LUISS

Department of Economics and Finance

Chair of Mathematical Finance

Portfolio Optimization

SUPERVISOR

Prof. Sara Biagini

CANDIDATE

Luca Simone 193321

ACADEMIC YEAR

2016/2017

Portfolio Optimization

Luca Simone

May 13, 2017

Contents

1 Introduction

The scope of the thesis is portfolio optimization. A portfolio is a set of various assets, its optimization is the process by which an investor can select the most appropriate allocation of wealth, among the different assets, for the purpose of achieving a specific goal. The optimization may reflect different aspirations of the investor, but overall the main goals are maximization of returns and minimization of risk. The aim of this discussion is to propose the right model and strategy for an investor. Topics treated will include a an introduction to stochastic processes, concepts of simulations and a virtual test of our conclusion using the Monte Carlo method.

The main topic will be the Merton model perspective. This model is named after its creator Robert C. Merton, Nobel Prize for Economics in 1997. What is the scope of this model? An investor is deciding how much of his wealth to allocate in either of two possible investment opportunities: a risk-free asset or a stock. His goal, provided of utility fuction and wealth process, is to maximize the final outcome in finite or infinite time horizons, consuming a certain amount of his wealth in each period.

Later on there will be an experiment of portfolio optimization. In this discussion we will describe Montecarlo simulations and, after preparing our data for analysis, we will use this type of procedure to reason over the real-life choices of an investor in the FTSE MIB index emulating a possibile outcome.

We will start from the basics in order to work our way up to the general picture, fully understanding what is going on. The work will be organized the following way:

1. A mathematical section in which we will consider the main tools used in finance to deal with portfolio optimization. Such tools include a probability background, stochastic processes and many properties that are of common use in these types of analysis

2. A description of the Merton problem, we will describe the issue and provide a solution to it, considering some of the many forms it can take.

3. Explanation of Montecarlo process and a brief description of its uses

4. We will make a simulation of the outcome of a portfolio created based on our criteria.

2 Probability and processes

Any variable that takes uncertain values in time is said to follow a stochastic process. Stochastic processes can be discrete-time, if the changes in the value of the variable can take place only at certain points in time, or continuous-time, if the changes are possible at all times. Moreover, stochastic processes can be discrete-variable, if the variable can take only specific discrete values, or continuous-variable, if the variable can take any value within a certain range. Stocks, which we will be discussing can be easily thought of following a continuous time and variable process. In reality, stocks can take discrete values (of multiples of a cent) and are not traded continuously (they are traded only when the markets are actually open). Nevertheless, reasoning in continuous terms is very important in achieving our result. We now start introducing the framework in which we will be immerged.

2.1 Probability Spaces

Stochastic processes take place in what is called a probability space. It has the form $(\Omega, \mathcal{F}, \mathbb{P})$ where:

- Ω is the set of all possible outcomes.
- $\mathcal{F}$, defined as sigma-algebra, represents the historical but not future ($\mathcal{F}$ is non-anticipative) information available on our stochastic process.
- $\mathbb{P}$ represents the normalized probability of outcomes.

Such characteristics of the probability space are strictly linked to the specific topic we are dealing with and, thus, vary with different processes meaning different $(\Omega, \mathcal{F}, \mathbb{P})$. The probability space is the big box in which all the action takes place. Consequently, it is important to know exactly how to move around in this environment.
Let's start by considering a real random variable X. X on $(\Omega, \mathcal{F})$ is a function on Ω which takes values in $\mathbb{R}$:

$$X : \Omega \to \mathbb{R}$$

and is $\mathcal{F}$-measurable. This means that the counter-image of any half line $(-\infty, x]$ is an event:

$$\{X \leqslant x\} \in \mathcal{F}$$

for all $x \in \mathbb{R}$. The information we possess up to a certain point x generates what is called a sigma-algebra, which we define as follows:

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

all the events that can be expressed in terms of X, for example $\{a \leqslant X \leqslant b\}$ belong to $\sigma(X)$.

2.2 Stochastic Processes

After this first glimpse at probability spaces, we can give a more rigorous definition of a stochastic process: it is a mapping of real valued random variables from $(\Omega, \mathcal{F})$ to $\mathbb{R}$.
In a filtered space $(\Omega, (\mathcal{F}_t)_{t\leqslant T}, \mathbb{P})$, if at time t we know the value of a real-valued stochastic process $S = (S(t))_t$, then $S(t)$ is $\mathcal{F}_t$-measurable, or, in other words, S is adapted to the filtration.
For S to be an adapted process the following conditions must hold:

- for any fixed time t,
$$S(t) : \Omega \longrightarrow \mathbb{R}$$
- for all fixed reals x, the set $\{S(t) \leqslant x\}$ belongs to $\mathcal{F}_t$

2.3 Expected Values

For a random variable X, the expected value, $E[X]$, is average of the possible outcomes of X, each weighted on the respective probabilities of actually happening.
For a discrete random variable X:

$$E[X] = \sum_i X_i \cdot p_i$$

if X is instead continuous, with density p_X then:

$$E[X] = \int x p_X(x) dx$$

We recall to the reader that expectation is a linear operation, meaning that the expectation of a linear combination is the linear combination of the expectations:

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

so an expectation over a linear combination of X and Y may be computed without knowing the joint distribution of the two variables. The same cannot be said, for example, for the computation of $E[XY]$. In this case a covariance matrix is required, or, at most, assumptions over the independence of the two variables.
Assume we have a continuous random variable, X, and function of such variable, Y. Then we can write:

$$Y = g(X)$$

In some cases it might be useful to compute the expectation of Y based on our knowledge of the distribution of X. There is no rule guaranteeing that Y necessairly has a density. Whether or not this is the case clearly depends on g. If we were to face a g that is of the

Bernoulli type, for instance, this wouldn't be the case. But, given a g that is invertible and differentiable, with $g' \neq 0$, we can state that Y has density, p_Y, with:

$$p_Y(y) = p_X(g^{-1}(y))\frac{1}{|g'(g^{-1}(y))|}$$

We take now an expected value of Y and by means of substitution we arrive to the following:

$$E[Y] = E[g(X)] = \int yp_X(g^{-1}(y))\frac{1}{|(g')g^{-1}(y)|}dy = \int g(x)p_X(x)dx$$

bearing in mind that by definition $x = g^{-1}(y)$. Note, moreover,that this last formula, being only dependent on x, is always valid and applicable to Y, even when it does not have a density.

2.4 Independence

As we mentioned this concept above while referring to joint distributions, we describe it a bit more in detail. Two random variables X and Y are independent if, taking two intervals I_1 and I_2, the probability of the intersection $X \in I_1$, $Y \in I_2$ factorizes to:

$$\mathbb{P}(X \in I_1, Y \in I_2) = \mathbb{P}(X \in I_1)\mathbb{P}(Y \in I_2)$$

which means that the joint density is the product of the marginal densities:

$$p_{(x,y)} = p_X(x)p_Y(y)$$

as we were mentioning above, in case of independence of our two random variables we can state:

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

moreover, we add that if two random variables are independent, then they are also uncorrelated. Attention must be paid to the fact that the opposite argument does not hold, as two variables that are uncorrelated are not necessairly independent. We can prove that, in the case of indpenence, they are uncorrelated by first computing their covariance which is:

$$E[(X - E[X])(Y - E[Y])$$

$$E[XY - YE[X] - XE[Y] + E[X]E[Y]]$$

$$E[XY] - E[Y][X] - E[Y]E[X] + E[X]E[Y] = 0$$

as, because of independence, $E[XY] = E[X]E[Y]$
Being the above covariance of the two variables equal to zero, also their correlation, defined as $\rho = \frac{\sigma_{x,y}}{\sigma_x\sigma_y}$ must be equal to zero.

2.5 Conditional Expectations

Conditional expectation of a random variable X is an expectation, therefore the weighted average of some possible outcomes, conditioned over some extra knowledge that is given. Conditional expectations are more precise in making a guess over a random variable than expected values as they allow us to take into account an extra bit of information that may be helpful for the task.
We list some of the main expected value properties:

1. $E[E[Y|X]] = E[Y]$
2. additivity: $E[Y_1 + Y_2]|X] = E[Y_1|X] + E[Y_2|X]$
3. random variables known when X is known can be considered as constants and taken out of the expectation, $E[f(X)Y|X] = f(X)E[Y|X]$
4. if two variables X and Y are independent we can state: $E[Y|X] = E[Y]$ and $E[X|Y] = E[X]$.

Let us now consider the case in which the conditioning happens over some information $\mathcal{F}$. Let's take the following situation as example: we deal with a random variable Y that will be known at some future date $t_2 > t_1$. What is the best guess we can make on Y at time t_1? The solution to this question is:

$$E[Y|\mathcal{F}_{t_1}]$$

It is actually our best guess because we are conditioning our future expectation on all the possible information available at t_1. This does not necessairly require us to be at that specific point of time t_1, as $\mathcal{F}_{t_1}$ doesn't represent current information but rather represents all the historical information we possess, or will possess, at that date. For this reason $E[Y|\mathcal{F}_{t_1}]$ is a random variable itself, which will be known precisely at time t_1.
It is useful now to consider the properties of conditional expectations, taking Y and W as random variables, known, at least, at some time T. Morever we fix a timeline such that $0 \leqslant t_0 < t_1 < t_2 \leqslant T$. Some of the properties with respect to information that are worth mentioning are:

1. $E[E[Y|\mathcal{F}_{t_1}]] = E[Y]$
2. if Y is known by time t_1, $E[Y|\mathcal{F}_{t_1}] = Y$
3. additivity: $E[Y + W|\mathcal{F}_{t_1}] = E[Y|\mathcal{F}_{t_1}] + E[W|\mathcal{F}_{t_1}]$
4. for any Z known at time t_1,$E[ZY|\mathcal{F}_{t_1}] = ZE[Y|\mathcal{F}_{t_1}]$
5. if Y is independent of $\mathcal{F}_{t_1}$, then $E[Y|\mathcal{F}_{t_1}] = E[Y]$ which is constant.

6. tower law: $E[Y|\mathcal{F}_{t_0}] = E[E[Y|\mathcal{F}_{t_1}]|\mathcal{F}_{t_0}]$; this means that our best prediction at time t_0 can be made directly or through an intermediate step, which is computing first the best prediction of Y for t_1 and then for t_0.

It is common practice in Finance to set $\mathcal{F}_0 = 0$, meaning that, at outset, we have no information whatsoever.

2.6 Martingales

It is a small step from the concept conditional expectation over information, $\mathcal{F}$, to the one of martingales. An adapted process M is defined as a martingale if:

$$E[M(t)|\mathcal{F}_s] = M(s)$$

for all $0 \leqslant s < t \leqslant T$.
To make it more clear we make use of an example. We claim that a martingale defines what we can call a "fair game". These games are defined as such because an entrant pays an entry price which is fair. If $S(t)$ models, at time $t \in [0, T]$, the entry price of a game with payoff X and $S(T) = X$ and S is a martingale, then the conditional expectation of the future payoff X at time t is exactly its current price:

$$E[S(T)|\mathcal{F}_t] = S(t)$$

2.6.1 Supermartingales and submartingales

These concepts are similar to a martingale and can be considered as an extension of this concept. Instead of being the exact value of the conditional expectation of a future value, the current value of the variable represents an upper bound (supermartingale) or a lower bound (submartingale) for such expectation. We can give a mathematical definition of these two concepts.
A discrete-time submartingale has the form:

$$E[X_{n+1}|X_1, ..., X_n] \geqslant X_n$$

in continuous terms:

$$E[X_t]|\{X_\tau : \tau \leqslant s\}] \geqslant X_s, \forall s \leqslant t$$

Similarly the concept of supermartingale can be expressed in discrete terms in the following way:

$$E[X_{n+1}|X_1, ..., X_n] \leqslant X_n$$

and in continuous terms:

$$E[X_t]|\{X_\tau : \tau \leqslant s\}] \leqslant X_s, \forall s \leqslant t$$

2.7 Markov Property

An adapted process, S, is a Markov process if for any t, δt we have:

$$E[S_{t+\Delta t}|\ \sigma(S_s : s \leqslant t)] \equiv f(S_t)$$

for an appropriate deterministic function f.
This definition explains how the expected value of our process S_t (which in our case will be the price of stock S) after an increase Δt is dependent only on the current time, t. In the definition above, past information is indicated with the expression σ. Our "sigma" represents the entire history of the process S up to the point of time s. We deduce that Markov processes are consistent with the weak form of market efficiency, which states that all past information, accessible to all investors, is, as a consequence, incorporated in prices of assets, for instance in the value of a stock. Therefore, traders cannot "beat the market" by using past information and technical analysis.

2.7.1 Markov Process vs. Martingales

Before moving on, it may be worth spending a few words on the meanings of these two concepts which at first sight may seem very much alike. What are the practical differences between a martingale and a Markov process? The key characteristic of the Markov property is that it is "memoryless", in the sense that how we reach a certain state of the world is irrelevant in predicting future states. This statement finds criticism especially from those believing in momentum. Taking the example of a stock, stating that the current price $S_t = 10$, all that matters for a Markovian process is this value, the way the stock reached it is irrelevant. A stock dropping in value from 100 to 10 in two days may be seen in the same way as a stock that has always fluctuated around the value of 10. Believers in momentum, based on such past data, would most likely bet on a further depreciation of the stock value. A Martingale on the other hand simply defines that the future expectation of a stochastic process, that is the mean of the future possible outcomes, is exactly the current value.

2.8 Standard Brownian Motion

A particular type of Markov process is the Wiener process, also called Brownian Motion. A Brownian motion is sometimes referred to as "Random Walk" process, because it is supposed to emulate the (random) path of a stock from an initial value S_0. It is a normally distributed Markov process. Some may be familiar with this concept as similar one is taught in a basic economics class, concerning portfolio choices, and has the name of IID variations. The IID is a typical assumption of the CAPM model and its acronym stands for "independently and identically distributed".

More formally, in a filtered space $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, taken t as a continuous time parameter, $W = (W(T))_{t\leqslant T}$ is a Brownian Motion if

- $W(0) = 0$
- W is adapted to the filtration
- for any $s < t$, the increment$W(t) - W(s)$ is independent of $\mathcal{F}_s$, and has distribution $N(0, t-s)$
- the paths $W(*, \omega)$ are continuous

consequences of these properties are:

- Marginal distributions are Gaussian, for any t we can write $W(t) - W(0)$ which is normally distributed with $N(0, t)$
- for any $u < s$ we can conclude that $W(u), W(t) - W(s)$ are independent and therefore have a joint normal distribution $N\left(\begin{pmatrix}0\\0\end{pmatrix}\begin{pmatrix}u & 0\\0 & t-s\end{pmatrix}\right)$

Extending this reasoning for an infinite number of increments, we come up with a general condition formalized in the following way:
Fixed $0 \leqslant t_1 < t_2 < ... < t_n \leqslant T$ we obtain n increments $W(t_1), W(t_2) - W(t_1), ..., W(t_n) - W(t_{n-1})$, independent and jointly Gaussian distributed, with:

$$N\left(\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}, \begin{pmatrix}t_1 & 0 & \cdots & \cdots & 0\\ 0 & t_2 - t_1 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & t_n - t_{n-1}\end{pmatrix}\right)$$

2.8.1 Wiener Process is a Martingale

We now prove that a Brownian Motion is a martingale. Fixed two dates $s < t$ we can write $W(t)$ as $W(t) - W(s) + W(s)$. We use this particular expedient to reach the conclusion that:

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

2.9 Transforms of a standard Brownian Motion

Brownian motions can be primarily characterized by two components, the drift μ, and volatility $\sigma > 0$.

2.9.1 Linear Brownian Motion

A linear transform of W, the standard Brownian motion, is:

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

this is also referred to as Brownian motion with drift.

2.9.2 Geometric Brownian Motion

We start from our linear Brownian motion B. The exponential transform, Y, of $B(t) = \mu t + \sigma W(t)$ is:

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

such process is called Geometric Brownian motion. Because of its form it is evident that it follows the lognormal property, that it is, the logarithm of the marginal distributions are normally distributed.

2.9.3 When is a Geometric Brownian Motion a martingale?

Let us prove that a Geometric Brownian motion, S, is a Markov process. Fixing once more $s < t$, we make use of the same expedient used for the proof of the standard Wiener process; so, $W(t) = W(t) - W(s) + W(s)$.

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

the expectation can be reduced to the exponential moment $E[e^{a \cdot \epsilon}]$ of a standard gaussian with $a = \sigma\sqrt{t-s}$. We obtain:

$$E[e^{a\epsilon}]$$

The calculation workout is as follows:

$$\int e^{ax} \cdot \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx$$

$$\frac{1}{\sqrt{2\pi}} \int e^{-\frac{1}{2}(x^2 - 2ax + a^2) + \frac{a^2}{2}} dx$$

$$e^{\frac{a^2}{2}} \int_{-\infty}^{+\infty} \frac{e^{-\frac{1}{2}(x-2)^2)}}{\sqrt{2\pi}} dx$$

$$= e^{\frac{a^2}{2}}$$

And since $exp(\mu t+\sigma W(S)+\frac{a^2}{2}$ is a random variable known at time s, the Geometric Brownian Motion is a Markov process We can also conclude conclude that:

$$\mu = -\frac{\sigma^2}{2}$$

so that at the exponent, which will be $\mu t + \sigma^2(t-s) + \sigma W(s)$, loses the part depending on t.

2.10 Itō's Formula

The goal of Itō's calculus is to furnish us with the dynamics of a smooth Markovian function of a Brownian Motion. Let us start from a deterministic smooth function F of (t, x). F varies only in response to changes in (t, x). A first-order approximation of the changes in our F is:

$$dF(t,x) = F_t(t,x)dt + F_x(t,x)dx$$

the philosophy behind these approximations is the use of a Taylor expansion. For example we could make a second-order approximation which would look like:

$$dF(t,x) = F_t(t,x)dt + F_x(t,x)dx + \frac{1}{2}(F_{xx}(t,x)(dx)^2 + 2F_{tx}(t,x)dtdx + F_{tt}(t,x)(dt)^2)$$

Usually second order elements of an approximation are negligible and are rarely considered. We now take t as the time paramater and consider a function Y which depends on time and on a Brownian motion W. We consider the following:

$$Y(t) = F(t, W(t))$$

Using our second order approximation we consider the following variation for Y in terms of t and W, our $dF(t, W(t))$ is therefore equal to:

$$F_t(t,W(t))dt+F_x(t,W(t))dW(t)+\frac{1}{2}(F_{xx}(t,W)(dW(t))^2+2F_{tx}(t,W(t))dtdW(t)+F_{tt}(t,W(t))(dt)^2)$$

In this case the second order approximations are important for the purpose of our study. Following the intuition that $dW(t) = W(t+dt) - W(t) \sim N(0, dt)$, we can approximate the square increment $(W(t))^2$ with its mean:

$$(W(t))^2 \sim dt$$

Itō's Lemma sums up what our findings where so far, and recites the following:

Let $F(t, x)$ be a smooth funciton (the minimal regularity required is $C^{1,2}(t, x)$). The Markov process defined by:

$$F(t, W(t))$$

has dynamics given by the following stochastic differential equation:

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

A diffusion, which is another name given to an Itō process, is any adapted process Y whose dynamics may be written as:

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

where α and β are two coefficients. The first one, α, is referred to as the drift of the process. In reality though, in Finance the practice is to call drift the fraction $\frac{\alpha(t)}{Y(t)}$. The second coefficient, β, is the diffusion of the process.
In the case of Brownian motions (B) with drift and Geometric Brownian motions (S) we have the following conditions. B verifies:

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

while the Geometric Brownian motion, $S(t) = exp(bt + \sigma dW(t))$ satisfies:

$$dS(t) = \left(b + \frac{\sigma^2}{2}\right) S(t)dt + \sigma S(t)dW(t)$$

where sometimes we can call $\mu = (b + \frac{\sigma^2}{2})$.

2.11 About Black-Scholes

We will now briefly consider the Black-Scholes-Merton enviroment as we will later on have to deal with its assets' processes. In this model we have only two assets, a money market risk-free bond and a capital market risky stock. The bond pays continuously an interest $r \geqslant 0$, ($B(t) = e^{rt}$) and has the following characteristics:

$$\begin{cases} dB(t) = rB(t)dt \\ B(0) = 1 \end{cases}$$

The risky asset on the other hand satisfies a stochastic differential equation with an initial condition (Cauchy's Problem). It is described as:

$$\begin{cases} dS(t) = \mu S(t)dt + \sigma S(t)dW(t) \\ S(0) = S_0 \end{cases}$$

here, S_0 is the observed, current, market price of the risky stock and the terms μ and σ are constants, with $\sigma \geqslant 0$.

Solving for the Cauchy problem, and finding its unique solution, leads us to find out that S must satisfy:

$$S(t) = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W(t)}$$

because of the lognormal property of the Geometric Brownian motion, we state that the marginals of the process $S(t)$ satisfy:

$$ln \frac{S(t)}{S_0} \sim N\left(\left(\mu - \frac{\sigma^2}{2}\right)t, \sigma^2 t\right)$$

meaning that the mean and the variance of the stock logreturns grow over time linearly.

2.11.1 Lognormal Property

Let's take a function, Y, of S such that:

$$Y = \ln S$$

practically we want to derive the process of the $\ln S$, where $dS = \mu dt + \sigma dW(t)$.
Using Itō's Lemma we can state that the process followed by Y is:

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

this is a standard Brownian motion with drift. For such reason we can justify the considration we made above and add that the logreturns of S_T are normally distributed with:

$$\ln S_T \sim N\left(\ln S_0 + \left(\mu - \frac{\sigma^2}{2}\right) T, \sigma^2 T\right)$$

2.12 Our Stochastic Process

The most important take away from this whole discussion is the form of our stock process,

$$S(t) = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W(t)}$$

We can see that the process is continuously compounded and, analyzing its exponent, we can identify two important elements. From now on we will refer to μ as drift and to σ as standard deviation or volatility. What is the meaning of these parameters? We can define by and large the first parameter, μ, as the expected return and the second, σ as a measure of risk.

3 Portfolio Optimization

3.1 Markowitz Model

The most basic approach to portfolio optimization is the one-period model described by Markowitz. It is a static model, meaning that the decisions taken at the initial time $t = 0$ remain unchanged and their outcomes are known only at the end. The investor is active only at outset, gathers the fruits of his investment at final time T, and is merely an observer in the meanwhile.
We start by considering a market with d securities with prices $p_1, p_2, ..., p_d > 0$ at time 0 and final prices $P_1, P_2, ..., P_d$ at the final time T, these prices are of course not foreseeable. They are random variables over the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.
The returns of the securities are computed as follows:

$$R_i(T) := \frac{P_i(T)}{p_i}$$

with $i = 1, 2, ..., d$. The mean, variance and covariance of the assets may be given or may be estimated. We define for the asets' return:

- $E[R_i(T)] = \mu_i$, as the mean
- $SD(R_i) = \sigma_i$, as standard deviation
- $Cov(R_i(T), R_j(T) = \sigma_{ij}$, as covariance

We proceede by making the following assumptions:

1. Securities are perfectly divisible, so we can hold also non-integers quantites of an asset. If we indicate with ϕ_i the number of shares invested in asset i, then this means that $\phi_i \in \mathbb{R}$.
2. Short-selling is not allowed, meaning that we cannot enter negative positions in an asset to finance the purchase of another. This ensures us that we won't end up with a negative final wealth. Stated: $\phi \geqslant 0$
3. There are no transaction costs in purchasing or selling securities.

If an investor starts off with an initial wealth $x > 0$, he is going to hold a number shares $\phi_i \geqslant 0$ of each asset. His budget constraint is going to be the following:

$$\sum \phi_i \cdot p_i = x$$

We shall refer to $\pi = (\pi_i, ..., \pi_d)$ as the portfolio vector, defined as:

$$\pi_i := \frac{\phi_i \cdot p_i}{x}$$

for $i = 1, 2, ..., d$. It is clear that the following condition must hold:

$$\sum_{i=1}^{d} \pi_i = 1$$

Return of our portfolio of assets is therefore:

$$R^{\pi} := \sum_{i=1}^{d} \pi_i \cdot R_i(T)$$

So we can now write the mean return of the portfolio as:

$$E[R^{\pi}] = \sum_{i=1}^{d} \pi_i \cdot \mu_i$$

Portfolio variance is:

$$Var(R^{\pi}) = \sum_{i=1}^{d} \sum_{j=1}^{d} \pi_i \cdot \pi_j \cdot \sigma_{ij}$$

Many portfolio can be created through a linear combination of the assets, each time attributing different weights to each component. The investor bases portfolio selection on two criteria:

1. He wants to maximize returns, meaning that the higher the return of a portfolio, the more appealing it is to his eyes
2. Because high returns are generally accompanied by high risk (volatility σ), the investor also consider the variance (as index of risk) of the portfolio. He will of course prefer low variance to high.

Because of this double objective, this model is also referred to as the mean-variance optimization model. An investor may intend to maximize returns constraining on a constant variance upper bound C_1. The problem is formulated in such a way:

$$\max_{\pi \in \mathbb{R}^d} E[R^{\pi}]$$

subject to $\pi_i \geqslant 0$, $\sum_{i=1}^{d} \pi_i = 1$, $Var(R^{\pi}) \leqslant C_1$.

Another investor, instead, might want to minimize the variance of the portfolio given a constant return lower bound C_2. In this case the problem formulation would be:

$$\min_{\pi \in \mathbb{R}^d} Var(R^{\pi})$$

subject to $\pi_i \geqslant 0$, $\sum_{i=1}^{d} \pi_i = 1$, $E[R^\pi] \geqslant C_2$.

The solution to these problems is mathematically obtained by simply using first order conditions (FOCs) but we will skip such passage to move on directly to the more complex Merton problem.

3.2 Merton Problem

The model put forth by Markowitz has the advantage of being simple to understand and widely taught and used. But one aspect in which it fails to be realistic is that it is a static model, meaning that the investor is active only once. Other models allow for continuous adjustments in the portfolio strategy and grant the investor a higher degree of freedom. This is the reason why we now introduce a new way of optimizing a portfolio, which derives from the solution of the so called "Merton Problem". Such topic is very broad as this problem has many variations which differentiate for time horizon, dividends streams, transaction costs, endowments, etc.

3.2.1 Wealth Process

An investor has the opportunity to invest in a risk-free asset and in one or more risky assets, say stocks. We provide first the general equation for the investor's wealth:

$$dw_t = r_t w_t dt + n_t(dS_t - r_t S_t dt + \delta_t dt) + e_t dt - c_t dt$$

which, in a simpler form, corresponds to:

$$dw_t = r_t(w_t - n_t S_t)dt + n_t(dS_t + \delta_t dt) + e_t dt - c_t dt$$

where e and δ account, respectively for the period endowment and dividend streams. These elements can be modeled in some variations of the Merton model but, in the basic cases (and also in our case) are assumed to be equal to zero. The consumption stream is represented by c and S stands for the stochastic process of the risky assets which has the form of a Geometric Brownian motion. The riskless rate is an adapted scalar, r. Usually r and S are given, just as the initial w_0, and the investment decisions of the individual concern the choice of consumption, c, and the portfolio process n. The model allows for changes in the quantity of wealth consumed and in the number of stocks purchased in each period of time. The investor initially chooses how much of w_0 to invest in stocks and how much in the risk-free asset. We can think of the latter as a bank account that returns an interest r. After setting such values, the total worth of the stocks bought will be $n_0 S_0$ and the bank account will add up to $w_0 - n_0 S_0$.

Beacause the values chosen for c and n can change over time, due to portfolio adjustments, it is important at this point to specify that the pair (n, c), which makes up the investor's strategy on how much to invest and to consume, has to be admissible at any time.

The definition for this condition is the following: a pair $(n_t, c_t)_{t>0}$ is said to be admissible for initial w_0 if the above wealth process w_t stays non-negative at all times. The set of all admissible strategies is therefore:

$$\mathcal{A} \equiv \cup_{w>0} \mathcal{A}(w)$$

We have seen what the investor has control over, but what is his objective? For a rational investor it is obvious that his desire is:

$$\sup_{(n,c)\in\mathcal{A}} E\left[\int_0^T u(t, c_t)dt + u(T, w_T)\right]$$

Which denotes the will of the investor to maximize the final expected value of the inter-temporal and final utilities over a specific time horizon (which could even be not finite).
As convention provides, we shall indicate the number of shares in the portfolio with n; we indicate the total value of the stocks with $\theta_t^i = n_t^i S_t^i$; with π_t^i we denote the proportion of wealth invested in asset i at time t. The relation among these variables is the following:

$$\theta_t^i = n_t^i S_t^i = \pi_t^i w_t$$

3.2.2 CRRA Utility function

We now specify the form and discuss the properties of the utility functions that invesors are assumed to possess in this model. They play a key role in determining how to maximize the outcome by means of our strategy. Merton considered among the possible families of utility functions a particular subset with the property of constant relative risk aversion (CRRA). What does this imply? CRRA are a particular set of utility functions belonging to the so called hyperbolic absolute risk aversion (HARA) family. The only functions of this wider set of functions that satisfy the CRRA property are referred to as isoelastic functions and are of the form:

$$u(t, x) = e^{-\rho t} \frac{x^{1-R}}{1-R}$$

or

$$u(t, x) = -\gamma^{-1} exp(-\rho t - \gamma x)$$

where ρ, γ, $R > 0$ and $R \neq 1$
The firt type of utility is the one we will focus on. The coefficient R is representative of risk aversion of the consumer. In behavioral finance such coefficient is considered relatively high or low based on the different models. Special cases for which we shall pay attention, bearing in mind that $R \geqslant 0$ (risk-aversion), are the following:

- $R = 1$, in this case we would have, solving through the use of l'Hôpital's rule, that the utility function would become $u(t.x) = e^{-\rho t}\log(x)$ and thus changes the function we are dealing with.
- $R \to \infty$, would correspond to infinite risk aversion
- Finally we consider, even if ruled out, the case in which $R = 0$. This would mean risk neutrality and therefore, because of our risk aversion context, doesn't concern our study.

CRRA means that our coefficint of risk-aversion R does not vary with scale, as we will see later on. In mathematical terms we can state that:

$$-c\frac{u''(c)}{u'(c)} = constant = R$$

The discount factor $-\rho t$ is made up of two components, an objective one and a subjective one. The first is time, the same for every agent, but the second, ρ is subjective discount rate for utilities and therefore varies among investors. It is a generalization, but we could assume it to be equal to the economy risk-free interest rate.
Using this particular type of functions is fundamental in our discussion as they are very practical to deal with. It is not surprising that they are in fact of common use in finance and economics in general.

3.2.3 Davis-Varaiya MPOC

The most commonly used method for solving this optimization problem is the value function approach which is based on the Martingale Principle of Optimal Control (MPOC). It states:

Suppose we have an objective of the form

$$\sup_{(n,c)\in\mathcal{A}} E\left[\int_0^T u(t,c_t)dt + u(T,w_T)\right]$$

and there exists a function $V : [0,T]\times\mathbb{R}^+ \to \mathbb{R}$ *which is continuous in at least its second derivative and such that* $V(T,\cdot) = u(T,\cdot)$*. Suppose also that for any* $(n,c) \in \mathcal{A}(w_0)$*:*

$$Y_t \equiv V(t,w_t) + \int_0^t u(s,c_s)ds \quad \text{is a supermatingale}$$

Then for some $(n^*,c^*) \in \mathcal{A}$*, the process* Y *is a martingale and such pair* (n^*,c^*) *is optimal. The value of the problem starting at* w_0 *is therefore:*

$$V(0,w_0) = \sup_{(n,c)\in\mathcal{A}(w_0)} E\left[\int_0^T u(t,c_t)dt + u(T,w_T)\right]$$

We prove this starting from the supermaringale property of Y, so for any $(n, c) \in \mathcal{A}(w_0)$,

$$Y_0 = V(0, w_0) \geqslant E[Y_T] = E\left[\int_0^T u(t, c_t)dt + u(T, w_T)\right]$$

holds. But in the case of (n^*, c^*) the inequality becomes an equality, proving the optimality of the strategy.

3.2.4 Value function approach

As mentioned before, this approach, most commonly used in solving the Merton Problem, lays its foundations on the MPOC principle.
Our asset follows a Geometric Brownian motion and its dynamics are the following:

$$dS_t^i = S_t^i\left(\sum_{j=1}^{N} \sigma^{ij} dW_t^j + \mu^i dt\right)$$

where σ^{ij} and μ^i are constants and W is a d-dimensional Brownian motion. Before proceeding in our discussion we restate that our risk-less rate r is constant and that the endowment e and the dividend payout δ are both equal to zero. For simplicity we write the process as:

$$S_t = S_t(\sigma \cdot dW + \mu dt)$$

And now we can express the wealth process in terms of θ (recall that $\theta_t^i = n_t^i S_t^i$):

$$dw_t = rw_t dt + \theta_t \cdot (\sigma dW_t + (\mu - r)dt) - c_t dt$$

Our goal is now to find a function V that satisfyies the MPOC condition. A good place to start is writing down our Y and perform an Itō expansion (we're assuming that V is sufficiently regular):

$$dY_t = V_t dt + V_w dw + \frac{1}{2}V_{ww}(dw)^2 + u(t, c)dt$$

follows that

$$dY_t = V_w \theta \cdot \sigma dW + \left\{u(t, c) + V_t + V_w(rw + \theta \cdot (\mu - r) - c) + \frac{1}{2}|\sigma^T \theta|^2 V_{ww}\right\} dt$$

the condition for the Itō expansion to be a supermartingale is that the drift be non-positive, where the drift is equal to zero, the Itō expansion is a martingale. This happens at the optimal strategy (θ^*, c^*), at this point the funtion V is our value function. So we consider

$$0 = \sup_{\theta, c} \left[u(t, c) + V_t + V_w(rw + \theta \cdot (\mu - r) - c) + \frac{1}{2}|\sigma^T \theta|^2 V_{ww}\right]$$

This particular partial differential equation (PDE) for the unknown value function V is named Hamilton-Jacobi-Bellman (HJB) equation. We will see that the HJB equation plays a key role in allowing us to find a solution. Yet, we don't know if there is any solution to this PDE, if it is unique and if the supremum (optimal strategy) is attained. In order to obtain some form of objective solution we need to assume that the utility function u is a CRRA function, of the form we saw before. Without such assumption it is difficult and rare for us to reach a defined solution.

3.2.5 Optimization over infinite horizons

Let us now try to tackle the Merton problem with an infinite time horizon. We will make use of a CRRA utility function u of the form:

$$u(t,x) \equiv e^{-\rho t} u(x) \equiv e^{-\rho t} \frac{x^{1-R}}{1-R}$$

The problem we now face is the following:

$$V(w) = \sup_{(n,c) \in \mathcal{A}(w)} E\left[\int_0^\infty e^{-\rho t} \frac{c_t^{1-R}}{1-R} dt\right]$$

This problem can be solved completely using a four-step methodology:

1. Use of special features to guess the form of the solution
2. Make use of the Hamilton-Jacobi-Equation to find the solution
3. Find a simple bound for the value of the problem
4. Verify that the bound is attained for our solution

Let us now analyze each of the steps individually.

1. We can write down the form of the solution by taking advantage of the scaling property, our dw_t is in fact linear. For any constant $k > 0$

$$\begin{aligned}
V(kw) &= \sup_{(n,c) \in \mathcal{A}(kw)} E[\int_0^\infty e - \rho t u(c_t) dt] \\
&= \sup_{(n,c) \in \mathcal{A}(w)} E[\int_0^\infty e - \rho t u(kc_t) dt] \\
&= \sup_{(n,c) \in \mathcal{A}(w)} k^{1-R} E[\int_0^\infty e - \rho t u(c_t) dt] \\
&= k^{1-R} V(w)
\end{aligned}$$

We call our constant γ_M and we have the form of our value function:

$$V(w) = \gamma_M^{-R} u(w) \equiv \gamma_M^{-R} \frac{w^{1-R}}{1-R}$$

for some constant $\gamma_M > 0$. The problem is now to find such γ_M.

2. We make use of the HJB equation to identify the unknown constant γ_M. We take

$$V(t,w) = \sup_{(n,c)\in\mathcal{A}(w)} E\left[\int_t^\infty e^{-\rho s} \frac{c_s^{1-R}}{1-R} ds | w_t = w\right]$$

because of the time-homogeneity of the problem we deduce that:

$$V(t,w) = e^{-\rho t} V(w)$$

where:

$$V(w) = \sup_{(n,c)\in\mathcal{A}(w)} E\left[\int_0^\infty e^{-\rho t} \frac{c_t^{1-R}}{1-R} dt\right]$$

We can therefore rewrite in the following form:

$$V(w) = \gamma_M^{-R} u(w) \equiv \gamma_M^{-R} \frac{w^{1-R}}{1-R}$$

and conclude that:

$$V(t,w) = e^{-\rho t} \gamma_M^{-R} u(w)$$

Now that we have the form of the solution, we need to optimize over the two factors that make up the investment strategy. To do this we return to our HJB equation, and we first start with optimization over θ. We procede in the following way:

$$(\sigma\sigma^T)\theta V_{ww} = -(\mu - r) V_w$$

From here we can solve for θ^* by rearranging the above equation. We find:

$$\theta^* = -\frac{V_w}{V_w w}(\sigma\sigma^T)^{-1}(\mu - r)$$

substituing $V(t,w) = e^{-\rho t}\gamma_M^{-R}u(w)$ we end up with the optimal wealth allocation in each asset i:

$$\theta^* = wR^{-1}(\sigma\sigma^T)^{-1}(\mu - r)$$

and now we can define the following N-vector as the Merton Portfolio.

$$\pi_M \equiv R^{-1}(\sigma\sigma^T)^{-1}(\mu - r)$$

θ^* tells us that each asset i held at any time $t > 0$ must have a total cash value proportional to the current wealth w_t with π_M^i as constant of proportionality, as we can see from:

$$(\theta_t^*)^i = w_t\pi_M^i$$

Optimization over c is achieved through the introduction of the following convex dual function of u:

$$\tilde{u}(y) \equiv sup\{u(x) - xy\}$$

for $u(x) = \frac{x^{1-R}}{1-R}$, so that:

$$\tilde{u}(y) = -\frac{y^{1-\tilde{R}}}{1-\tilde{R}}$$

where $\tilde{R} = R^{-1}$. So the optimization develops as:

$$\sup_c\{u(t,c) - cV_w\} = e^{-\rho t}\sup_c\{u(c) - ce^{\rho t}V_w\} = e^{-\rho t}\tilde{u}(e^{\rho t}V_w)$$

As we did for θ we now substitute the expected form of the solution and this gives us:

$$\sup_c\{u(t,c) - cV_w\} = e^{-\rho t}\tilde{u}(\gamma_M w)^{-R}) = -e^{-\rho t}\frac{(\gamma_M w)^{1-R}}{1-\tilde{R}} = e^{-\rho t}\frac{R}{1-R}(\gamma_M w)^{1-R}$$

Also in this case we obtain an optimizing c^* proportional to w which, before computing γ_M, we can momentairly write as:

$$c^* = \gamma_M w$$

We now put everything together and return the candidate value function $V(t,w) = e^{-\rho t}\gamma_M^{-R}u(w)$ to the Hamilton-Jacobi-Bellman equation. The result is:

$$\begin{aligned}0 &= e^{-\rho t}[\frac{R}{1-R}(\gamma_M w)^{1-R} - \rho\gamma_M^{-R}u(w) + rw\gamma_M^{-R}w^{-R} + \frac{1}{2}\gamma_M^{-R}w^{1-R}\frac{|\kappa|^2}{R}] \\ &= \frac{e^{-\rho t w^{1-R}\gamma_M^{-R}}}{1-R}[R\gamma_M - \rho - (R-1)(r + \frac{1}{2}\frac{|\kappa|^2}{R}]\end{aligned}$$

where $\kappa \equiv \sigma^{-1}(\mu - r)$ is the market price of risk vector. The value of γ_M is :

$$\gamma_M = R^{-1}\left\{\rho + (R-1)(r + \frac{1}{2}\frac{|\kappa|^2}{R}\right\}$$

The value function of the Merton Problem then is $V_M(w) \equiv V(t,w)$,

$$V_M(w) = \gamma_M^{-R}u(w)$$

The conclusions we make over the infinite-horizon problem is that we invest proportionally to wealth and we also consume proportionally to wealth. The constants of proportionality are the θ^* and c^* we computed.

What remains to do consider now is:

(a) What if γ_M is negative?

(b) Can we prove that γ_M is actually the optimal solution?

The first question relates directly to whether the Merton Problem is well-posed or not. It can be proven that the Merton problem is well posed if and only if $\gamma_M > 0$, therefore we discard the possibility that $\gamma_M \leqslant 0$. We will now address the second issue assuming that the problem is well-posed and that γ_M is positive.

3. Suppose we are given the initial w_0, we consider the wealth evolution under our conjectured optimal control. What we see is:

$$\begin{aligned}dw_t^* &= w_t^*\{\pi_M \cdot \sigma dW_t + (r + \pi_M \cdot (\mu - r) - \gamma_M)dt\} \\ &= w_t^*\{R^{-1}\kappa \cdot dW_t + (r + R^{-1}|\kappa|^2 - \gamma_M)dt\}\end{aligned}$$

which is solved by:

$$w_t^* = w_0 exp[R^{-1}\kappa \cdot W_t + (r + \frac{1}{2}R^{-2}|\kappa|^2(2R-1) - \gamma_M)t]$$

The proof of optimality is based on a trivial inequality:

$$u(y) \leqslant u(x) + (y - x)u'(x)$$

$(x, y) > 0$. which expresses the fact that the tangent to the concave function u at $x > 0$ lies everywhere above the graph of u. If we consider any admissible (n, c) then we can bound the objective to

$$\begin{aligned} E\left[\int_0^\infty e^{-\rho t} u(c_t)dt\right] &\leqslant E\left[\int_0^\infty e^{-\rho t}\{u(c_t^*) + (c_t - c_t^*)u'(c_t^*\}dt\right] \\ &= E\left[\int_0^\infty e^{-\rho t} u(c_t^*)dt\right] + E\left[\int_0^\infty (c_t - c_t^*)\zeta_t dt\right] \end{aligned}$$

where $\zeta_t \equiv e^{-\rho t}u'(c_t^*) \propto exp(-\kappa \cdot W_t - (r + \frac{1}{2}|\kappa|^2)t)$ is the state-price density, also referred to as stochastic discount factor. A property that follows is that for any admissible (n, c):

$$Y_T \equiv \zeta_t w_t + \int \zeta_s c_s ds$$

is a local martingale. This is verifiable through Itō's calculus. And from here we can be sure, given that wealth and consumption are non-negative, that Y is a non-negative supermartingale and thus:

$$w_0 = Y_0 \geqslant E[Y_\infty] \geqslant E\left[\int_0^\infty \zeta_s c_s ds\right]$$

4. Last step is to verify that

$$w_0 = E\left[\int_0^\infty \zeta_s c_s^* ds\right]$$

meaning that the optimal supreme is attained. Here c^* is our optimal consumption process. We use the bound condition and the conditons for w_0 to assert that:

$$E\left[\int_0^\infty e^{-\rho t} u(c_t)dt\right] \leqslant E\left[\int_0^\infty e^{-\rho t} u(c_t^*)dt\right]$$

assuring us that (n^*, c^*) is the optimal strategy.

3.2.6 Optimization on finite horizons

The solution to the finite horizon Merton problem is similar to the infinite horizon case. The first assumption we make in solving the problem is that the utility function is separable, and that constant relative risk aversion (CRRA) holds in consumption. THe objuective for the investor which we want to maximize has the form:

$$E\left[\int_0^T h(t)u(c_t)dt + Au(w_T)\right]$$

where h is strictly positive function and $A > 0$. Moreover, $u'(x) = x^{-R}$ for some $R > 0$, $R \neq 1$. We can again make use of the scaling properties of the CRRA functions and obtain that the value function:

$$V(t,w) = \sup E\left[\int_t^T h(t)u(c_t)dt + Au(w_T)|w_t = w\right]$$

which for some function f, must have the form:

$$V(t,w) = f(t)u(w)$$

The Hamilton-Jacobi-Bellman equation for this problem is:

$$0 = \sup_{\theta,c}[u(t,c) + V_t + V_w(rw + \theta(\mu - r) - c) + \frac{1}{2}\sigma^2\theta^2 V_{ww}]$$

If we substitute our scaled form of the function into the previous equation we obtain:

$$0 = \sup_{y,q} u(w)[f' + (r + y(\mu - r) - q)(1-R)f - \frac{1}{2}R(1-R)\sigma^2 y^2 f + hq^{1-R}]$$

where $y = \theta/w$ and $q = c/w$. In this case optimality is given by:

$$y = \pi_M \quad \text{and} \quad f = hq^{-R}$$

We moreover assign $f(T) = A$. We can then conclude that:

$$\theta_t^* = \pi_M w_t \quad \text{and} \quad c_t^* = w_t\left(\frac{h(t)}{f(t)}\right)^{1/R}$$

In general we take as $h(t) = e^{-\rho t}$ which allows us to compute both $f(t)$ and $f(T) = A$. Comparing to the infinite horizon case, we have the same investment strategy but now we no longer consume at a constant rate proportional to initial wealth.

4 Monte Carlo Simulation

In this section we will develop on the concept of Monte Carlo processes, describe what they are, what they are used for and how they are put forth. Later on we will make use of this particular concept to achieve a simulation of an investment's outcome.

4.1 Generics

A Montecarlo simulation is a stochastic method used in computing deterministic quantities. Imagine we were to approximate the expected value of a random variable X, assumed integrable and with expectation of α. We would be working on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $X : (\Omega, \mathcal{F}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. From now on our expected values are defined in conformity with $\mathbb{P}$. Let us denote with Λ the law regulating X. We assume that X is integrable in its square. We define as $\{X^q\}_q$ a succession of independent random variables, all having Λ as the law behind them. The law of large numbers allows us to state that:

$$\hat{\alpha} := \frac{1}{Q}\sum_{q=1}^{Q} X^q \longrightarrow E(X) = \alpha \quad \text{as } Q \to \infty$$

Then a good way to approximate our α is to compute the empirical average of a fairly large number of independent variables with law equal to X. This is exactly what a Montecarlo method does.
We point out that:

1. this method can be programmed fairly easily with a computer by using Excel or another programming language.
2. the estimate $\hat{\alpha}$ is subject to two main types of error. These are:
   (a) firstly the fact that we are fixing a finite number of observations Q (even if very large)
   (b) the second is related to the fact that $\hat{\alpha}$ is itself a random variable, so it can turn out different values for each time we repeat the simulation.

The Central limit theorem can be used to study the efficiency of the estimate, meaning that it can be used to study its convergence velocity.

$$\sqrt{Q}(\hat{\alpha} - \alpha) \to N(0, C)$$

We observe a convergence in law to a normal gaussian $(0, C)$; where C is the covariance matrix for the random vector $X = (X_1, ..., X_d)$ so where $C_{ij} = Cov(X_i, X_j)$. For very big values of Q, the following approximation also holds:

$$\hat{\alpha} \cong \alpha + \frac{1}{\sqrt{Q}} C^{\frac{1}{2}} Z$$

Where Z represents a random variable with standard gaussian over $\mathbb{R}$. From here we state that the random error we commit in estimating α with $\hat{\alpha}$ is of the order of $\frac{1}{\sqrt{Q}} C^{\frac{1}{2}} Z$. Thus, the convergence speed of a Montecarlo simulation is fairly slow $\frac{1}{\sqrt{Q}}$. Nevertheless, we see that an important element is its "flexibility", as the speed of convergenge is independent of the number of random variables taken into account,d. This is probably the most distinctive element of this type of simulation.
Let us fix a confidence level δ in order for us to compute double-sided confidence intervals. Taking, for the purpose of making an example, $d = 1$ we have:

$$\hat{\alpha} \cong \alpha + \frac{\sigma}{\sqrt{Q}} Z$$

Once we fix δ, a bilateral confidence interval for α at the δ level has the following extremes:

$$\hat{\alpha} \pm \frac{\sigma}{\sqrt{Q}} \Phi_{\frac{1+\delta}{2}}$$

where $\Phi_{\frac{1+\delta}{2}}$ denotes the $\frac{1+\delta}{2}$ quantile.
In the case of a confidence interval of 95% ($\delta = 0.95$), we have that:

$$\text{for } \delta = 0.95, \quad \text{we have } \Phi_{\frac{1+\delta}{2}} = \Phi_{0.975} = 1.96$$

for further description we state that for:

$$\delta = 0.9, \quad \text{we have } \Phi_{\frac{1+\delta}{2}} = \Phi_{0.95} = 1.645$$

Typically the confidence level required, if not otherwise specified, is 95%. We observe, though, that the variace of the sample is known. If thats not the case, as often happens, we compute an estimator for such variance which is computed in the following way:

$$\hat{\sigma}_Q^2 = \frac{1}{Q-1} \sum_{q=1}^{Q} (X^q - \hat{\alpha}_Q)^2 = \frac{1}{Q-1} (\sum_{q=1}^{Q} (X^q)^2 - Q\hat{\alpha}_Q^2)$$

In these type of situations our confidence interval is:

$$\hat{\alpha}_Q \pm 1.96 \sqrt{\frac{\hat{\sigma}_Q^2}{Q}}$$

4.2 Monte Carlo Simulation of a Brownian Motion

What we are interested in though is the simulation of a Geometric Brownian Motion. To get here let us first consider the normal Brownian Motion.
Let us recall the characteristics of a Brownian motion:

1. $B_0 = 0$
2. B independent increments
3. the increments $B_t - B_s$, $s < t$, are gaussian with $N(0, t-s)$

To generate a Brownian trajectory $[0, T]$ we need to divide such time span into N fairly small intervals. We can fix such width $h = T/N$, and call $t_k = kh$ the extremes of such small intervals, $k = 1, 2, ..., N$. We simulate B at the times t_k by taking N standard and independent gaussians $U_1, ..., U_N$ and setting:

$$B_0 = 0 \quad \text{and} \quad B_{t_k} = B_{t_{k-1}} + \sqrt{h}U_k$$

for $k = 1, 2, ..., N$. In other words:

$$B_{t_k} = \sqrt{h}\sum_{j=1}^{k} U_j$$

with $k = 0, 1, ..., N$. To simulate the Brownian motion we need to generate a d-vector of gaussian random variables. How do we do this? Many programs can do this for us, in this particular case we we will make use of Excel.

4.3 Monte Carlo Simulation of a Geometric Brownian Motion

At this point the simulation of a Geometric Brownian motion, which we need for our purposes, comes at hand in an easier way. Let's first remind the reader that a process S_t follows a Geometric Brownian motion if:

- $S_0^i = x^i$
- $dS_t = \mu^i S_t^i dt + \sum_{j=1}^{d} \sigma_j^i S_t^i dB^j(t)$

with $i = 1, ..., d$ and where σ denotes the volatility of the motion and is a $d \times d$ invertible matrix. It follows that:

$$S_t^i = x^i exp\left(\left(\mu^i - \frac{1}{2}\sum_{j=1}^{d} \sigma_j^{i^2}\right) t + \sum_{j=1}^{d} \sigma_j^i B_t^j\right)$$

Again we fix $h = \frac{T}{N}$ and $t_k = kh$, $k = 1, ..., N$ and in analogy with the previous case we obtain:

$$S_0^i = x \quad \text{and} \quad S_{t_k} = S_{t_{k-1}} exp((\mu^i - \frac{1}{2}\sum_{j=1}^{d} \sigma_j^{i^2})h + \sqrt{h}\sigma_j^{i^2})t + \sum_{j=1}^{d} \sigma_j^i U_k^j)$$

for $i = 1, ..., d$ and where $U_1, ...U_N$ are N independently distributed random variables, with standard gaussian distribution in $\mathbb{R}$,$(U_k = U_k^1, ..., U_k^1)$.

5 Simulation of Portfolio

In this section we will provide a simulation of the paths followed by a Merton portfolio. The scope is to simulate what would be the expected result of an investment made in the near future (so that business cycle interferances would be minimized). The experiment could be well applied to the 40 main Italian stocks, contained in the FTSE MIB index. This is why we will focus first on the Index but then we will make the procedure shorter by simulating our optimization for a single stock. Notice that the procedure remains exactly the same and solving for the 40 stocks can be done in an analogous way. We will start off with an initial Index value (also referred to as price)and assign to it a stochastic process in the form of a geometric Brownian motion,

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

5.1 Data

In the computation of the coefficients μ and σ we will be using data obtained through the Bloomberg Terminal. The data set of the values of the index have a time span of a bit less than one year. Why such choice? The reason behind it is that the FTSE MIB like most indeces reflects the performance not only of a set of securities but of the economy (even global) as a whole. Because of the 2008 crisis and the soverign debt crisis of 2012, many values characterizing the index have been influenced by exogenous factors. We assume that, starting from the period specified, these shocks have a lesser impact and therefore we consider this period as proxy for the years to come, or better, for the near future. Considering a time span of, say, 30 years could lead to a bias in the return and volatility coefficients also because of the changing historical and institutional factors (wars, EU creation, ".com" bubble and so on...). After this preliminary reasoning, data on the FTSE MIB for the last year has been gathered. We recall that the FTSE MIB is an index consisting of the 40 most important Italian listed companies, weighted for their market capitalization. The data used for the purpose of our computations would be too much to be inserted but is accessible through sources such as the above mentioned Bloomberg Terminal. Anyways we will describe the process just for the purpose of clarity. On an Excel worksheet, 50 weekly values of the index have been chronologically listed starting from the 01/07/2016 to the 05/06/2017. The logreturns of such value were computed using

$$\ln r_t = \ln \left(\frac{Pt}{P_{t-1}} \right)$$

The values we obtained are the following:

$$\mu = 0,53\% \quad \sigma = 9,56\%$$

as a result of the following computations. The drift μ stemmed out of the average of the logreturns per period (the average of the 49 different $\ln r_t$). The standard deviation of our sample has been computed using the Excel formula: STDEV(). The volatility was computed in percentage terms as the ratio of the standard deviation of our sample over the average closing value of the index. After we obtained the data needed we turned to the very simulation. The Brownian motion (W_t) component of our Geometric Brownian motion has been computed as $\sigma\epsilon\sqrt{t}$. Because our intervals are all unitary, of a week each, the square root term has been disregarded, σ is our volatility and ϵ is the random component. The Excel function used to obtain the random values with gaussian distribution $N(0,1)$ for the purpose of simulation, is the combination of two embeded commands: NORMSINV(RAND()). The outcome is a random number with our desired distribution. The Excel command RAND() returns a random number uniformly disributed across $[0,1]$. The command NORMSINV() is the inverse of the standard normal cumulative distribution. The procedure of simulating has been repeated for 15 times, every time for a time span of 50 weeks (from time t_1 to t_{51}, where t_1 is the initial time). All the paths start from the opening price of the index which is $20.721,04$, the closure price of the 05/06/2017. Before simulating our portfolio optimization we look at the possible paths taken by the index to familiarize with the concept of Monte Carlo and even to give some extra knowledge over the possible investment in the Index.

5.2 Path of the index

5.2.1 Creating Random Numbers in Excel

How are we actually sure we are generating a random output Y with a desired distribution F?

Proposition

Let X be uniformly distributed on $[0,1]$. Assume F is invertible. Then, $Y := F^{-1}(X)$ has distribution F.

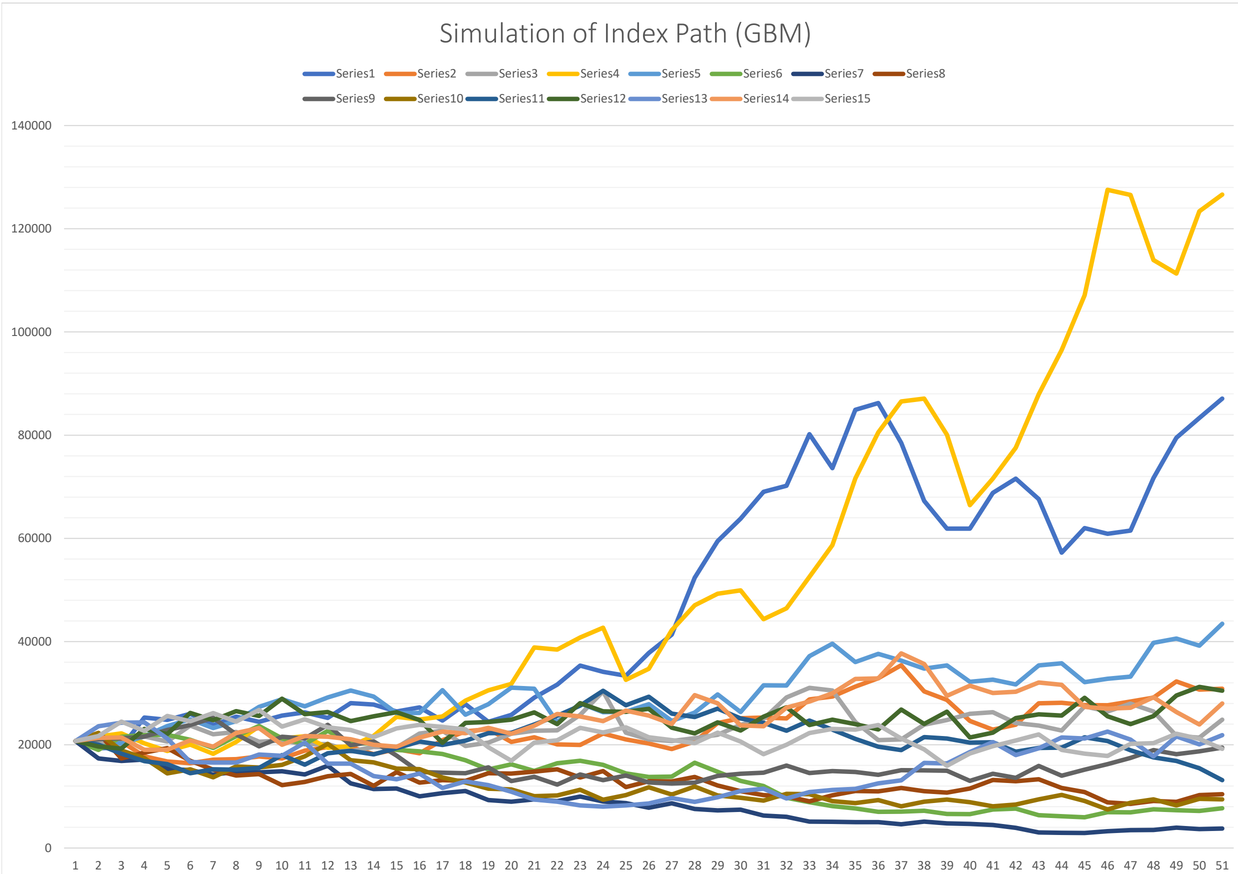
Proof

We show that Y has distribution F in the following way:

$$P(Y \leqslant y) = P(F^{-1}(X) \leqslant y) = P(X \leqslant F(Y)) = F(Y)$$

because X is uniformly distributed over $[0,1]$. This explains the use of the Excel function NORMSINV() over the function RAND() in our computation.

5.2.2 The path



Each of the series shown in the graph is a possible outcome after 50 weeks and to see 15 of them together can give an idea of what to expect from the future. The Monte Carlo simulation obtained shows that investing in the index can be extremely profitable in the best outcome scenarios but, in other cases, it could lead to losses sometimes close to the full amount (series 7). The determinant of the outcome is aleatory, the uncertainty component $\epsilon \cdot \sigma$. We notice that, since the drift is a constant percentage of the closing value of the previous period, increases tend to be magnified (explains the very high returns of the series 1 and 4) while losses tend to be more contained as they are mitigated by the drift component.

5.3 Merton Simulation

The next step is considering the 40 different stocks in the FTSE MIB as 40 different assets, assign an initial wealth w_0 and an utility function for the investor. The utility funciton we will use is of the kind $u(t,x) = e^{-\rho t}\frac{x^{1-R}}{1-R}$ and, based on empirical evidence we can assume a coefficient of risk aversion $R = 5$. The risk-free rate, in the light of the not so happy global economic situation will be considered as $r = 0,1\%$. Because of the absecnce of the subjective utility discount factor (ρ), we will take it as equal to the risk-free interest rate (considering

the latter as the average of the subjective utility discount factors). Let's assume that our investor has an initial wealth of 30000 euros. What we need is the variance-covariance matrix to compute the Merton portfolio which will allow us to arrive to the final values in terms of utility for the investor. The data for the weights and the composition of the FTSE MIB is available on the website of the *Borsa Italiana* and simple calculations through an Excel spreadsheet will furnish us with the necessairy information required for the optimization.
To keep our analysis more simple we will make such simulation considering only the risk-free asset and a stock, Intesa San Paolo, one of the biggest companies for capitalizzation percentage in the Italian stock index. We took the historical data regarding Intesa San Paolo from Yahoo Finance, and such data regards the same time period as before, from the 03/07/2016 to the 04/06/2017. The data has been analyzed as before, using an Excel spreadsheet. The drift and volatility for the stock which we obtained are:

$$\mu = 0,65\% \quad \sigma = 5,77\%$$

computed with the same methodology applied to the Index.
We can now use the discussion made in Section 3 over the Merton Portfolio optimization to compute our results. The full set of the variables we will use is:

1. $\mu = 0,65\%$
2. $\sigma = 5,77\%$
3. $r = 0,1\%$
4. $u(t,x) = e^{-\rho t}\frac{x^{1-R}}{1-R}$ with $R = 5$

Our wealth process is going to be:

$$dw_t = r_t(w_t - n_t S_t)dt + n_t dS_t - c_t dt$$

We verify whether or not the Merton problem is well-posed by computing the value of γ_M. Bearing in mind that we need the condition $\gamma_M > 0$ to procede, we move on:

$$\begin{aligned}
\gamma_M &= R^{-1}\left\{\rho + (R-1)(r + \frac{1}{2}\frac{|\kappa|^2}{R}\right\} \\
&= 5^{-1}(0,001 + 4(0,001 + \frac{1}{2}\frac{(0,0577^{-1}\cdot 0,0055)^2}{5}) \\
&= 0,0017269 > 0
\end{aligned}$$

here notice that we used as discount factor for utility (ρ) the risk-free rate ($r = 0,1\%$) as we mentioned before. We simply plugged in our values and the result obtained ensures us that the problem can be solved.

Considering the infinite horizon Merton problem we would have the following result for our investment strategy:

$$\begin{aligned}\theta^* &= w \cdot \frac{\mu - r}{\sigma^2 R} \\ &= 30000(0,0055)/(5 \cdot 0,0577^2) \\ &= 30000 \cdot 0,3304 \\ &= 9912\end{aligned}$$

which is the wealth allocation in Intesa San Paolo. The rest would be placed in the bank account and generate an interest of 0,1%. If the assets had been N (or 40 as in the case of the Index) instead of only one, we would have computed the weights of each asset and obtained an N-dimensional Merton portfolio π_M. In our case, with only one asset, we obtain a unitary vector which would be:

$$\pi_M = \big(\text{Intesa San Paolo}\big) = \big(0.3304\big)$$

This result, as we saw, also holds for the case of finite horizons. What is changing between the two cases is the consumption strategy. In the case of the infinite horizon we would consume proportionally to wealth. The optimal strategy would be given by:

$$c^* = \gamma_M w = 0,0017269 \cdot 30000 = 51,807$$

Since we are reasoning in weekly terms, our investor would consume 51,807 euros of wealth per week thanks to his investment strategy. This may seem a bit restricting but recall that we are disregarding any form of endowment (e) per period which could significally boost up this number (nevertheless a consumption of 51 euros per week could grant at least substinance in many countries).
In the finite horizon case we face a different problem, our conclusion was the following:

$$c_t^* = w_t \left(\frac{h(t)}{f(t)}\right)^{1/R}$$

We are going to assume a one year investment in which we adjust strategies every week. This means that for us $T = 52$ and the increments are $t = 1$ (in weeks as mentioned). The functions assigned to h is $h(t) = e^{-\rho t}$. The computations here are more complex as consumption is not a fixed proportion of the initial wealth but rather varies over time, for this reason we will do the calculus for the simulation using a worksheet (again Excel). In the end, we are going to be able to graphically analyze the outcomes in order to better acknowledge the processes for wealth and for consumption in the one year period. Let us before describe our procedure. We started a table at time zero, at outset we have an initial wealth of 30000, consumption is computed with our formula for c as a function of (W, t). The new wealth, at $t = 1$ is going to be given by our wealth process. It will include, apart from the bank account balance, the variations of our stock and the consumption of the previous period will be taken away. We simulate the perfomance of Intesa San Paolo stocks utilizing a Monte Carlo simulation, in which the starting point is given by the closure of the stock at 05/06/2017 ($2,56$ euros) and the drift and volatility are the ones previously computed ($\mu = 0,65\%$ and $\sigma = 5,77\%$). The number of experiments (Q) for the Monte Carlo simulation is 15 as we already stated. From the graphs shown we notice how consumption per period is not fixed. It is, in fact, adjusted in each period and turning out to be very volatile on the one hand. On the other hand, though, we see a tendency to increase over time,at the same pace of wealth. We do not consume proportionally to the initial wealth but consumption is still proportional (even if not with a fixed constant rate) to the latter. As wealth tends to generally increase over time, we can see that the quantity consumed basically comoves with it. We repeat the the wealth evolution in each perido (wealth process) is the result of the sum of the total value in stocks and the bank account minus the consumption of the previous period.

5.4 Conclusions

We see that using this portfolio optimization method is fairly profitable. The final computations show an average final wealth of 31.561 meaning an overall return in wealth of 5.203%, and an average consumption per period of 53.07 euros. The strategy seems very attractive but it is important to bear in mind that it lies on a few assumptions. These include for example CRRA utility functions, the existance of risk-free assets or the possibility of buying infinitesimal amounts of shares. But overall the solution to this widely spread problem of optimization is done taking into considerations much many aspects Even though assumptions make models less realistic they are a necessary condition for the model to have a sense in the first place. The aim of the thesis, which hopefully has been fulfilled, has been to show how this model, working with only a few necessary assumptions (which benefit from a relatively high degree of plausibility) reaches a solution which is more practical than other models. The final simulation makes us reach the conclusion that the Merton model of optimization is generally reliable, at least in the current market conditions.

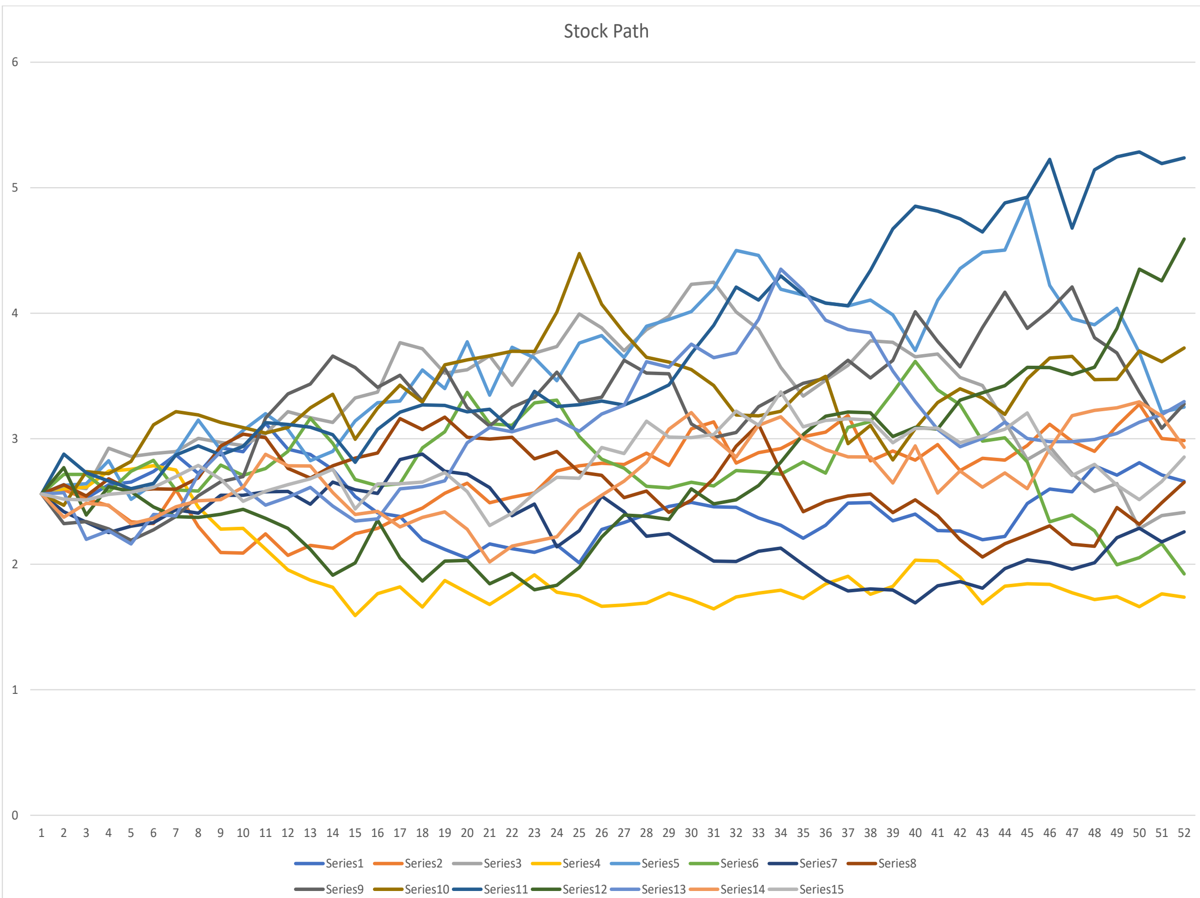


Figure 1: The simulation of the possible path from here to one year from now of Intesa San Paolo, weekly intervals of time

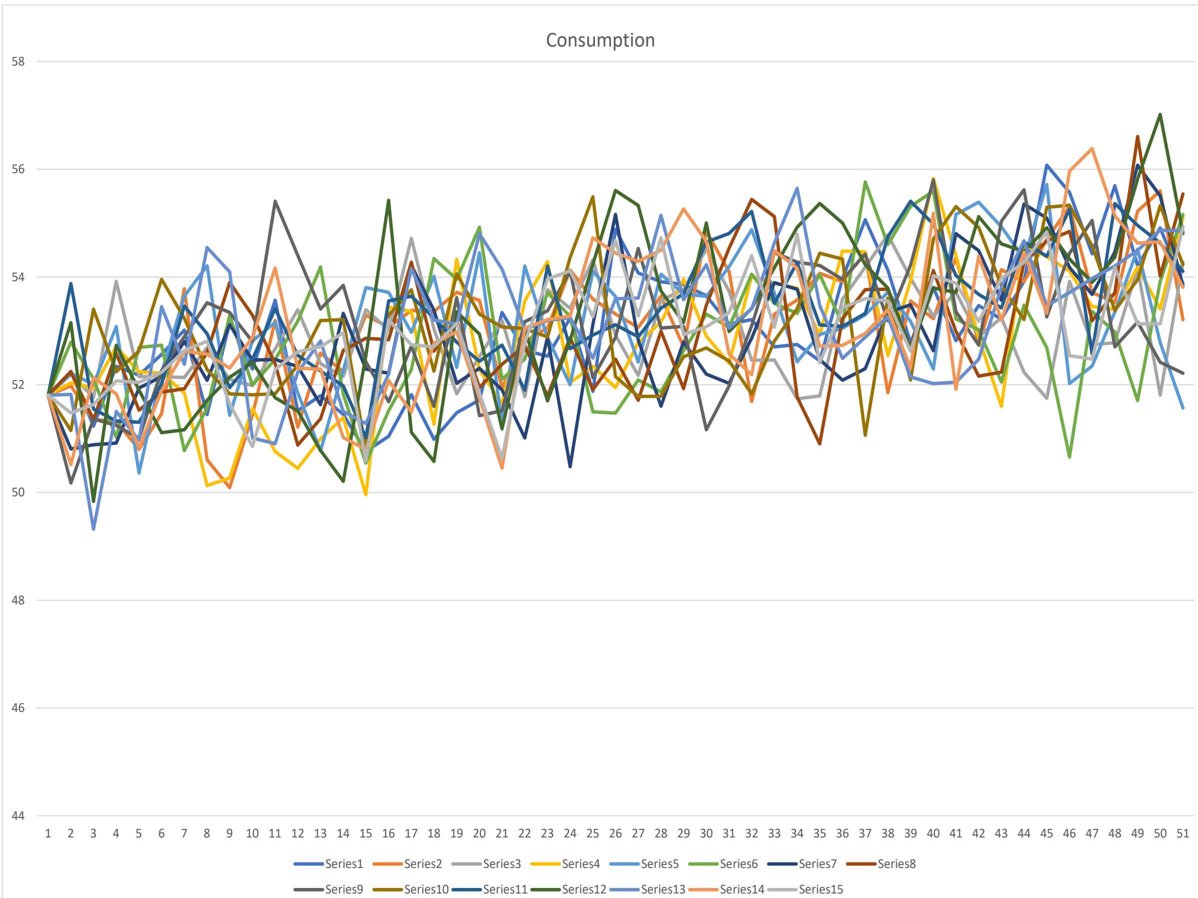


Figure 2: The consumption per week from here to one year following our Portfolio strategy

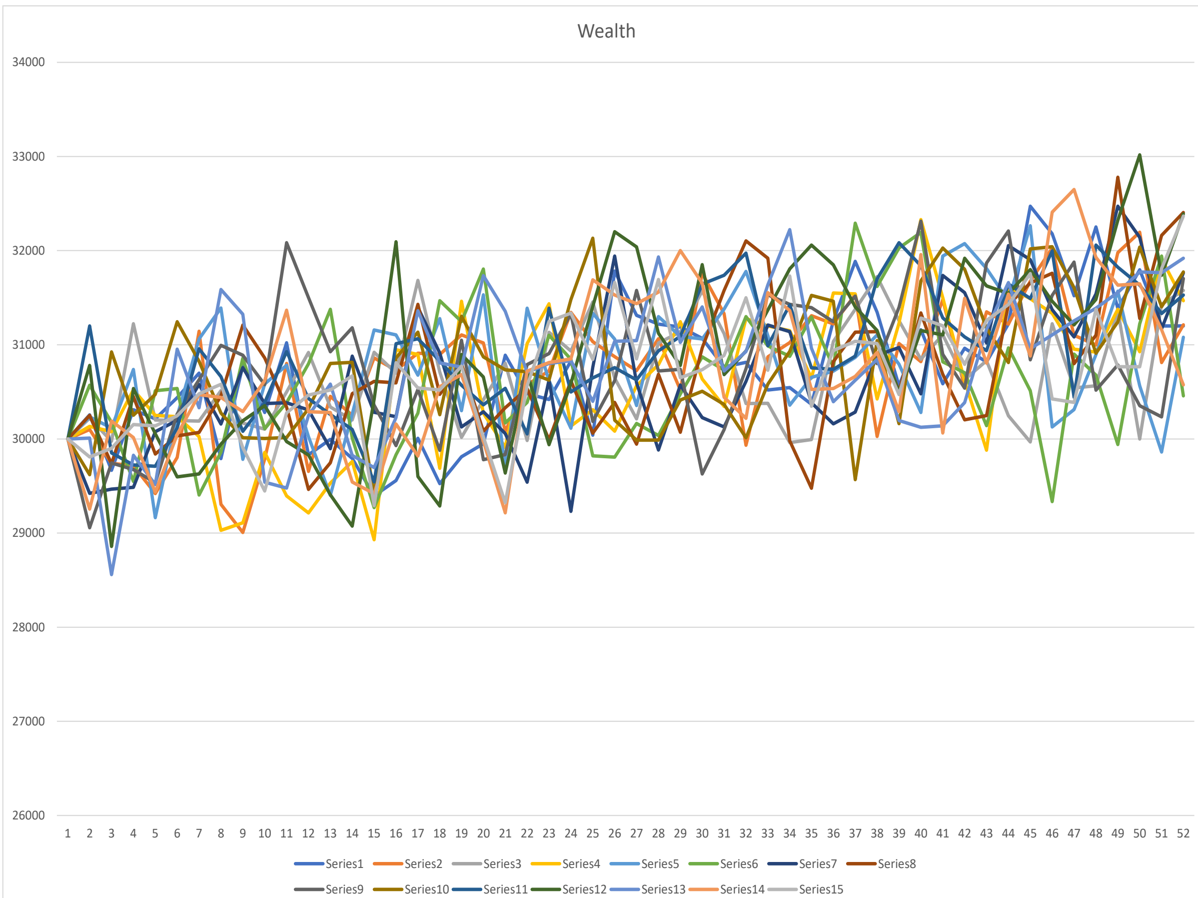


Figure 3: The evolution per period of the total amount of our wealth in the 52 weeks from today

References

[1] Sara Biagini, *Lecture Notes of Mathematical Methods for Financial Markets*, Department of Economics and Management, Pisa University

[2] Ralf Korn, Elke Korn, *Option Pricing and Portfolio Optimization*, American Mathematical Society, Providence, Rhode Island, 2000

[3] L.C.G. Rogers, *Optimal Investment*, SpringerBriefs in Quantitative Finance, 2013

[4] John C. Hull, *Options, Futures and Other Derivatives*, University of Toronto, 9th edition, 2014

[5] Lucia Caramellino, *Metodi Monte Carlo in Finanza*, Dipartimento di Matematica, Università di Roma Tor Vergata

[6] www.investopedia.com, *Monte Carlo Simulation With GBM*, accessed[14-04-2017] *Volatility*, accessed [18-05-2017], *Greeks*, accessed [21-05-2017]

[7] Lars Peter Hansen, Kenneth J. Singleton *Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns*

[8] Ralf Korn, Elke Korn, *Monte Carlo Methods in Finance and Insurance*, Chapman & Hall, Springer, 2nd Edition

[9] Philip E. Protter, *Stochastic Integration and Differential Equations*, Springer, 2nd edition

[10] Bloomberg Terminal, *Index graph for FTSE MIB from 01/07/16 to 05/06/17*, *Stock Price for Intesa San Paolo from 03/07/16 to 05/06/17*, Bloomberg database, Luiss Guido Carli University, Rome

[11] www.investexcel.net, *Simulate Geometric Brownian Motion with Excel*, accessed [17-05-2017]