NAME: LORENZO CONTI

ID: 244291

TITLE: OPTION PRICING AND THE MONTECARLO METHOD

1) INTRODUCTION

Options are financial instruments with a centennial history. The first examples of such contracts can be traced back to the 14th century with conditional sales where the insurer agreed to purchase ship or cargo if it failed to arrive (a sort of put option). In the 1400s the Cerchi bank of Florence was already involved in the trading of call options.

An option is a financial contract which gives the owner the right to buy or sell the underlying asset at a pre-specified price (strike price K) on a specified date t. The underlying asset might simply be a stock or even an index, currency or commodity. Regardless of the nature of the underlying asset, the price of the option is affected by the price changes of the underlying asset. We can distinguish two different types of options:

- CALL OPTION: it gives the owner the right to buy the underlying asset at the pre-specified price k. The payoff of a call option with maturity t, strike price k can be expressed as max (S_t - k,0) where S_t is the price of the option at time t.
- PUT OPTION: it gives the owner the right to sell the underlying asset at the pre-specified price k. The payoff of a put option can be written as max (k-S_t,0).

A further distinction can be made between European-style and American-style options where the former can only be exercised at the expiration date whereas the latter can be exercised at any time until the expiration date. A third famous type of options is represented by the Asian ones which are an example of path-dependent options: a path dependent option is an option whose payoff is affected by how the price of the underlying asset at maturity was reached, in other words the price path of the underlying asset. More specifically, the payoff of an Asian option depends on the average price of the underlying asset over a pre-specified period of time during the lifetime of the option. Regardless of the kind of option that we are considering (American, European, Asian etc.)

we want to come up with a "fair" prediction for the option's initial price, fair in the sense that it does not introduce any arbitrage possibility.

One of the first and most important studies in the field that is worth citing is represented by Bachelier's Thesis in which Louis Bachelier came up with concepts that are fundamental for the problem of option pricing such as:

- price increments are independent and normally distributed (so the price process is Brownian motion as the diffusion limit of a random walk), by applying the CLT.
- the importance of recognizing and taking into account the concept of arbitrage.

The work was later cited and used by Doob who is generally considered the "father" of martingales.

2) DETAILED OVERVIEW OF UNDERLYING MATHEMATICAL CONCEPTS

In this section, i will go through a detailed presentation of all the underlying mathematical concepts that will be cited in subsequent sections.

1) PROBABILITY SPACE

A probability space can be thought of as a mathematical model for the phenomenon of interest. Choosing a suitable model is a fundamental step as it allows us to compute all the relevant quantities such as mean and variance besides the dependence structure among the various variables involved.

A probability space is a triplet (Ω,F,P) where:

1) Ω is the set of possible outcomes. In our case Ω will be a path space since we want to model the price evolution of some asset's price. (options's underlying asset)

2) F is a collection of subsets of Ω which we would like to measure. F is also referred to as 'the information'. This collection contains Ω and is required to be closed for countable unions and complement

3) P is the probability. If A is an event, P(A) expresses the chances of A to happen. The properties of P are:

- $P : F \to [0, 1]$
- $P(\Omega) = 1$ (normalization property)
- P is countably additive: if $(A_n)_n$ is a sequence of disjoint events, then the probability of the sequence is equal to the sum of each event's probability.

$P(\sqcup_n A_n)$ = sum over n $P(A_n)$

$P(A^c) = 1 - P(A)$ where A^c is the complement $\Omega \setminus A$ of the event A.

If two events are not disjoint:

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

2) RANDOM VARIABLES

A random variable on (Ω,F) is a function on Ω which takes values in R:

$$X: \Omega \to R$$

and is F-measurable, meaning that the counter-image of any half line $(-\infty,x]$ is an event:

$$\{X \leq x\} \in F$$

for all $x \in R$

If we only need a sub-collection of events, we can define a sub-sigma-algebra G of F if it contains Ω (the set of possible outcomes) and is closed for countable unions of events and complement. A sub-sigma algebra is a collection of events which is a sigma-algebra in its turn. If a random variable X verifies the following condition:

$$\{X \leq x\} \in G$$

for all $x \in R$,

we say that X is G-measurable.

Lastly, the sigma-algebra generated by a random variable X is

$$\sigma(X) := \sigma(\{X \leq x\} | x \in R)$$

and it is a sub-sigma algebra of F. Any event which can be written in terms of X, for example:

$$\{a \leq X \leq b\}$$

belongs to σ(X).

3) FILTRATION

A filtration is an increasing collection of sub-σ-algebras F_t of F such that:

$$F_{t1} \subseteq F_{t2} \text{ if } t_1 < t_2$$

t is a time parameter, which may be discrete or continuous.

If t_1 is fixed (for example 3 months from now) then F_{t1} contains all the information about what may happen in $[0, t_1]$. If t_2 = 6 months from now, then F_{t2} describes what may happen in $[0,t_2]$. As any event happened by time t_1 may be seen as an event happened before t_2, F_{t2} is a bigger collection of events than F_{t1}.

In Finance, a finite time horizon T is often fixed. We may choose T as the longest maturity of the derivatives in our portfolio. If we assume $F_T = F$ (information set), we then have a filtered probability space identified by the triplets $(\Omega, (F_t)_{t \leq T}, P)$ which is also called a stochastic basis.

4) STOCHASTIC PROCESS

Given a stochastic basis identified by the triplets $(\Omega, (F_t)_{t \leq T}, P)$, a real-valued stochastic process S $= (S(t))_t$ is simply a collection of real valued random variables, measurable functions from (Ω, F_T) to R (or to R^d for d-dimensional processes).

A process S is said to be random, non-anticipative and Ft -measurable if the following conditions hold

1. for any fixed time t, $S(t) : \Omega \rightarrow R$
2. for all fixed real x, the set $\{S(t) \leq x\}$ belongs to F_t.

Furthermore, a process S can be thought of as a function of two variables, time t and outcome ω:

$$S = S(t)(\omega) = S(t, \omega)$$

1. when $t = t^*$ is fixed, the result is the random variable $S(t^*) = S(t^*, \cdot)$, the t^*-marginal of the process S;

2. when $\omega = \omega^*$ is fixed, the result is a function of time $S(\cdot, \omega^*) : [0,T] \rightarrow R$ which is called a path of the process.

Giving a formal definition of a stochastic process is fundamental because, as we will see in later sections, in particular in the Black-Scholes option-pricing model, the log-returns of the stock price S are modeled as a Brownian motion, which is one of the most famous stochastic processes. The Brownian motion paths are continuous functions ω on [0, T], and therefore Ω in this case is the space of continuous paths on [0, T].

5) DISTRIBUTION OF A PROCESS

The cumulative distribution function of the process S(t) can be written as:

$$P(S_t \leq x)$$

We also need to define a joint distribution for those cases in which we want to monitor a process at two or more stages (dates t1, t2 etc), for example path dependent options whose payoff depends on how the price of the underlying asset at maturity was reached. The joint distribution can be computed as follows:

$$P(S(t_1) \le x_1, S(t_2) \le x_2)$$

as the sets $\{S(t_1) \le x_1\} \cap \{S(t_2) \le x_2\}$, for any x_i belonging to $F_{t2} \subseteq F_T = F$.

6) EXPECTATION

The expected value or mean of a random variable X can be defined as the average of its realizations X(ω), weighted by the probability of happening. If X is discrete:

E[X] = sum over i $x_i P(X = x_i)$

If X is a continuous random variable with density p_X

$P(x < X \le x + dx) = p_X(x)dx$ and

E[X] = integral of $xp_X(x)dx$

Expectation is a linear operation thus the expectation of a linear combination of random variables is the linear combination of the expectations:

E[aX + bY] = aE[X] + bE[Y]

An important consequence is that we do not need to joint distribution of two random variables X,Y if we want to compute the expectation of a linear combination of them. However, the knowledge of the joint distribution is essential if we want to compute quantities which involve both random variables, for example E[XY].

If X is a continuous random variable and Y is a function of X, say

$$Y = g(X)$$

as far as pricing an option is concerned, it is important to be able to compute the expectation of Y given the distribution of X. For example, if we want to fairly price a call option with payoff function g(x) = max(X-K,0) where X is the price of the stock at time T, the price Y = g(X) can be computed and expressed as an expectation.

If X has a density, it is not necessarily true that Y has a density too. However, If g is invertible and differentiable, with $g' \neq 0$, then Y has a density given by the formula:

$$p_Y(y) = p_X(g^{-1}(y)) * 1/|g'(g^{-1}(y))|$$

If Y has a density, its expectation, when g is regular and invertible can be written as:

$$\text{integral of } y\, p_X(g^{-1}(y)) * \frac{1}{|g'(g^{-1}(y))|}\, dy$$

Substituting $x = g^{-1}(y)$ leads to the following formula for the expectation:

$$E[Y] = E[g(X)] = \text{integral of } g(x)\, p_X(x)dx$$

This formula is very important because it is always valid, even in those cases in which Y doesn't have a density.

7) INDEPENDENCE

Two random variables X, Y are said to be independent if taken any couple of intervals I_1, I_2 the probability of the intersection $X \in I_1$, $Y \in I_2$ is equal to the the product of the probabilities.

Formally,

$$P(X \in I_1, Y \in I_2) = P(X \in I_1) * P(Y \in I_2)$$

If the two random variables X,Y have a joint density (X,Y), independence can be expressed using the following two conditions:

1. the marginals X and Y have probability densities p_X , p_Y

2. the joint density is the product of the marginal densities:

If X, Y are independent, then:

$$p(x, y) = p_X(x) * p_Y(y)$$

Another important implication is that, if two random variables are independent, the expectation of their product can be expressed as the product of the expectations:

$$E[XY] = E[X] * E[Y]$$

and therefore they are uncorrelated:

$$E[(X-E[X])*(Y-E[Y])] = 0$$

Independence implies uncorrelation but the opposite is not true except in the Gaussian context where the two notions of independence and uncorrelation are equivalent.

8) CONDITIONAL EXPECTATION

The conditional expectation of a random variable can be defined as the average of all outcomes (weighted by the probability of happening), given some extra knowledge/information that we have.

The result of this operation is not a number but a random variable whose value is known at the time the extra information is revealed. It is possible to identify four properties of the conditional expectation:

- $E[E[Y|X]] = E[Y]$

- additivity: $E[Y_1 + Y_2 | X] = E[Y_1 | X] + E[Y_2 | X]$

- $E[f(X)Y | X] = f(X)E[Y | X]$, meaning that random variables which are known when X is known can be treated like constants and taken out of the expectation.

- If two variables (X,Y) are independent then the conditional expectation of one with respect to the other is constant, and coincides with the mean:
  $E[Y |X]=E[Y]$ or $E[X|Y]=E[X]$

All the above properties can and should be expressed in terms of filtered space since, as we already said, in the world of Finance, it is quite a common practice to fix a finite time horizon T and look at the information available at any time t where t <= T. As we know, a filtered space is identified by the triplet $(\Omega, (F_t)_{t \leq T}, P)$. If we consider two dates t1, t2 with t1<t2, the conditional expectation of Y given the information available at time t1 is nothing more than a random variable which represents the best prediction of Y considering the information that we possess at t1. We can express the conditional expectation properties in terms of filtered space as follows:

- $E[E[Y |F_{t1}]]=E[Y]$

- if Y is known by time t_1, then $E[Y |F_{t1}]=Y$

- additivity: $E[Y +W |F_{t1}]=E[Y |F_{t1}]+E[W |F_{t1}]$

- If Z is known at time t1 it can be treated as a constant and taken out of the expectation:
  $E[ZY | F_{t1}] = ZE[Y | F_{t1}]$

- If Y is independent of F_{t1} , then $E[Y |F_{t1}] = E[Y]$.

Last but not least, we should mention the "Tower Law" also known as law of iterated expectations or smoothing theorem. The "Tower Law" tells us that if we want to predict Y at time t0 we can first compute the best prediction for Y at time t1 ($E[Y|F_{t1}]$) and then compute the best prediction at time t0 (this is why the law is called law of iterated expectations).

The general assumption that we make is that F0 (the set of information available at time t0) is nothing more than the set of all possible outcomes (Ω). At time t0, we know nothing about future realizations of the random variable hence all the variables known at this time are treated as constants.

9) MARTINGALE PROCESS

An adapted process M is said to be a martingale if:

$$E[M(t) \mid F_s] = M(s)$$

for all $0 \le s < t \le T$.

A martingale represents the formalization of a fair game in which the current entry price M(s) is equal to the conditional expectation of the future payoff M(t) given the information available at time s (we assume that $s < t$).

One of the most famous martingale processes is the Brownian Motion.

10) BROWNIAN MOTION

A process $W = (W(t))_{t \le T}$ is said to be a (standard) Brownian Motion on a filtered space identified by the triplet $(\Omega, (F_t)_{t \in [0,T]}, P)$ if the following conditions hold:

1. $W(0)=0$

2. W is adapted to the filtration

3. for any $s < t$, the increment $W(t) - W(s)$ is independent of F_s, and has distribution $N(0,t-s)$

4. the paths $W(\cdot, \omega)$ are continuous

It follows that:

- Marginal distributions are Gaussian. For any t, $W(t)$ can be written as $W(t) - W(0)$ since $w(0) = 0$ and it has distribution $N(0,t)$.

- For any $u \leq s < t$ the variables $W(u), W(t)-W(s)$ are independent, and therefore jointly Gaussian. This reasoning can be extended to any number of increments.

Proof: A Brownian Motion is a Martingale process.

We fix two arbitrary dates $s < t$ and write $W(t)$ as $W(t) - W(s) + W(s)$. It follows that:

$E[W(t) \mid F_s]$ (we rewrite $W(t)$ as $W(t) + W(s) - W(s)$) =

$E[W(t) - W(s) + W(s) \mid F_s]$ ($E[W(s)|Fs] = W(s)$ and the increment $W(t)-W(s)$ is independent of Fs)

=

$W(s) + E[W(t) - W(s)]$ $W(t)-W(s)$ is $N(0, t-s)$ =

$W(s)$

It is possible to apply different transformations to the standard Brownian motion.

10.2) LINEAR BROWNIAN MOTION

It's a linear transformation of the standard Brownian motion. A Brownian motion with drift b and volatility $\sigma > 0$ is the process

$$B(t) = bt + \sigma W(t)$$

10.3) GEOMETRIC BROWNIAN MOTION

We consider a Brownian motion with drift b and volatility σ. Recalling that $B(t) = bt + \sigma W(t)$ it is possible to apply an exponential transformation to this linear Brownian motion therefore coming up with a process Y such that:

$$Y(t) = \exp(B(t)) = \exp(bt + \sigma W(t))$$

This process is known as geometric Brownian Motion. Marginal distributions are lognormal, meaning that the logarithm of marginal distributions is normally distributed.

11) MARKOV PROCESS

A Markov process S is an adapted process such that, for every deterministic function $g = g(x)$, and for any arbitrary dates $s < t$, the conditional expectation of $g(S(t))$ satisfies

$E[g(S(t)) \mid F_s] = E[g(S(t)) \mid S(s)] = h(S(s))$ with h being an appropriate function that depends on g and the distribution of the process

The formula can be explained as follows: to predict the future value of S at time t (namely $g(S(t))$) we only need the information available at s with $s < t$. The information available in the interval $[0,s)$ is not necessary for the purpose.

12) ITO'S LEMMA

In stochastic calculus, we need to extend the traditional rules that we apply to compute derivatives and integrals when dealing with non-stochastic calculus. Suppose that we have a function of x $u(x)$ and that we want to express the differential du in terms of the differential dx. We consider the Taylor expansion of $u(x)$ about some value x bar.

u(x) = u(x bar) + u'(x bar) (x- x bar) + 1/2 u''(x bar) (x- x bar)^2 + …

x - x bar can be expressed as Δx.

In terms of differential we have:

du = u'(x bar) dx + 1/2 u''(x bar) dx^2 + …

In standard non-stochastic calculus, we compute a differential simply by keeping the first-order terms. For small changes in the variable, second-order and higher terms are negligible compared to the first-order terms and the change in u is proportional to the change in x. In stochastic calculus, we must also keep second-order terms. This leads to Ito's formula:

du = u' dx + 1/2 u'' dx^2

If we use z to denote the Brownian motion and t to denote time, we can compute Ito's formula using the following rules:

d(z)^2 = dt

dzdt = 0

d(t)^2 = 0

In general, we consider a function F(t, x). The Markov process defined by F(t,W(t)) has dynamics given by the stochastic differential equation:

$$dF(t,W(t))= (F_t(t,W(t))+1/2F_{xx}(t,W(t)))dt+F_x(t,W(t))dW(t)$$

where thee green term above is known as Ito's term.

12.2) ITO'S PROCESS

An Ito process, also called diffusion, is any adapted process Y whose dynamics may be described by the function:

$$dY (t) = \alpha(t)dt + \beta(t)dW (t)$$

where α and β are adapted processes which are called the coefficients of the stochastic differential equation mentioned in the Ito's lemma. The first coefficient, α, is called the drift although in Finance the fraction a(t)/Y(t) is referred to as "the drift". The second coefficient, β, is the diffusion coefficient.

SECTION 3: OPTION PRICING MODELS

3.1) BLACK-SCHOLES MODEL

The Black-Scholes model is a model for the relative pricing of financial derivatives: derivatives (in our case of interest options) are priced relatively to the price dynamics of the underlying financial instruments. We have two basic financial instruments in the BS model: a risk-less bond B which constantly pays an interest rate $r >= 0$ and a risky stock S which must satisfy the stochastic differential equation and the following conditions known as Cauchy problem:

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

$$S(0) = S0$$

where S0 is the initial market price and μ and σ are constants known as drift and volatility respectively (we already found this concepts in the definition of a Brownian motion). The Cauchy problem has a unique solution that is:

$$S(t) = S0 * e^{(\mu - 1/2\sigma^2)t + \sigma W(t)}$$

and the marginal S(t) satisfies

$$\ln S(t)/S0 \sim N((\mu - 1/2\sigma^2)t, \sigma^2 t)$$

This condition can be formalized as follows: the stock log-return follows a normal distribution and the mean and the variance of the stock log-return grow linearly with time. As we already mentioned, the BS model, besides satisfying the SDE, must also satisfy the Cauchy problem. But what do the parameters of drift and volatility represent in this model? In Finance, it is a common practice to use the standard deviation as measure of volatility (standard deviation of the stock returns for example). Hence, the volatility σ is nothing more than the standard deviation of the annual stock leg-return. On the other hand, the meaning of the parameter μ is not so straightforward. We can deduct it from the following known expression for the expected value of the stock price as a function of t:

$$E[S(t)] = S_0 e^{\mu t}$$

It follows that the drift is the exponential growth of the average stock price. As we already said in the “Introduction” part of the Thesis, we want to come up with an option’s price that is fair. In Finance, we say that a pricing is fair if and only if it does not introduce any arbitrage possibility. An arbitrage can be defined as a risk-less profit opportunity that the investor can take advantage of when a miss-pricing of financial instruments has occurred. In the BS model that we just described, we may think of two possible arbitrage opportunities:

- if $\mu >= r$, the exponential growth of the average stock-price is greater or equal than the constant interest rate r paid by the bond. An arbitrage may be built by short-selling the bond and purchasing the stock.

- if $\mu < r$, the constant interest rate paid by the risk-less bond is greater than the growth of the average stock price. We may come up with an arbitrage opportunity by short-selling the stock and purchasing the bond.

However, we can safely conclude that the BS model is arbitrage free since neither of these two strategies leads to an arbitrage. The reason is the following: while the bond B increases deterministically at rate r, the stock increases at rate μ on average. It follows that the marginal S(t) (which is lognormal) can be below S0*e^rt even if μ > r and it can be above S0*e^rt even if μ < r. The concepts of arbitrage and self-financing strategy are fundamental and thus deserve a thorough explanation. We start with the notion of self-financing strategy and then we move to arbitrage since the latter is nothing more than an example of the former broader class.

A portfolio strategy in general is a couple of adapted processes (H,K) where H(t) identifies the number of stocks held at time t and B(t) is the number of bonds held. The portfolio value is given by:

$$V(t) = H(t)S(t) + B(t)K(t)$$

The simplest example of portfolio strategy is represented by buy-and-hold. This strategy is represented by the couple (K0,H0), meaning that the number of stocks and bonds held by the investor are fixed at time 0 and do not change till maturity. However, the BS model that we just described is a continuous-time option pricing model in the sense that the trading occurs along a finite time grid and not only at the start date 0 and the final date T. The notion of self-financing strategy is strictly related with the concept of wealth. To implement any strategy we need an initial wealth w0:

$$w0 = V(0) = K0 + H0S0$$

and for the following steps t1, t2:

$$w_1 = \lim t\text{->}t1\text{-}\ V(t) = K_0B(t_1) + H_0S(t_1)$$

$$w_2 = V(t_1) = K(t_1)B(t_1) + H(t_1)S(t_1)$$

If w1 = w2 at the date t1 and all the following trading dates, we say that the portfolio is self-financing, we do not need to inject or consume any extra money. At any trading date tn we just re-shuffle and re-allocate money between stocks and bonds.

More formally, the portfolio value V must satisfy a global stochastic differential equation (we already found this concept when we mentioned that the stock S must satisfy an equation of this kind) on the continuous time interval [0,T]:

$$dV(t) = K(t)dB(t) + H(t)dS(t)$$

where H(t) and K(t) are the the number of stocks and bonds held at time t. Furthermore, any self-financing portfolio is a diffusion. We recall that an Ito process, also called diffusion, is any adapted process Y whose dynamics may be described by the function:

$$dY(t) = \alpha(t)dt + \beta(t)dW(t)$$

where α and β are coefficients of the stochastic differential equation. The first coefficient, α, is called the drift and the second coefficient, β, is the diffusion coefficient. In our case Y(t) = V(t), α = K(t) and β = H(t).

Introducing the concept of self-financing strategy/portfolio is extremely important because it makes the notion of arbitrage easier to grasp. An arbitrage is nothing more than a self-financing strategy V(0), H such that:

- $V(0)=0$
- $V(T)\geq 0$ and $P(V(T)>0)>0$

This kind of strategy starts with zero money and leads to a non-negative wealth, positive with positive probability. There is the possibility to gain with positive probability at no cost and without any risk. If such an opportunity existed, all investors would try to take advantage of it and prices would not be at equilibrium because of the demand-supply shifts. Prices would continue to fluctuate until any risk-less profit opportunity disappears. This the reasoning behind NA (short for non-arbitrage) models. One of the most famous examples of such kind of models is the Cox-Ross-Rubinstein model, also known as binomial option pricing model.

BINOMIAL OPTION PRICING MODEL

In the binomial option pricing model, the stock price follows a simple, stationary binomial process. At each moment in time, the price can either go up or down by a given percentage (we have an "up" factor u and a "down" factor d). When the stock price follows such a process and when there exists a risk-free asset (bond that constantly pays an interest rate r), options written on the stock can be easily priced. Furthermore, under certain pre-specified conditions, the binomial process converges to a lognormal process and thus the binomial formula converges to the Black-Scholes formula. The model is arbitrage free iff:

$$\mathbf{d < 1+r < u}$$

NO-ARBITRAGE FROM A MATEMATICAL VIEWPOINT

The NA concept shall be formalized using mathematical notation. Given a finite time filtered probability space identified by the triplet $(\Omega, (F_{t_i})_{i=0...n}, P)$ and a market model (B,S), where S is the risky asset and B is the risk-less bond, NA is equivalent to the existence of a probability Q on F_T with the following two properties:

1. Q(A)=0 if and only if P(A)=0 [Q is equivalent to P, notation Q~P]

2. the discounted stock price process $\hat{S} = S/B$ is a Q-martingale,

$$E^{Q}[S(t)/B(t)|Fs] = S(s)/B(s)$$

Such Q, which is not necessarily unique, is called an equivalent martingale measure (EMM) for $\hat{S}$.

The second condition can be restated as follows: the current price for S is equal to the Q-expectation of discounted future values:

$$S(s) = E^{Q}[B(s)*S(t)/B(t) \mid F_s] = E^{Q}[e^{-r(t-s)} S(t) \mid F_s]$$

when s = 0:

$$S(0) = E^{Q}[e^{-rt} S(t)]$$

In Finance, the existence and use of martingale methods is bound to the fundamental theorem of asset pricing. According to this theorem, the NA condition implies the existence of at least one equivalent martingale measure Q. If we want to launch a derivative (for example a European-style option that cannot be exercised at intermediate dates) written on the stock S, with payoff Ψ and maturity T, we should place the derivative on the market with a fair price P . Fair price means that the extended market (B, (S, P)) must also verify NA, in the sense that there exists no arbitrage strategy (K, H, J) based on the stock-bond-derivative triplet and the latter has price P. Obviously, at the maturity T P(T) = Ψ. The solution is to set:

$$\hat{P}(t) := P(t)/B(t) = E^{Q}[\Psi/B(T)|Ft]$$

If the price P verifies the above condition, the probability Q, which is an equivalent martingale measure for $\hat{S}$, is an EMM also for the discounted $\hat{P}$ thanks to its definition above and to the terminal condition P(T) = Ψ. This implies that Q is an equivalent martingale measure for the bi-dimensional process (S,T). As a consequence, the fundamental theorem of asset pricing can be reversed and also stated as follows: the existence of at least one equivalent martingale measure implies no arbitrage. A fair price is:

$$P(0) = E^{Q}[e^{-rT}\Psi]$$

This expression tells us that prices must be computed as Q-expectations of discounted payoffs.

- Q VS P

Given the reasoning that we just made, we may be tempted to think that Q and P are two sides of the same coin. This is completely wrong as Q may be very different from P. We should think of Q just as a pricing probability tool. Thus, if we want to answer questions like:

- What is the probability that a call option with maturity T and strike price K has a positive payoff?
- What utility should were expect from wn (terminal wealth) of a given strategy?
- How can we choose between different portfolios of securities and how can we optimize the composition of the chosen portfolio?

All this questions must be addressed under P and not under Q by:

- computing P(S(t)>K)
- computing E[U(V(T))] where U is a utility function and E is a **P-expectation**
- choosing the portfolio according to a performance criterion such as Markowitz's mean-variance criterion (choose portfolios that lie on the efficient frontier, the set of portfolios which offer the highest expected return given a level of risk expressed as standard deviation of returns)

- **ARBITRAGE IN GENERAL MARKETS**

The way in which we just presented the concept of arbitrage is valid for markets open on a finite number of dates (such as the binomial option pricing model). For general markets (open on discrete but infinite dates or continuous-time models such as Black-Scholes) we must add one condition. In general markets, an arbitrage strategy is a self-financing strategy such that:

• $V(0)=0$

• there exists a constant $c>0$ such that $V(t)\geq -c$ for all $t\in(0,T)$. C here is a finite credit line which must be respected during the trading, and it may depend on H.

• $V(T)\geq 0$ and $P(V(T)>0)>0$

Not only arbitrage strategies, but also certain limits of self-financing strategies, which are 'asymptotic arbitrages' and are called Free Lunches with Vanishing Risk (FLVR) shall be excluded to get a consistent fair-pricing rule. The fundamental theorem of asset pricing can be restated as follows:

Existence of a Q equivalent martingale measures⇔ no Free Lunches with Vanishing Risk (NFLVR) ⇒ NA

The pricing tool based on Q equivalent martingale measure is very powerful. Suppose that we want to price a European-style option with payoff Ψ, maturity T. If:

$$P(t) = E^{Q}[e^{-r(T-t)}\Psi \mid F_t]$$

the extended model identified by the triplet (B,S,P) verifies NFLVR and also the weaker condition of NA. As we were saying, this pricing approach is very powerful since it applies to all market models and to all kinds of options, even those with no early exercise feature including path-dependent ones (options whose payoff is affected by how the price of the underlying asset at maturity was reached, the price path of the underlying asset) like Asian options whose payoff depends on the average price of the underlying asset over a pre-specified period of time during the lifetime of the option.

This discussion suggests us that there is a link between pricing and hedging. If we have a derivative Ψ with maturity T that can be replicated/perfectly hedged, there exists a self financing portfolio V (composed of stocks and bonds) such that.

$$V(T) = \Psi$$

In words, the payoff of the replicating self-financing portfolio must equal the payoff of the derivative. It follows that there exists a unique way of pricing the derivative without introducing arbitrage in the market represented by the triplet (stock, bond, derivative) that is:

$$P(t) = V(t)$$

All replicable derivatives can be reproduced using portfolios of stocks and bonds. The number of stocks used for the replication argument is usually referred to as the "delta" of the derivative. Of course, the approach based on martingale methods and the replicating/hedging method must yield the same result. This leads us to the definition of complete market:

A market is complete if and only if all the derivatives are replicable.

The market is complete in the sense that any derivative has a unique fair price which corresponds to the value of the replicating portfolio. The notion of "completeness" is important since the Cox-Ross-Rubinstein model and the Black-scholes model are both complete market models. These models are based on a number of assumptions. In reality, markets are incomplete in the sense that it is impossible to replicate derivatives accounting for all the risk involved in the process.

The notion of "completeness" leads us to the definition of the second fundamental theorem of arbitrage pricing:

If the market model verifies NFLVR then the following two conditions are equivalent:

- the market is complete
- there exists a unique equivalent martingale measure for S^

APPLICATION OF MARTINGALE METHOD TO THE BLACK-SCHOLES MODEL

As we said, the Black-Scholes model is a complete market model: any derivative has a unique fair price which equals the value of the replicating portfolio. This amounts to say that there exists a unique Q equivalent martingale measure for S^, according to the second fundamental theorem of arbitrage pricing. We want to prove mathematically the uniqueness of Q for S^, the discounted stock price. We can do it in the following way:

Since

$$dS = \mu S dt + \sigma S dW,$$

then S^ = S/B has dynamics given by

$$dS\text{^} = (\mu - r)\hat{S} dt + \sigma \hat{S} dW$$

We now introduce the new concept of market price of risk lambda.

$$\text{lambda} = \mu - r / \sigma$$

The market price of risk expresses the excess return with respect to the risk-free rate of return on the bond r (μ - r) per unit of volatility. It is reminiscent of the Sharpe Ratio (the ratio of the stock excess return with respect to the risk-free borrowing and lending rate per unit of standard deviation of the stock return). We set:

$$W^{*}(t) = W(t) + \lambda t$$

and

$$dW^{*} = dW + \lambda dt$$

The stochastic differential equation for S^ can be read as

$$dS\text{^} = \sigma \hat{S} dW^{*}$$

If there exists a pricing probability Q ~ P such that W* were a Brownian-motion, we can conclude that Q is a martingale probability for the discounted stock price S^. The Girsanov Theorem ensures the uniqueness of Q and provides a link between Q and P. According to the second fundamental theorem of arbitrage pricing, we know that the market model is complete and that any derivative with no early-exercise feature can be replicated.

The Girsanov Theorem provides a random variable which is the density of Q with respect to P:

$$dQ/dP = \exp(-\lambda W(T) - 1/2\lambda\text{^}2T) > 0$$

The strict positivity ensures the equivalence of Q and P. This density function allows us to compute the Q probabilities as "deformed" P probabilities.

$$Q(A) = E^{Q}[I_A] = E[dQ/dP I_A]$$

IA is multiplied by the "deformation" density function dQ/dP prior to taking the P-expectation. S^, the discounted stock price, becomes a martingale geometric Brownian-motion under Q since:

$$\hat{S}(t) = \exp(-\sigma^2 t/2 + \sigma W^*(t))$$

which is equal to saying that

$$dS = rSdt + \sigma S dW^*$$

The drift of the stock price S under Q is nothing more than the risk-less rate r. This means that under Q the stock is a geometric Brownian-motion with the same volatility but grows at the risk-free rate constantly paid by the bond. This is why Q is called a risk-free probability measure. Even if the investor was risk-averse, he would behave as if she was risk-neutral when pricing.

UNIQUE FAIR PRICING OF A DERIVATIVE UNDER THE BLACK-SCHOLES MODEL

If the derivative has maturity T and is of the Markovian form g(S(T)) then the price today is:

$$P(0) = e^{-rT} E^{Q}[g(S(T))] = e^{-rT} E^{Q}[g(S_0 \exp((r - \sigma^2/2)T + \sigma W^*(T)))]$$

The following integral has to be solved:

P(0)= e^-rT * integral of $g(S_0$exp((r−σ^2 /2)T +σ T^1/2x))e^-x^2/(2pi)^1/2dx

The integral may be impossible to solve analytically. However, it is possible to resort to simulation. If

$$P(t) = E^{Q}[e^{-r(T-t)} g(S(T)) \mid F_t]$$

The idea is to write

$$S(T)=S(t)\exp((r-\sigma^2/2)(T-t)+\sigma(W^*(T)-W^*(t))),$$

In this way, the dependence of the Markovian function from the past is contained in S(t) whereas the lengthy term at the exponent is independent of the information set Ft. Solving the integral leads to a resulting price process of the following form:

$$G(t, S(t))$$

that is a Markovian function of S.

FINAL RESULTS FOR THE BLACK-SCHOLES MODEL

The final formula for the price C(0) of a call with maturity T and strike price K is:

$$C(0) = S_0\Phi(d_1) - Ke^{-rT}\Phi(d_2)$$

where Φ is the standard Gaussian CDF.

$$d1 = (\ln S0/K + (r+1/2\sigma\hat{\ }2)T)/\sigma T\hat{\ }1/2$$

$$d_2=d_1-\sigma T\hat{\ }1/2$$

In general, at any intermediate date t < T:

$$C(t, S(t)) = S(t)\Phi(d_1(t)) - Ke^{-r\tau}\Phi(d_2(t))$$

where $\tau = T - t$ and $d_i(t)$, i = 1, 2 are

$$d1(t) = (\ln S(t)/K + (r + 1/2\sigma\hat{\ }2)\tau)/\sigma T\hat{\ }1/2$$

$$d_2(t) = d_1(t) - \sigma T\text{^}1/2$$

The price of the corresponding put option can be derived through direct calculations, as for the call option or more easily through the put-call parity which expresses the relationship between the price of a call-option and a put-option written on the same stock S and with the same maturity T and the same strike price K:

$$Ct - Pt = St * Ke\text{^}{-r(T-t)}$$

which is true for all $t < T$.

4) THE IMPORTANCE OF SIMULATION

If the payoff of the option is given by a very complex expression, for example in the path-dependent case, the only feasible way to price it is to resort to simulation.

Why should we use simulation?

First of all, we start by defining what a simulation is. It is the imitation of a real-world process or system. It is often a mathematical model of a process. Simulations, just like any other mathematical model, usually make assumptions about the behavior of the system being modeled. The model is constructed in such a way that inputs can be changed over a set of values and this allows a complete picture of the possible outcomes.

The main advantage of simulation is that it transfers work to a computer. Models with greater complexity can be handled, making far less assumptions. A more faithful representation of reality is also possible, going beyond the results that can be achieved through mere mathematical analysis.

Simulation can also be used in order to perform the so-called “stress testing”: determining what would happen under extreme circumstances. It is a very powerful tool since the scenarios that we are interested in are so rare that we do not have substantial experience of.

There are different steps involved in the construction of a simulation. Some of the most important are:

- Formulation of the problem

- Definition of the objectives
- Choose a model and write computer code for it
- Verify, validate the model and run the simulation
- Analyze the output and present the results in a qualitatively-informative way

MONTECARLO SIMULATION METHOD

The Montecarlo simulation method is a computer technique based on performing numerous experiments with random numbers. It can be applied in many fields of research and it doesn't need a special knowledge of probability theory. The only information needed is the relationship between the output and input quantities:

$$y = f(x) \text{ or } y = f(x1,x2,x3,\ldots)$$

and the knowledge of probability distributions of the input variables. The method repeats trials with computer-generated random numbers processed by the relevant mathematical operations. The input variables $x_1, x_2, \ldots, x_n$ are assigned random values each time but their distributions must correspond to the probability distribution of each variable. The output quantity y is calculated as a function of these values. The distribution of y may be represented using graphical tools such as histograms.

The generated values can be used for example to determine the average value or the probability that y will be lower or higher than a chosen value y*.

AN APPLICATION OF THE MONTECARLO METHOD: PRICING AN ASIAN OPTION

An asian option has value determined not by the closing price of the underlying asset but on the average price of the asset over a time interval. For example, an Asian call option on an asset with price process S(t) pays an amount equal to max(0,Sk bar − K) at maturity where Sk bar = 1/k sum from i = 1 to k of S(iT/k) is the average asset price over k equally spaced time points over the interval (0,T). If the price process S(t) follows a geometric Brownian motion then Sk bar is the sum of log-normally distributed random variables and the distribution of such sum is very difficult to express analytically. This is why we resort to pricing the Asian option using simulation. The geometric average has a distribution which can be easily obtained. The geometric mean of n values $X_1,...,X_n$ is $(X_1X_2...X_n)^{1/n}$ =exp{1/n sum from i = 1 to n of ln(Xi)} and if the random variables are log-normally distributed this amounts to adding the normally distributed random variables ln(Xi) in the exponent. The sum in the exponent is normally distributed hence the geometric average will have a lognormal distribution. On the other hand, pricing an arithmetic-average Asian option is much more difficult than the geometric counterpart and a closed-form formula has not been derived yet.

In order to price Asian options, we need to agree on a specific risk-neutral model which will be the Black-Scholes one that we already discussed. The Montecarlo method is a robust, yet computationally demanding approach to pricing arithmetic-average Asian options. It consists of:

- using random number generators to generate outcomes for a random variable with known distribution.
- applying the LLN, estimating the unknown expected value with the average of the generated outcomes.
- increase the sample size in order to determine a confidence interval for the estimate.

Montecarlo simulation is very flexible, allowing for complex stock movements and path-dependency which are otherwise extremely hard to handle by analytical methods. However, there is also a bad part of the story. Simulation methods are known for their slow speed of convergence. Relying on the LLN and estimating the unknown expected value by the average of generated outcomes implies the standard deviation of the estimate to be proportional to n^-0.5. This means that for each additional decimal digit precision, we need to increase the sample size by 100. As a consequence, the Montecarlo simulation method is infeasible beyond the 4-6 decimal digits of precision. To reduce the magnitude of this issue, we need to resort to variance reduction techniques to reduce the variance of the estimate. Some well known techniques are antithetic variates, control variates and quasi Montecarlo simulation. I will focus on the antithetic variates technique.

ANTITHETIC VARIATES TECHNIQUE

The basic Montecarlo method consists of taking the arithmetic-mean as unbiased estimator of the unknown expected value. The antithetic variates method allows us to try and reduce the variance of the estimator. The idea is using each generated random number twice, creating two trajectories (outcomes) which are negatively correlated (antithetic). It is usually a good idea to find a step in the simulation when we deal with symmetrically distributed random variates and then mirror them about their expected value, For example, in the case of log-return sequences, we can take the negative of the log-returns in order to generated a "mirrored" trajectory. As long as the original and antithetic payouts have a negative covariance, the variance of the estimator will be reduced with respect to the basic method and the estimator remains unbiased. Furthermore, since we use each random number twice, we are only generating n/2 outcomes instead of n, making the antithetic variates method twice as efficient as the basic method.

OPTION PRICING WITH PYTHON: CLOSED-FORM FORMULAS VS SIMULATION

In the following Python script, we construct a simulation to derive an approximation for the price of a call and put option. We then compare it with the results coming from the application of Black-Scholes closed formula. The code contains comments to explain how we proceed step by step.

```
# We set values for the following parameters
S = 100 # spot price
T = 1 # expiration date of the option, 1 year from now
r = 0.05 # discount rate to price the option
sigma = 0.15 # volatility (standard deviation) of the stock/general underlying asset
K = 100 # exercise price
#We now set a number of iterations for the Montecarlo simulation. The higher the
number of iterations, the more precise the result but the slower the process.
# We set a number of simulations equal to 2000.
n = 2000
n_trad = 250 # This is the approximate number of trading days in one year

dt = T/n_trad # We consider a brownian motion for daily prices

# General brownian motion formula that we apply: S_t+1 = S_t * e^((r-(sigma^2)/2)dt +
sigma*dt^1/2*r) where (r-(sigma^2)/2)dt is the drift and sigma*dt^1/2 is the time component.
```

We call the latter a. The r at the exponent is a standard normal variable having mean 0 and standard deviation 1

```
import numpy as np
drift = (r-(sigma**2)/2)* dt
a = sigma* np.sqrt(dt)
x = np.random.normal(0,1, (n, n_trad))
S_matr = np.zeros((n, n_trad)) # we start from a zero stock price matrix. This is a 5000*250
matrix filled with zeros.
S_matr [:,0] += S # at time t = 0 the stock price is 100. On the x axis of the matrix
we have the number of simulations, on the y-axis we have the number of steps represented by
the number of trading days
for i in range (1, n_trad): # We use a loop to repeatedly apply the brownian motion formula.
    S_matr [:,i] = S_matr[:,i-1] * np.exp(drift + a * x[:, i])
C = S_matr[:, -1] - K # payoff of a call option at the final date
for i in range(len(C)):
    if C[i] < 0:
        C[i] = 0
    else:
        C[i] = C[i]
# We remove all negative values since the payoff of a call option cannot be negative
# We repeat the same process for a put option
P = K-S_matr[:,-1]
for i in range(len(P)):
    if P[i] < 0:
        P[i] = 0
    else:
        P[i] = P[i]
# we compute the payoff of the options as an average of the values derived from the loops
call_payoff = np.mean(C)
put_payoff = np.mean(P)
# We apply continuous discounting to derive the price of the options today.
call = call_payoff * np.exp(-r * T)
put = put_payoff * np.exp(-r * T)
```

```
# We make a comparison with the closed-form expression from Black-Scholes model.
import scipy.stats as st
# These are the formulas we already explained in detail in the previous section
def d1(s, K, r, vol, t):
    return (np.log(s/K) + (r+(vol**2)/2)*t)/(vol*np.sqrt(t))
def d2(s, K, r, vol, t):
    return d1(s, K, r, vol, t) - vol*np.sqrt(t)
def callval (s, K, r, vol, t):
    return s * st.norm.cdf(d1(s, K, r, vol, t)) - K * st.norm.cdf (d2(s, K, r, vol, t))*np.exp(-r*t)
def putval (s, K, r, vol, t):
    return -s * st.norm.cdf (-d1(s, K, r, vol, t)) + K * st.norm.cdf(-d2(s, K, r, vol, t)) * np.exp(-r*T)
# We compute the prices using the same values of the simulations
call2 = callval(100,100,0.05, 0.15, 1)
put2 = putval(100,100,0.05,0.15,1)
```

FINAL RESULTS

The price of the call with the simulation method is 8.44 whereas it is 8.59 using BS formula. The price of the put with the first method is 3.76 and 3.71 with the second.
The Montecarlo simulation method provides a fairly accurate approximation although its execution time depends on the number of iterations. With a very large number of simulations, the method becomes too slow to be competitive.

CONCLUSIONS

Determining the price of an option is a very difficult task since there are many parameters involved that affect the price: the most important are the price of the underlying asset, time and volatility. Any change in one of these three variables will alter the option's value. Starting with the introduction of the relevant mathematical concepts, we then moved to the analysis of three of the most famous models for option pricing: Black-Scholes, binomial option pricing and Montecarlo simulation. All these models start from some relevant variables (price of the stock, strike price, volatility, interest rate, time to maturity) in order to theoretically value an option. They all provide an estimation of the option's fair value. The first two models are based on a series of assumptions. Although they both consider a portfolio consisting of two financial instruments (bonds and stocks), the Black-Scholes model assumes that the risk free rate that the bond constantly pays and the volatility of the underlying asset are constant, the returns of the underlying asset follow a log-normal distribution and that the option can only be exercised at expiration. On the other hand, the fundamental assumption of the binomial model is that at each moment in time the price of the stock can either go up (according to an up factor u) or down (according to a down factor d). These assumptions allow us to come up with a closed-form expression that can be immediately computed given the value of the other parameters. When the payoff of the option is more complicated, simulation is the way to go. In the last section, we discussed a practical application of the Montecarlo simulation method to Asian options, path dependent options which are more difficult to value since their value depends on the average price of the underlying asset over a period of time. Finally, we showed that the simulation approach is very powerful especially for large number of trials but this comes at the expense of a slow execution speed. Whenever the simplifying assumptions hold and an analytical formula is available, it should be preferred since simulation may become too slow to be competitive.

REFERENCES

[1] Stanley R. Pliska, *History of options from the middle ages to Harrison and Kreps*

[2] Sara Biagini, *Lecture Note of Mathematical Methods for Financial Markets*, Pisa University, 2014

[3] Simon Benninga, Zvi Wiener, *The Binomial option pricing model*, 1997

[4] Jaroslav Menčík, *Montecarlo simulation method*, University of Pardubice

[5] Don L. McLeish, *Montecarlo simulation and Finance chapters 1-6*, 2004

[6] Akos Horvath and Peter Medvegyev, *Pricing Asian Options: A Comparison of Numerical and Simulation Approaches Twenty Years Later,* Journal of Mathematical finance, 2016

[7] Karatzas, Steven E. Shreve, Brownian motion and stochastic calculus, Springer 2000

[8] Hull, Options, futures and other derivatives

[9] Steven E. Shreve, Calculus for finance 1 & 2, Springer 2005