}{\mu t_I\tau_\varepsilon}$$

So that the best prediction of s varies with p :

$$\mathbb{E}\left(\widetilde{s}/p\right) = \frac{\left[(1-\mu)t_U\tau_\theta + \mu t_I\tau_\varepsilon\right]p - \left((1-\mu)\,\tau_\theta t_U\right)\overline{\theta}}{\mu t_I\tau_\varepsilon}$$

In fact, Uninformed (but sophiticated) traders should have used this information to set their demand (which they have not at this first naïve stage).

But if they modify their demand function accordingly to take into account this information, this will modify the formula of the price equilibrium, function of s and u! this will in turn modify the information inferred by Uninformed...and so on!

The idea of Rational Expectation Equilibrium consists in finding a price formula which is "self fulfilling". If the price formula is $p = f(s,u)$ and if the traders use this information then the equilibrium price will be precisely $f(s,u)$!

Definition 47. A REE is a price function $p = f\,(s,u)$ such that, if traders know this price function, they infer information from price and set their demand accordingly. Doing this it turns out that equilibrium price will be precisely $f(s,u)$. In some sense, this type of equilibrium is the limit of the sequence of inference mentioned above.

5.2.2. Rational Expectation Equilibrium (Grossman and Stiglitz). The idea is to find a linear price formula $p = a + bs - \lambda u$ which is self fulfilling.

Let us be more precise.

Uninformed only observe p :

$$\mathbb{E}\left(\widetilde{\theta}/\mathcal{J}_U\right) = \mathbb{E}\left(\widetilde{\theta}/a + b\widetilde{s} - \lambda\widetilde{u} = p\right) = \mathbb{E}\left(\widetilde{s} + \widetilde{\varepsilon}/b\widetilde{s} - \lambda\widetilde{u} = p - a\right)$$

which gives, using the fact that all variables are normal :

$$\mathbb{E}\left(\widetilde{\theta}/\mathcal{J}_U\right) = \overline{\theta} + \frac{b\sigma_s^2}{b^2\sigma_s^2 + \lambda^2\sigma_u^2}\left(p - a - b\overline{\theta}\right)$$

or, equivalently

$$\mathbb{E}\left(\widetilde{\theta}/\mathcal{J}_U\right) = \overline{\theta} + \frac{b\tau_u}{b^2\tau_u + \lambda^2\tau_s}\left(p - a - b\overline{\theta}\right) = \overline{\theta} + \frac{b\tau_u}{k}\left(p - a - b\overline{\theta}\right)$$

In the above formula we have set $k = b^2\tau_u + \lambda^2\tau_s$, (which depends on b and λ).

In the same way :

$$\operatorname{var}\left(\widetilde{\theta}/\mathcal{J}_U\right) = \operatorname{var}\left(\widetilde{s} + \widetilde{\varepsilon}/b\widetilde{s} - \lambda\widetilde{u} = p - a\right) = \sigma_s^2 + \sigma_\varepsilon^2 - \frac{b^2\sigma_s^4}{b^2\sigma_s^2 + \lambda^2\sigma_u^2}$$

$$= \sigma_\varepsilon^2 + \frac{\lambda^2\sigma_u^2\sigma_s^2}{(b^2\sigma_s^2 + \lambda^2\sigma_u^2)} = \sigma_\varepsilon^2 + \frac{\lambda^2}{k}$$

The Uninformed demand is hence :

$$X_U(p/\mathcal{J}_U) = t_U \frac{1}{\sigma_\varepsilon^2 + \frac{\lambda^2}{k}} \left[\overline{\theta} + \frac{b\tau_u}{k}\left(p - a - b\overline{\theta}\right) - p\right]$$

The Informed has the same demand function :

$$X_I(p/\mathcal{J}_I) = t_I \tau_\varepsilon \left(s - p\right)$$

Market clearing gives :

$$\mu t_I \tau_\varepsilon \left(s - p\right) + (1-\mu) t_U \frac{1}{\sigma_\varepsilon^2 + \frac{\lambda^2}{k}} \left[\overline{\theta} + \frac{b\tau_u}{k}\left(p - a - b\overline{\theta}\right) - p\right] = u$$

Identifying with $p = a + bs - \lambda u$ gives for the coefficients of u and s , and for the constant a :

$$\mu t_I \tau_\varepsilon \lambda + (1-\mu) t_U \frac{1}{\sigma_\varepsilon^2 + \frac{\lambda^2}{k}} \left[\frac{-b\lambda\tau_u}{k} + \lambda\right] = 1$$

$$-\mu t_I \tau_\varepsilon b + (1-\mu) t_U \frac{1}{\sigma_\varepsilon^2 + \frac{\lambda^2}{k}} \left[\frac{b\tau_u}{k} b - b\right] = -\mu t_I \tau_\varepsilon$$

$$\left(-\mu t_I \tau_\varepsilon - (1-\mu) t_U \frac{1}{\sigma_\varepsilon^2 + \frac{\lambda^2}{k}}\right) a + (1-\mu) t_U \frac{\overline{\theta}\tau_s}{\sigma_\varepsilon^2 + \frac{\lambda^2}{k}} \left(\frac{\lambda^2}{k}\right) = 0$$

That gives :

$$\lambda \mu t_I \tau_\varepsilon = b \tag{5.2.1}$$

Such that :

$$p = a + \lambda\left(\mu t_I \tau_\varepsilon s - u\right)$$

reinjecting 5.2.1 in $k = b^2\tau_u + \lambda^2\tau_s$ gives :

$$k = \lambda^2 \left((\mu t_I \tau_\varepsilon)^2 \tau_u + \tau_s \right)$$

We have hence :

$$\mathrm{var}\left(\widetilde{\theta}/J_U\right) = \sigma_\varepsilon^2 + \frac{1}{(\mu t_I \tau_\varepsilon)^2 \tau_u + \tau_s} = \frac{1}{\tau}$$

which allows to find equations giving λ and a :

$$\mu t_I \tau_\varepsilon \lambda + (1-\mu) t_U \tau \left[-\frac{\mu t_I \tau_\varepsilon \tau_u}{\frac{1}{\tau} - \frac{1}{\tau_\varepsilon}} + \lambda \right] = 1$$

$$(-\mu t_I \tau_\varepsilon - (1-\mu) t_U \tau)\, a + (1-\mu) t_U \tau \overline{\theta} \left(\frac{\tau_s}{\frac{1}{\tau} - \frac{1}{\tau_\varepsilon}} \right) = 0$$

It is interesting that $w = s - \frac{1}{\mu t_I \tau_\varepsilon} u$ is the "noisy" information conveyed by p on s. Indeed :
$\mathbb{E}\,(w/s) = s$ and $\mathrm{var}\,(w/s) = \frac{1}{\mu t_I \tau_\varepsilon \tau_u}$

- Informed perfectly informed

Obviously, as in the naïve equilibrium, when τ_ε is infinite (Informed are perfectly informed) then the price gives perfect information on s (and $p = s$).

- No noisy traders

More interestingly, if there are no "noisy" traders $\sigma_u = 0$, then the price gives also perfect information on s. then $\mathrm{var}\left(\widetilde{\theta}/J_U\right) = \sigma_\varepsilon^2 = \mathrm{var}\left(\widetilde{\theta}/J_I\right)$

This gives also $p = s$! Indeed market clearing gives :

$$\mu t_I \tau_\varepsilon (s-p) + (1-\mu)\, t_U \tau_\varepsilon \left[\overline{\theta} + \frac{\left(p - a - b\overline{\theta}\right)}{b} - p \right] = 0$$

which implies $b = 1$ and $a = 0$ as soon as $\mu > 0$.

- No insider

What happens when $\mu = 0$ (no insider)? In that casewe have :

$$\lambda = \frac{\sigma_\theta^2}{t_U}$$

$$a = \overline{\theta}$$

Which means :

$$p = \overline{\theta} - \frac{\sigma_\theta^2}{t_U} u$$

To sum up :

PROPOSITION 48. *In the Rational Expectation Equilibrium, if $\sigma_u = 0$ (no noisy traders), and $\mu > 0$ then the equilibrium price is $p = s$. That means that the market is strongly efficient : the price reflects all public and private information. When $\mu = 0$ the equilibrium price is $\overline{\theta} - \frac{\sigma_\theta^2}{t_U} u$, which gives obviously $p = \overline{\theta}$ when there are no noisy traders.*

REMARK 49. The Grossman Stiglitz paradox. Suppose that prior to trading people decide weather to acquire or not the signal s at a fix cost k, and that there are no noisy traders. If $\mu = 0$ and if k is not too large, it is interesting to buy information : it can be easily shown indeed that the expected utility achieved when informed is larger than the one when everybody is non informed. But as soon as $\mu > 0$, the price becomes fully informative and it is not worth while to buy information since this information becomes free through price!

5.3. Bid ask spread (Glosten and Milgrom)‘

Asymmetry of information can be the source of some market characteristics.

Consider the following model. There is one risky asset whose liquidation value at date 1 is either $\underline{\theta}$ with probability π or $\overline{\theta}$ with probability 1-π. One unit is traded at time zero. There are three types of traders :

- informed traders : they know in advance the value θ at time 1 : they buy if the price is lower than θ and sell if the price is larger.

- liquidity traders: at any price they sell with probability $\frac{1}{2}$ and buy with probability $\frac{1}{2}$.

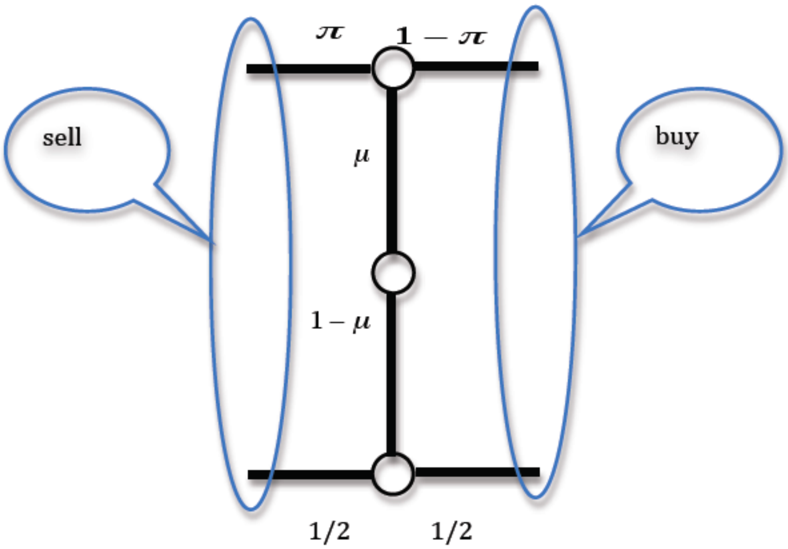
- market makers : they post a bid b (max price to buy) and an ask a (min price to sell) $b \leq a$.

The proportion of informed traders is μ.

The market maker does not know if he faces an informed or a liquidity trader.

Assume $\underline{\theta} \leq b \leq a \leq \overline{\theta}$

We have the following "tree" :



We read the diagram as the following : with probability μ the market maker faces an informed. In this case he will buy if $\theta = \overline{\theta}$ (with probability $1-\pi$) and sell if $\theta = \underline{\theta}$ (with porobability π). With probability $1-\mu$ this is an uninformed who sells or buys with probability $1/2$.

Assume first that $a = b$

With prob $1-\mu$ the market maker faces a non informed trader. The expected profit of the market maker will be : $\frac{1}{2}(a - \mathbb{E}[\theta]) + \frac{1}{2}(\mathbb{E}[\theta] - a) = 0$

With prob μ it is an informed : he buys if $\theta = \overline{\theta}$ and sells if $\theta = \underline{\theta}$

That is with probability $1-\pi$ the profit is $a - \overline{\theta}$ and with probability π the profit is $\underline{\theta} - b$ which are both negative values!

This comes from the asymmetry of information : inferior information of market makers implies negative profit if ask and bid are equal. A bid-ask spread allows to restore non negative

profit.

Indeed assume $\underline{\theta} \leq b < a \leq \overline{\theta}$. It is usefull to set the random variable D "demand" of the trader : this variable has two states $D =$sell and $D =$buy. In the graph above there are four terminal nodes : 2 when $D =$sell and 2 when $D =$buy.

The probabilities of each node is easy to compute.

D	Sell	Buy	total
Informed	$\mu\pi$	$\mu(1-\pi)$	μ
Liquidity	$\frac{1}{2}(1-\mu)$	$\frac{1}{2}(1-\mu)$	$1-\mu$
Total	$\mu\pi+\frac{1}{2}(1-\mu)$	$\mu(1-\pi)+\frac{1}{2}(1-\mu)$	1

The market maker profits are in each case :

Π	Sell	Buy
Informed	$\overline{\theta}-a$	$b-\underline{\theta}$
Liquidity	$\mathbb{E}[\theta]-a$	$b-E[\theta]$

To determine the values of a and b, Glosten and Milgrom assume that the "average" (expected) profit made when he sells (resp buys) is zero :

$$\mathbb{E}[\Pi/D]=0$$

We have $\Pr(\text{Informed}/D=\text{sell})=\frac{\mu\pi}{\mu\pi+\frac{1}{2}(1-\mu)}$ and $\Pr(\text{Liquidity}/D=\text{sell})=\frac{\frac{1}{2}(1-\mu)}{\mu\pi+\frac{1}{2}(1-\mu)}$

So that $\frac{\mu\pi}{\mu\pi+\frac{1}{2}(1-\mu)}\left(\overline{\theta}-a\right)+\frac{\frac{1}{2}(1-\mu)}{\mu\pi+\frac{1}{2}(1-\mu)}\left(\mathbb{E}[\theta]-a\right)=0$ which implies :

$$a=\frac{\mu\pi\overline{\theta}+\frac{1}{2}(1-\mu)\,\mathbb{E}[\theta]}{\mu\pi+\frac{1}{2}(1-\mu)}$$

In the same way :

$\frac{\mu(1-\pi)}{\mu(1-\pi)+\frac{1}{2}(1-\mu)}\left(b-\underline{\theta}\right)+\frac{\frac{1}{2}(1-\mu)}{\mu(1-\pi)+\frac{1}{2}(1-\mu)}\left(b-E[\theta]\right)=0$

$$b=\frac{\mu(1-\pi)\,\underline{\theta}+\frac{1}{2}(1-\mu)\,\mathbb{E}[\theta]}{\mu(1-\pi)+\frac{1}{2}(1-\mu)}$$

When μ is large, a is close to $\overline{\theta}$ and b is close to $\underline{\theta}$. Conversely when μ is small, a and b are close to $\mathbb{E}(\theta)$.

5.4. High Frequency trading : arm's race

The question we adress in this section are the following : does the prevalence of fast traders (who practice high frequency trading) enhance or deteriorate the functioning of markets? Do market forces lead to an optimal amount of investment in fast trading technologies? Is policy intervention called for?

On the one hand, highspeed market connections and information processing improve ability to seize trading opportunities, which raises gains from trade. From this viewpoint, HFT (high frequency trading) allows a better adjustment to equilibrium and hence is supposed to improve efficiency. But, on the other hand, when slow and fast traders are both active on a market, fast traders can process information before slow traders, which can generate information asymmetry and hence adverse selection. This can imply market inefficiency as for instance slow trader eviction. To avoid this "adverse selection" phenomenon, slow traders can be induced to (over)invest in fast trading technologies. Investment waves can arise, where institutions invest in fast trading technologies just to keep up with the others. When some traders become fast, it increases adverse selection costs for all, i.e., it generates negative externalities. Therefore equilibrium investment can exceed its welfare maximizing counterpart.

5.4.1. The model. Several institutions are present on the market : fast traders, slow traders and market makers that offer counterpart. Institutions can buy or sell only one share of one asset or abstain from trading. Their valuations for the asset are the sum of a common value component and a private value component. The common value, v, can be equal to $\mu + \varepsilon$ or $\mu - \varepsilon$ with equal probability. Private values (that means willingness to pay or to be payed) are i.i.d across institutions and can be equal to $v + \delta$ or $v - \delta$ with equal probability. Differences in private values capture in a simple way that other considerations than expected cashflows.

Fast traders access and process information on v before slow institutions. To capture this in the simplest possible way, we assume that, before trading, fast institutions observe whether $v = \mu + \varepsilon$ (good news) or $v = \mu - \varepsilon$ (bad news). Fast traders know in advance the liquidation value.

Slow institutions don't observe whether $v = \mu+\varepsilon$ or $v = \mu-\varepsilon$ and find a trading counterparty with probability $\rho < 1$.

The proportion of fast traders is α, of slow ones is $1 - \alpha$.

When an institution gets a trading opportunity, it decides whether to buy one share ($\omega = 1$), sell one share ($\omega = -1$), or abstain from trading ($\omega = 0$). We assume that the trade of the institution is executed at a price equal to the expectation of the common value conditional on its order: $\mathbb{E}(v/\omega)$, computed with rational expectations about all equilibrium strategies, as in Glosten and Milgrom. It amounts to asay that there is perfect competition among market makers. So that :

- the ask equilibrium price is :

$$a = \mathbb{E}\left(v/\omega = 1\right)$$

- the bid equilibrium price is :

$$b = \mathbb{E}\left(v/\omega = -1\right)$$

To solve the model, compute first the valuations for the different traders :

	good news	bad news
Fast traders with high valuation	$\mu+\delta+\varepsilon$	$\mu+\delta-\varepsilon$
Fast traders with low valuation	$\mu-\delta+\varepsilon$	$\mu-\delta-\varepsilon$
Slow traders with high valuation	$\mu+\delta$	
Slow traders with low valuation	$\mu-\delta$	

To develop the model, assume that $\varepsilon > \delta > \frac{\varepsilon}{2}$ which implies $\delta > \varepsilon - \delta$

So that we have :

$$\mu-\delta-\varepsilon < \mu-\delta < \mu+\delta-\varepsilon < \mu < \mu-\delta+\varepsilon < \mu+\delta < \mu+\delta+\varepsilon$$

The market maker faces different types of traders with the probabilities depicted in the following table.

		good news	bad news	
	probability	$\frac{1}{2}$	$\frac{1}{2}$	
Fast traders with high valuation	$\frac{\alpha}{2}$	$\frac{\alpha}{4}$	$\frac{\alpha}{4}$	
Fast traders with low valuation	$\frac{\alpha}{2}$	$\frac{\alpha}{4}$	$\frac{\alpha}{4}$	
Slow traders with high valuation	$\rho\frac{1-\alpha}{2}$	$\rho\frac{1-\alpha}{2}$		
Slow traders with low valuation	$\rho\frac{1-\alpha}{2}$	$\rho\frac{1-\alpha}{2}$		
Slow traders with no counterparty	$(1-\rho)(1-\alpha)$	$(1-\rho)(1-\alpha)$		
Total	1			

We are studying the case case of $\omega = 1$, that is the case where the trader, facing an ask price a decides to buy or not.

- If $\mu+\delta < a < \mu+\delta+\varepsilon$, only fast trader with high valuation and good news buy (cell 1 on the following table)
- If $\mu-\delta+\varepsilon < a < \mu+\delta$, fast trader with high valuation and good news and slow traders with high valuation buy (cells 1 and 2)
- $\mu+\delta-\varepsilon < a < \mu-\delta+\varepsilon$, fast trader with good news and slow traders with high valuation buy (cells 1,2 and 3)
- $\mu-\delta < a < \mu+\delta-\varepsilon$, fast trader with good news, fast traders with bad news and high valuation, and slow traders with high valuation buy (cells 1,2, 3 and 4)
- $\mu-\delta-\varepsilon < a < \mu-\delta$, everybody buys except fast trader with low valuation and bad news (cells 1,2, 3 ,4 and 5)

When price decreases	good news	bad news
Fast traders with high valuation	$1:\left(\frac{\alpha}{4}\right)$	$4:\left(\frac{\alpha}{4}\right)$
Fast traders with low valuation	$3:\left(\frac{\alpha}{4}\right)$	don't buy
Slow traders with high valuation	$2:\left(\rho\frac{1-\alpha}{2}\right)$	
Slow traders with low valuation	$5:\left(\rho\frac{1-\alpha}{2}\right)$	

On the following graph we show the demand (price on the vertical axis)

The demand D is hence :

- If $\mu+\delta<a<\mu+\delta+\varepsilon$: $D=\frac{\alpha}{4}$
  - $a=\mu+\delta$: $D=\frac{\alpha}{4}+\beta_H\rho\frac{1-\alpha}{2}$, where β_H is the probability a slow trader with high valuation buys.
- If $\mu-\delta+\varepsilon<a<\mu+\delta$, $D=\frac{\alpha}{4}+\rho\frac{1-\alpha}{2}$
  - $a=\mu-\delta+\varepsilon$, $D=\frac{\alpha}{4}+\rho\frac{1-\alpha}{2}+\beta_{GL}\frac{\alpha}{4}$, where β_{GL} is the probability a fast trader with low valuation and good news buys
- $\mu+\delta-\varepsilon<a<\mu-\delta+\varepsilon$, $D=\frac{\alpha}{2}+\rho\frac{1-\alpha}{2}$
  - $a=\mu+\delta-\varepsilon$, $D=\frac{\alpha}{2}+\rho\frac{1-\alpha}{2}+\beta_{BH}\frac{\alpha}{4}$, where β_{BH} is the probability a fast trader with high valuation and bad news buys
- $\mu-\delta<a<\mu+\delta-\varepsilon$, $D=\frac{3\alpha}{4}+\rho\frac{1-\alpha}{2}$
  - $a=\mu-\delta$, $D=\frac{3\alpha}{4}+\rho\frac{1-\alpha}{2}+\beta_L\rho\frac{1-\alpha}{2}$, where β_L is the probability a slow trader with low valuation buys

Let us compute the different probabilities. The first one is the probability that the trader buys, facing $a<\mu+\delta+\varepsilon$, that is the cells 1,2,3,4 and 5.

$$\Pr(\omega=1)=\frac{\alpha}{4}+\beta_{BH}\frac{\alpha}{4}+\beta_{GL}\frac{\alpha}{4}+\rho\frac{1-\alpha}{2}\left(\beta_H+\beta_L\right)$$

Among these, some cases are such that $v=\mu+\varepsilon$:

$$\Pr\left(v=\mu+\varepsilon/\omega=1\right)\Pr\left(\omega=1\right)=\frac{\alpha}{4}+\beta_{GL}\frac{\alpha}{4}+\rho\frac{(1-\alpha)}{4}\left(\beta_H+\beta_L\right)$$

And others where $v=\mu-\varepsilon$

$$\Pr\left(v=\mu-\varepsilon/\omega=1\right)\Pr(\omega=1)=\beta_{BH}\frac{\alpha}{4}+\rho\frac{(1-\alpha)}{4}\left(\beta_H+\beta_L\right)$$

So that the value $\mathbb{E}\left(v/\omega=1\right)$ is easily computed :

$$\mathbb{E}\left(v/\omega=1\right)=\mu+\frac{\alpha\left(1+\beta_{GL}-\beta_{BH}\right)}{\alpha\left(1+\beta_{BH}+\beta_{GL}\right)+2\rho\left(1-\alpha\right)\left(\beta_H+\beta_L\right)}\varepsilon$$

There is a first (intuitive) equilibrium where $a^* = \mu + \varepsilon$. In that case $\beta_{BH} = \beta_{GL} = \beta_H = \beta_L = 0$, and we have $\mu + \delta < a^* < \mu + \delta + \varepsilon$, only fast traders with good news and high valuation buy. That means that no trade occurs in the other cases.

THEOREM 50. *For any value of the parameter α there always exists a "crowding out" equilibrium where only fast traders with high valuation and good news buy : .$a^* = \mu + \varepsilon$. We call that equilibrium P1*

Now describe other possible equilibria

THEOREM 51. *When $\alpha \leq \frac{2\rho\delta}{\varepsilon-\delta+2\rho\delta}$, then there exists an equilibrium such that $\beta_{GL} = \beta_{BL} = \beta_L = 0$ and $\beta_H = \frac{\alpha(\varepsilon-\delta)}{2\rho(1-\alpha)\delta}$, for which the equlibrium price is $a^* = \mu + \delta$. This equilibrium is called M1*

For lower values of α:

THEOREM 52. *For lower values of α $\frac{2\rho(\varepsilon-\delta)}{\delta+2\rho(\varepsilon-\delta)} \leq \alpha \leq \frac{2\rho\delta}{\varepsilon-\delta+2\rho\delta}$ there exists an equilibrium such that $\beta_{GL} = \beta_{BL} = \beta_L = 0$ and $\beta_H = 1$ for which the equlibrium price is $a^* = \mu + \frac{\alpha}{\alpha+2\rho(1-\alpha)}\varepsilon$. This equilibrium is called P2*

There are other equilibria for lower values of α. We don't derive them and study now the incentive to invest in fast technologies. The question is : when α is given, does a low trader have incentive to invest in fast trading strategies?

The expected payoffs are for fast traders : $\frac{1}{2}(\mu + \varepsilon + \delta - a^*) + \frac{1}{2}(\mu + \varepsilon - \delta - a^*)\beta_{GL}$, and for low traders $\rho(\mu + \delta - a^*)\beta_H$, the gives for P1 and P2 equilibria :

For P1 : fast obtain $\frac{1}{2}\delta$ and low obtain 0.

For P2 : fast obtain $\frac{1}{2}\left(\varepsilon + \delta - \frac{\alpha\varepsilon}{\alpha+2\rho(1-\alpha)}\right)$ and slow : $\rho\left(\delta - \frac{\alpha\varepsilon}{\alpha+2\rho(1-\alpha)}\right)$,

We remark that that when $\frac{2\rho(\varepsilon-\delta)}{\delta+2\rho(\varepsilon-\delta)} \leq \alpha \leq \frac{2\rho\delta}{\varepsilon-\delta+2\rho\delta}$, P2 equilibrium is better than P1 for both types of traders. We assume that P2 will be the actual equilibrium for this values of α.

Consider a slow trader. What is is expected increase of surplus if he invests in fast technologies? The values above give that :

When $\frac{2\rho(\varepsilon-\delta)}{\delta+2\rho(\varepsilon-\delta)} \leq \alpha \leq \frac{2\rho\delta}{\varepsilon-\delta+2\rho\delta}$ the gain is : $\frac{1}{2}\left(\varepsilon + \delta - \frac{\alpha\varepsilon}{\alpha+2\rho(1-\alpha)}\right)$ -$\rho\left(\delta - \frac{\alpha\varepsilon}{\alpha+2\rho(1-\alpha)}\right)$ $\epsilon\rho + \delta(1-2\rho)$

: δ -$\rho(2\delta - \varepsilon)$

5.5. The capital asset pricing model

5.5.1. The principle of diversification. The principle of diversification is a central principle in finance and insurance.

The purpose of this first section is to give a meaning to the phrase “do not put all your eggs in one basket” , which is the popular expression of the principle of diversification.

The idea of diversification principle is actually quite simple, and makes a clear statement: (unless they are perfectly correlated) half the sum of two identically distributed random variables, having a finite expectation, is less risky than each of them. for example, if we imagine two baskets each with the same probability p to fall (and thus cause the loss of eggs) , put an egg in each basket is less risky than putting them both in one .

Indeed in the first case it will be possible to eat 0 eggs with probability p^2 , 2 eggs with probability $(1-p)^2\}$, And 1 egg with probability $2p(1-p)$. In the second 0 with probability p and 2 with probability $1-p$. The probabilities of the extreme, 0 and 2 , decreased : p to p^2 a decrease of $p(1-p)$) and $(1-p)$ to $(1-p)^2$, while the medium event, 1 egg , has its probability increased of exactly $2p(1-p)$.

If $\widetilde{X}$ is the random variable giving 1 if the first basket remains intact and 0 if it falls, $\widetilde{Y}$ defined in the same manner for the second rack , is a $2\widetilde{X}$ and 2 $\widetilde{Y}$ are riskier than $\widetilde{X}+\widetilde{Y}$.

This notion of risk reduction is associated to a “concentration” of the probability distribution function. It can be shown that if $\widetilde{A}$ and $\widetilde{B}$ have the same finite expectation, $\widetilde{A}$ is more risky than $\widetilde{B}$ if and only if there exists a random variable $\widetilde{\varepsilon}$ such that $\widetilde{A}=\widetilde{B}+\widetilde{\varepsilon}$ with $\mathbb{E}\left(\widetilde{\varepsilon}/\widetilde{B}\right)=0$. $\widetilde{A}$ is a noisy transformation of $\widetilde{B}$. Obviuously, if the variables have finite variances this implies that $\operatorname{var}\left(\widetilde{A}\right)\geq\operatorname{var}\left(\widetilde{B}\right)$

In the case of n independent and identically distributed random variables we can state the following general result :

Proposition 53. *if $\widetilde{x}_i$ are n independent real random variables and identically distributed such that $|\mathbb{E}(\widetilde{x}_i)|<+\infty$, then $\forall\alpha_i$, N positive real numbers of sum 1 ($\sum\alpha_i=1$) $\widetilde{x}_{\frac{1}{n}}=\frac{\sum\widetilde{x}_i}{n}$ is less risky than $\widetilde{x}_\alpha=\sum\alpha_i\widetilde{x}_i$. (and in particular than each of the $\widetilde{x}_i$) : that means that the probability distribution of $\widetilde{x}_{\frac{1}{n}}$ is more concentrated than that of $\widetilde{x}_\alpha$ or alternatively that, for all α, there exists $\widetilde{\varepsilon}$ such that $\widetilde{x}_\alpha=\widetilde{x}_{\frac{1}{n}}+\widetilde{\varepsilon}$ with $\mathbb{E}\left(\widetilde{\varepsilon}/\widetilde{x}_{\frac{1}{n}}\right)=0$.*

In particular if the variables have a finite variance σ^2 we have :

$$\operatorname{var}(\widetilde{x}_\alpha)=\left(\sum\alpha_i^2\right)\sigma^2$$

Which is minimum for $\alpha_i=\frac{1}{n}$

Obviously the situation is slightly more complex when the random variables are correlated. First, consider two random variables $\widetilde{x}_1$ and $\widetilde{x}_2$ with the same variance σ^2 but not necessarily independent . A study of variance allows to get an idea of the risk of a convex combination of the two variables.

Consider $t(\alpha)$:

$$t(\alpha)=\frac{\operatorname{var}(\alpha\widetilde{x}_1+(1-\alpha)\widetilde{x}_2)}{\sigma^2}$$

$$= \alpha^2 + (1-\alpha)^2 + 2\alpha(1-\alpha)\frac{\mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2)}{\sigma^2}$$

$t(\alpha)$ is less than 1, because $\mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2) < \sigma^2$. it is minimal for $\alpha = 1/2$.

But now if , $\widetilde{x}_1$ and $\widetilde{x}_2$ have not the same variance, with for example $\mathrm{var}(\widetilde{x}_1) \leq \mathrm{var}(\widetilde{x}_2)$. One computes :

$$\Sigma^2(\alpha) = \mathrm{var}\left(\alpha\widetilde{x}_1 + (1-\alpha)\widetilde{x}_2\right)$$

$$= \alpha^2 \mathrm{var}(\widetilde{x}_1) + (1-\alpha)^2 \mathrm{var}\left(\widetilde{x}_2\right) + 2\alpha(1-\alpha)\ \mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2)$$

$$= \alpha^2\left(\mathrm{var}(\widetilde{x}_1) + \mathrm{var}\left(\widetilde{x}_2\right) - 2\mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2)\right) + 2\alpha\left(\mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2) - \mathrm{var}\left(\widetilde{x}_2\right)\right) + \mathrm{var}\left(\widetilde{x}_2\right)$$

As this must be always positive, the discriminant must be negative :

$$|\mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2)| \leq \sqrt{\mathrm{var}\left(\widetilde{x}_1\right) \mathrm{var}\left(\widetilde{x}_2\right)} \leq \mathrm{var}\left(\widetilde{x}_2\right)$$

This means that the correlation coefficient $\frac{\mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2)}{\sqrt{\mathrm{var}(\widetilde{x}_1)\mathrm{var}(\widetilde{x}_2)}}$ lies between -1 and 1!

We have :

$$\frac{d\Sigma^2}{d\alpha} = \alpha\left(var(\widetilde{x}_1) - \mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2)\right) - (1-\alpha)\left(var(\widetilde{x}_2) - \mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2)\right)$$

The minimum is obtained for :

$$\frac{\alpha^*}{1-\alpha^*} = \frac{var(\widetilde{x}_2) - \mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2)}{var(\widetilde{x}_1) - \mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2)}$$

That is

$$\alpha^* = \frac{var(\widetilde{x}_2) - \mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2)}{var(\widetilde{x}_1) + var(\widetilde{x}_2) - 2\mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2)}$$

α^* belongs to $[0, 1]$ if :

$$\mathrm{var}\left(\widetilde{x}_1\right) - \ \mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2) \geq 0$$

That we write :

$$\frac{\mathrm{cov}(\widetilde{x}_1, \widetilde{x}_2)}{\mathrm{var}\left(\widetilde{x}_1\right)} \leq 1$$

$\frac{\operatorname{cov}(\widetilde{x}_1,\widetilde{x}_2)}{\operatorname{var}(\widetilde{x}_1)}$ is the "beta" coefficient $\beta(\widetilde{x}_2/\widetilde{x}_1)$ of 2 with respect to 1 : even if variance of 2 is larger than that of 1, a combination of the two assets allows to decrease the risk below that of 1 if $\beta(\widetilde{x}_2/\widetilde{x}_1) \leq 1$. If $\tau = \frac{\operatorname{cov}(\widetilde{x}_1,\widetilde{x}_2)}{\sqrt{\operatorname{var}(\widetilde{x}_1)\operatorname{var}(\widetilde{x}_2)}}$ is the correlation coefficient, this means that the risk can be diminihed below the lowest risk if $\tau\sqrt{var\,(\widetilde{x}_2)} \leq \sqrt{var\,(\widetilde{x}_1)}$ (which is always true if $\tau \leq 0$).

However, if it is not the case, the minimal risk is obtained for $\alpha \notin [0,1]$.

5.5.2. Portfolio choice. In this section we propose to analyze the problem of portfolio choice in general . An investor has 1 euro, how should it be allocated among the various assets available? Obviously the answer depends on his attitude to risk. We will assume here that our investor uses the mean-variance criterion, i.e. he evaluates $\mathbb{E}(\widetilde{v}) - \frac{1}{2}\theta\operatorname{var}(\widetilde{v})$ to compare the random variables. This is the case, in particular, when all variables are gaussian and the decision maker has a concave utility function.

Consider K financial assets ,$k = 1, ..., K$. Income from asset k is a real random variable $\widetilde{a}_k$: the income (cash) that provides the risky asset in the future. This random variable is assumed to be known through statistical studies. We note p_k the market price of asset k . It is quite convenient to define the return on assets k as the random variable that measures the income for one euro : $\widetilde{R}_k = \frac{\widetilde{a}_k}{p_k}$. There is also a risk-free asset , the asset 0 , which givesR_0 (nonrandom) euros per euro invested . A risky portfolio is a vector θ, each component θ_k measuring the amount of asset k held.

The random income of a portfolio is $\sum \widetilde{a}_k\theta_k$ and its cost ${}^t\theta p$.

One can write the income as a function of yields :

$$\widetilde{v} = \frac{\widetilde{a}_k}{p_k}p_k\theta_k = \widetilde{R}_k x_k = {}^t x\widetilde{R}$$

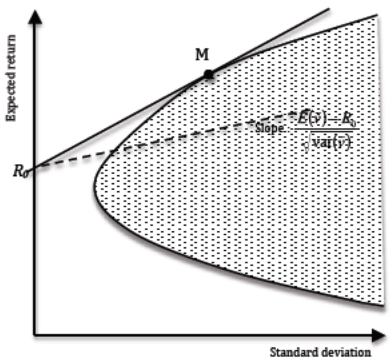
Where$x_k = p_k\theta_k$ is the expenditure to purchase the asset k.

Suppose our investor has a euro to be shared among different assets. He must choose to spread this euro between risk-free asset (x_{0}) and risky assets$x = (x_1, ...x_K)$ with $\mathbf{1}$ is the vector with$K$ components all equal to 1 :$x_0 + {}^t x\mathbf{1} = 1$. This strategy gives him for one euro , an income equal to $x_0R_0 +^t x\widetilde{R}. = R_0 + {}^t x(\widetilde{R} - R_0\mathbf{1})$.

We se t$\widetilde{\rho} = \widetilde{R} - R_0\mathbf{1}$, the vector of excess returns over the risk-free asset . Income from invested euro is equal to $R_0 + {}^t x\widetilde{\rho}$.

The problem for the investor is to choose the optimal x.

Before doing computations, use a graphical reasonning. The following graph represents each portfolio by a point whose coordinates are the standard deviation (that is the root square of the variance $\sqrt{\operatorname{var}(\widetilde{v})}$) of its return on the horizontal axis and the expectation $\mathbb{E}\,(\widetilde{v})$ on the vertical.



The shaded area is the set of all the possible portfolios containing only risky assets. Consider then a portfolio in that area and combinate it with the risk free asset. Say for instance that you put $1 - x_0$ euros on this portfolio and x_0 on the risk free asset. The expectation will be $\mathbb{E}(\widetilde{w}) = x_0 R_0 + (1-x_0)\mathbb{E}(\widetilde{v})$ and the standard deviation $\sqrt{\text{var}(\widetilde{w})} = (1-x_0)\sqrt{\text{var}(\widetilde{v})}$. The expected excess return is $\mathbb{E}(\widetilde{w}) = (1-x_0)(\mathbb{E}(\widetilde{v}) - R_0)$ so that this strategy is such that $\frac{\mathbb{E}(\widetilde{w})-R_0}{\sqrt{\text{var}(\widetilde{w})}} = \frac{\mathbb{E}(\widetilde{v})-R_0}{\sqrt{\text{var}(\widetilde{v})}}$. The point obtained is on the line between the point $(0, R_0)$ and the point $\left(\sqrt{\text{var}(\widetilde{v})}, \mathbb{E}(\widetilde{v})\right)$ whose slope is exactly $\frac{\mathbb{E}(\widetilde{v})-R_0}{\sqrt{\text{var}(\widetilde{v})}}$ which is called "the Sharpe ratio" of the portfolio. Hence, all the possible combinations are obtained by drawing lines between the shaded area and the point $(0, R_0)$.

For portfolio with a given standard deviation (a given risk) we seek those with the maximum expected return. To do this consider the point M which has the maximal Sharpe ratio in the shaded area. combining this portfolio with the risk free asset give portfolios on the bold line. Obviouly, the region above this line cannot be reached : there are no feasible combination that gives these expectation and standard deviation. Conversely, the points below are all feasible.

The bold line is hence the "efficiency frontier" of the market : efficient points (those with maximal expected return at agiven risk) are on this line. M is the market portfolio : all the investors share their investment between the risk free asset and this particular portfolio.

This result can be derived more formally.

We have :

$$\begin{cases} E(\widetilde{w}) = R_0 + {}^t x E(\widetilde{\rho}) \\ \text{var}(\widetilde{w}) = E\left[(\widetilde{v} - E(\widetilde{v}))^2\right] \end{cases}$$

$$\begin{cases} \text{var}(\widetilde{w}) = E\left[({}^t x\widetilde{R} - {}^t xE(\widetilde{R}))^2\right] \\ = E\left[{}^t x(\widetilde{R} - E(\widetilde{R}))^t(\widetilde{R} - E(\widetilde{R}))x\right] \\ = {}^t xE\left[(\widetilde{R} - E(\widetilde{R}))^t(\widetilde{R} - E(\widetilde{R}))\right] x \end{cases}$$

set

$$\Omega = E\left[(\widetilde{R} - E(\widetilde{R}))^t(\widetilde{R} - E(\widetilde{R}))\right]$$

Ω is called the matrix (symmetric) of variance-covariance of assets. The ij element equals $\sigma_{ij} = E\left[(\widetilde{R}_i - E(\widetilde{R}_i))(\widetilde{R}_j - E(\widetilde{R}_j))\right]$ σ_{ij} is the covariance between assets i and j . The formula above shows that this symmetric matrix is positive (the associated quadratic form is positive : $x^t\Omega x$ is a variance , which is positive) .

PROPOSITION 54. *To sum up, the strategy* (x_0, x) *gives* $R_0 +^t xE(\widetilde{\rho})$ *on average with a variance equal to* ${}^t x\Omega x$.

In the following if $\widetilde{v}$ is a random variable , we note v its expectation. Here $\rho_i = E(\widetilde{\rho}_i), \quad R_i = E(\widetilde{R}_i)$.

How to choose between all possible strategies ? Clearly two strategies give the same expectation , any risk-averse investor prefers the strategy of minimum variance. Therefore fix our investor expected return at m and seek the vectors x of $\mathbb{R}^K$ that minimizes the variance and give m as expected return . Consider the optimization problem (P) :

$$(P)\begin{cases} \min\limits_x({}^t x\Omega x) \\ {}^t x\rho + R_0 = m \end{cases}$$

Define a new scalar product :

DEFINITION 55. $< x, y >= {}^t x\Omega y$ is a scalar product (quadratic positive definite form) note $\| x \|$, the associated norm.

The problem becomes :

$$(P)\begin{cases} \min\limits_x \| x \| \\ < \Omega^{-1}\rho, x >= m - R_0 \end{cases}$$

This amounts to find the point of the affine hyperplane ($< \Omega^{-1}\rho, x >= m - R_0$), that is the closest from 0. This point is the orthogonal projection x^* of 0 on this hyperplane . It is defined by two equations with unknown x^* and λ :

$$\begin{cases} x^* = \lambda\Omega^{-1}\rho \\ < \Omega^{-1}\rho, x^* >= m - R_0 \end{cases}$$

The first one says that x^* is colinear with the orthogonal vector $\Omega^{-1}\rho$ of the hyperplane, The second says that tis projection belongs to this affine hyperplane.

- An important note should be made x^* is a vector which is proportional to the vector $\Omega^{-1}\rho$ which does not depend on m . In other words, regardless of the expected return required , the structure of the risky portfolio is identical. Structure refers to the relative proportion of different risky assets .
- Of course we can easily solve the above system :

$$x^* = \frac{(m - R_0)}{{}^t\rho\Omega^{-1}\rho}\Omega^{-1}\rho$$

- How our investor chooses the level he m ? Obviously this is determined by his tradeoff between mean and variance , maximize wrt to m :

$$E(\widetilde{w}) - \frac{1}{2}\theta var(\widetilde{w}) = m - \frac{1}{2}\theta(m - R_0)^2\frac{1}{{}^t\rho\Omega^{-1}\rho}$$

Definition 56. One calls market portfolio the portfolio $x^m = \frac{\Omega^{-1}\rho}{{}^t\mathbf{1}\Omega^{-1}\rho} = \mu\Omega^{-1}\rho$, that portfolio contains only risky assets in the relative proportions defined by the solutions of problems (P). The yield of this portfolio is : $\widetilde{R}_m = R_0 + {}^tx^m\widetilde{\rho}$, The variance is :

$$var(\widetilde{R}_m) = {}^tx^m\Omega x^m$$

Moreover:

$$\begin{aligned}(\Omega x^m)_i &= \sum_j \sigma_{ij}x_j^m \\ &= cov(\widetilde{R}_i, \sum \widetilde{R}_j x_j^m) \\ &= cov(\widetilde{R}_i, \widetilde{R}_m)\end{aligned}$$

This portfolio is called the market portfolio because, under the investors mean-variance assumption , the above shows that all individuals demand a portfolio whose risky component is proportional to this portfolio. It follows that the total demand of all the portfolios held the same structure (in their risky part) . Of course, a very risk-averse individual will ask a relative little amount of this risky portfolio and focus his investment on the risk-free asset . Instead, a less risk-averse individual will choose a x_0 smaller. It is as if each investor was buying a piece of the total market capitalization , piece more or less according to risk aversion !

5.5.3. Capital asset pricing model formula. The above leads to find one of the most famous of finance formulas.

One has :

$$\begin{aligned}\Omega x^m &= \mu\rho \\ var(\widetilde{R}_m) &= {}^tx^m\Omega x^m = \mu{}^tx^m\rho = \mu(R_m - R_0)\end{aligned}$$

This implies :

$$\frac{\Omega x^m}{var(\widetilde{R}_m)}(R_m - R_0) = \rho$$

Coordinate by coordinate:

$$\frac{cov(\widetilde{R}_i, \widetilde{R}_m)}{var(\widetilde{R}_m)}(R_m - R_0) = (R_i - R_0)$$

PROPOSITION 57. *For any asset i its average outperformance over the risk-free asset $R_i - R_0$ is proportional to the outperformance of the market portfolio$R_m - R_0$. The proportionality factor is the coefficient β_i , relative to the market portfolio .*

$$\begin{aligned} R_i - R_0 &= \beta_i(R_m - R_0) \\ \beta_i &= \frac{cov(\widetilde{R}_i, \widetilde{R}_m)}{var(\widetilde{R}_m)} \end{aligned}$$