

POLITECNICO DI MILANO
School of Industrial and Information Engineering
Master of Science in Mathematical Engineering



TITLE: VERY INTERESTING SUBJECT,
AIN'T IT?

Supervisors: Prof. Daniele Marazzina
Prof. Ferdinando M. Ametrano

Master thesis by:
Samuele Vianello
ID: *****

Academic year 2017-2018

*The secret to happiness is freedom.
And the secret to freedom is courage.*

Thucydides

Contents

List of Tables	iv
List of Figures	v
List of Algorithms	vi
Abstract	vii
Acknowledgements	viii
1 Introduction	1
1.1 Thesis structure	1
2 Correlation Analysis	2
2.1 Empirical Correlation of Returns	2
2.2 Correlation Significance	3
2.2.1 t -test	4
2.2.2 Permutation test	4
2.2.3 Significance results	4
2.3 Rolling Correlation	4
3 Presentation of the Models	7
3.1 Preliminary Notions	7
3.1.1 Geometric Brownian Motion	7
3.1.2 Poisson Process and Compound Poisson Process	8
3.1.3 CIR Process	10
3.2 Merton Model	11
3.2.1 Original Univariate Model	12
3.2.2 Multivariate Model	13
3.3 Heston Model	14
3.3.1 Univariate Heston Model	14
3.4 Bates Model	17

4	Calibration of the Models	18
5	Markowitz Portfolio Optimization	19
6	Conclusions	20
A	Characteristic Function of a Compound Poisson Process	21

List of Tables

List of Figures

2.1	Plots of rolling correlation for the different asset classes (on top) and significance for each value (on bottom). Blue lines are the 3-year rolling correlations, while the black ones have a window of 18 months. Both computations are updated monthly.	5
3.1	Poisson processes and compound Poisson processes with different λ .	9
3.2	Two trajectories of the same CIR process: $dr_t = 3(2 - r_t)dt + 0.5\sqrt{r_t}dW_t$ with different starting point. We can see the mean-reverting effect that attracts both trajectories to the value $\theta = 2$.	11

List of Algorithms

Abstract

Including Bitcoin in an investment portfolio increases portfolio diversification.

Acknowledgements

add acknowledgements

Thank you.

Chapter 1

Introduction

1.1 Thesis structure

Chapter 2

Correlation Analysis

In order to get an initial insight on how Bitcoin is correlated with other assets, we will perform a correlation analysis based on the empirical time series of our data. We will focus our attention on the logarithmic returns it is the standard practice. We will often refer to logarithmic returns simply as returns, only specifying their nature when it is necessary to avoid confusion.

2.1 Empirical Correlation of Returns

We first start by performing some statistical analysis on the data in order to estimate the distribution from which they are sampled. For this part, we will consider our data as successive samples of a N -dimensional vector in \mathbb{R}^N , where N is the number of assets:

$$\mathbf{x}_j = \begin{pmatrix} x_{1,j} \\ x_{2,j} \\ \vdots \\ x_{N,j} \end{pmatrix}, j = 1 \dots N_{sample}$$

Each element i of the vector \mathbf{x}_j represents the j^{th} realization of the returns for asset i .

Following basic statistics, we can now compute the *sample mean* of our vectors of returns as:

$$\bar{\mathbf{x}} = \frac{1}{N_{sample}} \sum_{j=1}^{N_{sample}} \mathbf{x}_j = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_N \end{pmatrix}$$

where $\bar{x}_i = \frac{1}{N_{sample}} \sum_{j=1}^{N_{sample}} x_{i,j}$ is the sample mean of component i .

Now we compute the *sample covariance matrix* through the following formula:

$$\bar{\Sigma} = \frac{1}{N_{sample} - 1} \sum_{j=1}^{N_{sample}} (\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})^T$$

where $\bar{\mathbf{x}}$ represent the sample mean of the returns just introduced.

All the information needed to obtain the *correlation matrix* C are already included in $\bar{\Sigma}$, we only need to perform some further calculations:

$$C_{i,j} = \frac{\bar{\Sigma}_{i,j}}{\sqrt{\bar{\Sigma}_{i,i}\bar{\Sigma}_{j,j}}} \quad (2.1)$$

We have thus obtained an empirical estimate of the correlation between our assets returns. The formula in (2.1) is often referred to as *Pearson correlation coefficient*, from the name of the English mathematician Karl Pearson who first formulated it.

Results are reported in the following tables.

***** ADD RESULT TABLES *****

We are mainly interested in the correlation between Bitcoin and other assets returns, so we will now focus on the first row (or equivalently column, by symmetry) of the correlation matrix.

All values are fairly close to zero, never exceeding 10% towards the positive or the negative side. One may thus wonder whether these correlations are *statistically significantly* different from zero. To answer this question, we will introduce two statistical tests to check the correlation significance.

2.2 Correlation Significance

The very core of Inferential Statistics, the branch of statistics that allows to draw conclusions from the information contained in a set of data, is hypothesis testing.

In our case, we are specifically interested in testing if the sample correlation coefficients are significantly different from zero or not. Both of the following tests are presented in the most general form for a sample of two variables, their distribution correlation ρ and their sample correlation $\hat{\rho}$.

Following standard testing procedure, we specify the *null hypothesis* and the *alternative hypothesis*:

$$\mathbf{H}_0 : \quad \rho = 0 \quad vs. \quad \mathbf{H}_1 : \quad \rho \neq 0$$

These will be common to both presented tests.

2.2.1 t -test

Our first test is based on Student's t -distribution and the following t -statistic:

$$t = \hat{\rho} \sqrt{\frac{n-2}{1-\hat{\rho}^2}} \quad (2.2)$$

which under the null hypothesis is distributed as a Student's t with $n-2$ degrees of freedom, where n stands for the cardinality of the sample. We can thus proceed by computing the relative p-value and compare it to a given level of confidence α (usually $\alpha = 95\%$). The result of the test will be deduced as follows:

- $p - value < 1 - \alpha$: we have statistical evidence to state that the correlation is *significantly* different from zero;
- $p - value \geq 1 - \alpha$: there is *no statistical evidence* to state that the correlation is different from zero.

2.2.2 Permutation test

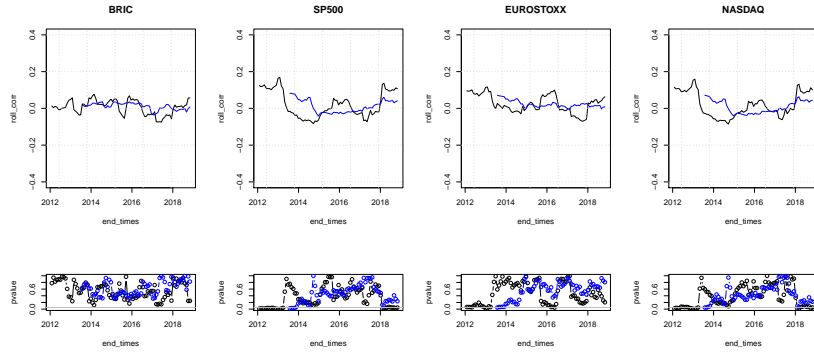
The permutation test is based on building an empirical distribution of values for the correlation by sampling different pairs of X and Y variable and then computing Pearson's correlation. If this is done a large enough number of times, we obtain an empirical distribution of possible values. From this distribution we can then obtain the p-value of the test and thus the final result in the same way as in the previous case.

2.2.3 Significance results

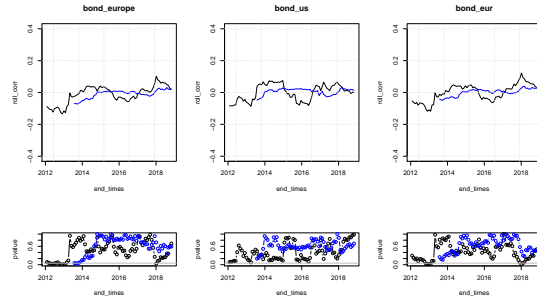
*****Comments and numerical results*****

2.3 Rolling Correlation

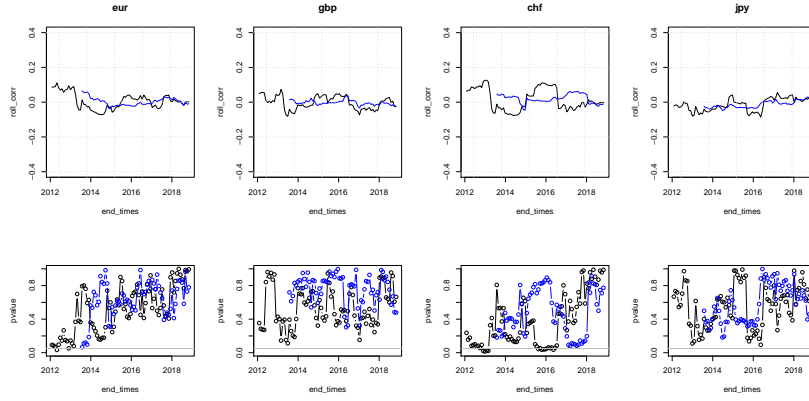
Our study so far has focused on the analysis of the dataset as a whole, with values spanning from 2010 to 2018. This is clearly important if we want to obtain a general overview of the period, but it is also interesting to see how the correlation between the assets has evolved through. Therefore, we present in Figure 2.1 the results obtained from calculating the correlation between Bitcoin and the other assets using rolling windows of 36 and 18 months, updated monthly.



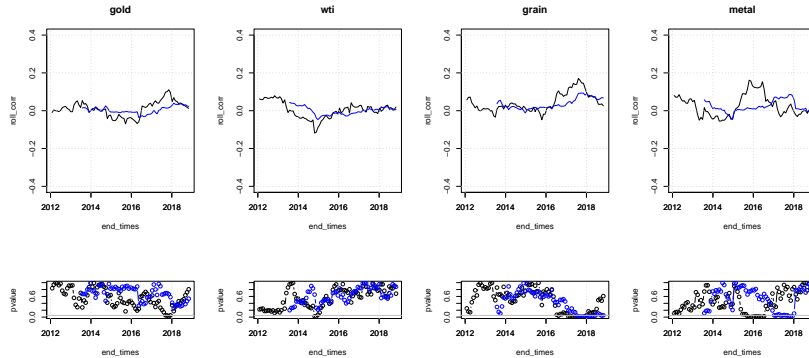
(a) Stocks



(b) Bonds



(c) Currency exchange



(d) Commodities

Figure 2.1: Plots of rolling correlation for the different asset classes (on top) and significance for each value (on bottom). Blue lines are the 3-year rolling correlations, while the black ones have a window of 18 months. Both computations are updated monthly.

There are two graphs for each asset: in the top plots levels of the rolling correlations are represented using two different colours, blue for the 3-year and black for the 18-month windows; in the bottom plots we included the significance of each rolling correlation through its p-value. The grey horizontal line represents the 5% level of significance¹.

The main conclusion we can draw from these images is that the correlation of any asset with Bitcoin is hardly ever significantly different from zero, and when it is, its absolute level is never more than greater than 20% for a small period of time.

To confirm the fact that Bitcoin is not correlated with any asset, we can also take a look at the path of the rolling correlations: there is no line that is always above zero, nor below. This indicates that there is no underlying trend, whether positive or negative, and the correlation one might find is only temporary.

**** maybe add more comments on the results ****

¹As we explained in the previous section, to check that a sample correlation is *significantly* non zero, we compare the p-value of the test to a given level, here $1 - \alpha = 5\%$. Graphically, whenever the dots are above the grey line in Figure 2.1, the corresponding correlation is *not* significantly different from zero.

Chapter 3

Presentation of the Models

In this chapter we will present the stochastic frameworks in which we developed our analysis. We first introduce a *jump diffusion* (JD) model presented in 1976 by R.C. Merton: he added log-normal jumps to the simple B&S dynamics of the asset price. Then we move to the *stochastic volatility* (SV) model of Heston 1993. Heston introduced a new stochastic process that accounts for the variance of the underlying which evolves as a B&S with a stochastic volatility term. The last model we will present was introduced by Bates in 1996 and it is the combination of the former two: an asset dynamics which include jumps and is driven by a stochastic volatility. All models are first introduced in the one dimensional case and then generalised to the n dimensional case which was then implemented in our code.

3.1 Preliminary Notions

In this section we will briefly present the equation of a geometric brownian motion, introduce the notion of Poisson process and present the CIR process. All of these building blocks will be required to fully understand the models to follow.

3.1.1 Geometric Brownian Motion

The simplest continuous dynamics to describe the price of an asset is that of a geometric brownian motion:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \quad (3.1)$$

where S_t represents the price of the asset at time t , μ is the (constant) drift and σ is the (constant) volatility. W_t is a Wiener process. This is the standard and most widespread stochastic differential equation to model asset dynamics, so we will only present those results that will be later used in our study.

Applying Ito's lemma to the previous equation, we can also explicitly express the dynamics of the log-returns $X_t = \log(S_t)$, obtaining:

$$dX_t = (\mu - \frac{\sigma^2}{2})dt + \sigma dW_t \quad (3.2)$$

This stochastic differential has a simple solution which can be computed via stochastic integrals and allows us to describe the dynamics of the log-returns at each instant t starting from $t = 0$:

$$X_t = X_0 + (\mu - \frac{\sigma^2}{2})t + \sigma W_t \quad (3.3)$$

Thanks to (3.3) we can now express the price dynamics of the asset by inverting the relation with the returns: $S_t = e^{X_t}$. We thus obtain the solution to (3.1):

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t} \quad (3.4)$$

Given that $S_t = e^{X_t}$, that X_t by its equation is a generalized brownian motion and hence we have $X_t - X_0 \sim \mathcal{N}(\mu - \frac{\sigma^2}{2}, \sigma^2)$, the resulting distribution of prices at time t as units of the initial value is distributed as a log-normal.

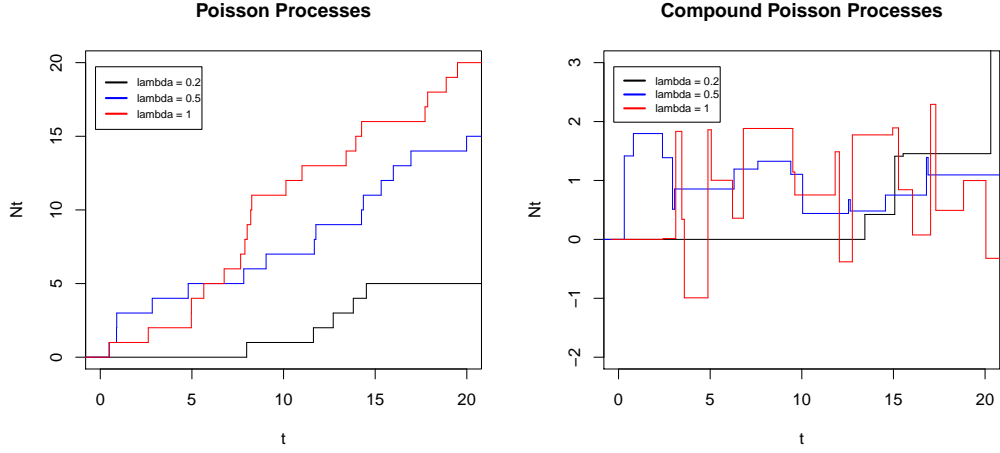
The great success of these framework comes from the simplicity of its dynamics. In particular, since the log-returns follow a Gaussian distribution, μ and σ are easy to calibrate from data and the formulas for pricing options are often explicit. As one can imagine, a simple model can only explain simple phenomena: that's why we have such a great deal of *new and improved* version of equation 3.1.

3.1.2 Poisson Process and Compound Poisson Process

Consider a sequence of *independent* exponential random variables $\{\tau_i\}_{i \geq 1}$ with parameter λ^1 and let $T_n = \sum_{i=1}^n \tau_i$. Then we can define the *Poisson process* N_t as

$$N_t = \sum_{n \geq 1} \mathbb{1}_{t \geq T_n} \quad (3.5)$$

¹An exponential random variable τ of parameter λ has a cumulative distribution function of the form: $\mathbb{P}(\tau \geq y) = e^{-\lambda y}$



(a) Poisson processes.

(b) Compound with Gaussian jumps.

Figure 3.1: Poisson processes and compound Poisson processes with different λ .

where $\mathbb{1}_{condition}$ is 1 if the *condition* is true, 0 if it is false.

N_t is thus a piece-wise constant RCLL² process with jumps that happen at times T_n and are all of size 1, as we can see from Figure fig. 3.1a. An important property of Poisson processes is that they have independent and stationary increments, meaning that the increment of $N_t - N_s$ (with $s \leq t$) is independent from the history of the process up to N_s and has the same law of N_{t-s} . At any time t , N_t is distributed as a Poisson of parameter λt , which means it is a discrete random variable on the integer set with

$$\mathbb{P}(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \quad (3.6)$$

When working with jump diffusion process, it is often the case that there is no explicit formula for its density, thus usually we resort to characteristic functions. The characteristic function of N_t is given by

$$\phi_{N_t}(u) = e^{\lambda t(e^{iu} - 1)} \quad (3.7)$$

The computations to get 3.7 from 3.6 are carried out in Appendix A.

For financial applications, it is of little interest to have a process with a single possible jump size. The *compound* Poisson processes are a generalization of Poisson processes where the jump sizes can have an arbitrary

²RCLL is shorthand for right continuous with left limit.

distribution. More precisely, consider a Poisson process N_t with parameter λ and a sequence of i.i.d³ variables $Y_{i \geq 1}$ with law $f_Y(y)$. Then the process defined by

$$X_t = \sum_{n=1}^{N_t} Y_n \quad (3.8)$$

is a compound Poisson process. Examples of this kind of process are plotted in Figure fig. 3.1b.

As before, we developed the computations to obtain an expression for the characteristic function of X_t in Appendix A. The resulting expression depends on the distribution of Y , specifically from its characteristic function $\phi_Y(u)$:

$$\phi_{X_t}(u) = e^{\lambda t(\phi_Y(u)-1)} \quad (3.9)$$

Both in Merton's and in the Heston's models there will be a jump component driven by a compound Poisson process with Gaussian jump sizes, as we will see in the following paragraphs.

3.1.3 CIR Process

The CIR process was introduced in [3] in 1985 by Cox, Ingersoll and Ross (hence the name CIR) as a generalization of a Vasicek process to model the mean reverting dynamics of interest rates. Following their notation, the differential equation for the evolution of the rate is given by:

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t \quad (3.10)$$

where we have three parameters that characterize it: θ is the *long-term value* of the rate, the asymptotic level which it tends to settle at in the long run; κ is the *mean-reversion rate*, the speed at which the rate is pulled back to the θ value; finally σ accounts for the *volatility* of the stochastic component. When $\kappa, \theta > 0$, equation (3.10) represents a first order mean-reverting autoregressive process. Moreover, thanks to *******add citation*******, we know that the process will not hit zero if the following condition is satisfied:

$$2\kappa\theta > \sigma^2 \quad (3.11)$$

This condition is usually referred to as *Feller* condition, from the author of the cited paper in which this result was first presented.

An example of a CIR process is shown in Figure (3.2), where the mean-reverting effect is clearly visible.

³i.i.d stands for independent and identically distributed.

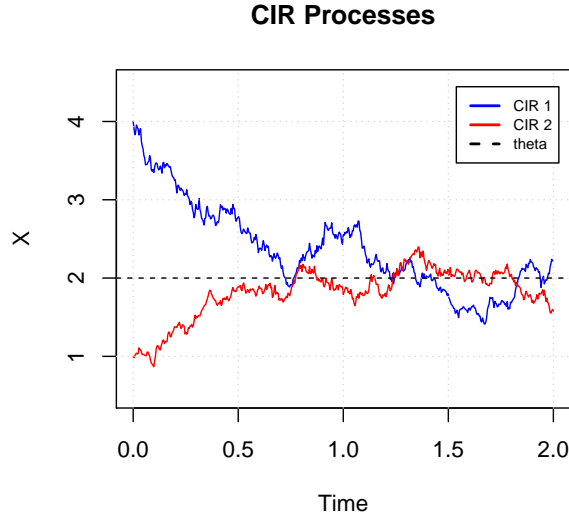


Figure 3.2: Two trajectories of the same CIR process: $dr_t = 3(2 - r_t)dt + 0.5\sqrt{r_t}dW_t$ with different starting point. We can see the mean-reverting effect that attracts both trajectories to the value $\theta = 2$.

When modelling interest rate in continuous time, having r_t hit zero is not an issue since when $r_t = 0$ equation (3.10) reduces to $dr_t = \kappa\theta dt$ which immediately brings the level back to positive values and hence the square root in the dynamics never loses meaning. Conversely, considering a *discretized* version of (3.10), as is the case in a simulation framework, one needs to pay attention on how he models the increments since a simple Euler discretization scheme may cause r_t to reach *negative* values⁴ and invalidate the whole model representation.

The non negativity of the CIR process will become fundamental when we introduce Heston and Bates SV models, where the (stochastic) variance of the process driving the asset price will be model as a CIR process.

*****AGGIUNGERE DISTRIBUZIONE DI r_t e di r_{∞} (gamma)*****

3.2 Merton Model

*****maybe add some words right here *****

⁴This model was introduced having in mind the certainty that interest rates could never be negative, hence the introduction of a square root in the dynamics. Given recent years interest rates levels, this is no more the case.

3.2.1 Original Univariate Model

The first jump diffusion model was originally introduced in [8] in order to account for the leptokurtic distribution of real market returns and to model sudden falls (or rises) in prices due to the arrival of new information. The asset price dynamics S_t is modelled as a GBM to which a jump component driven by a compound Poisson process is added:

$$\frac{dS_t}{S_t} = \alpha dt + \sigma dW_t + (Y_t - 1)dN \quad (3.12)$$

where α and σ are respectively the drift and the diffusion of the continuous part, Y_t is a process modelling the intensity of the jumps and $N(t)$ is the Poisson process driving the arrival of the jumps and has parameter λ .

We can rewrite (3.12) in terms of the log-returns $X_t = \log(S_t)$ and obtain, following the computations in [7] and using theory from [9]:

$$dX_t = \left(\alpha - \frac{\sigma^2}{2}\right)dt + \sigma dW_t + \log(Y_t) \quad (3.13)$$

that has as solution:

$$X_t = X_0 + \mu t + \sigma W_t + \sum_{k=1}^{N(t)} \eta_k \quad (3.14)$$

where X_0 is the initial value of the log-returns, $\eta_k = \log(Y_k) = \log(Y_{t_k})$ and t_k is the time when the k^{th} Poisson shock from $N(t)$ happens. We use $\mu = \alpha - \frac{\sigma^2}{2}$ for ease of notation throughout the paper. Following [8], we take η_k *i.i.d.* (independent and identically distributed) and Gaussian, in particular $\eta \sim \mathcal{N}(\theta, \delta^2)$. Another choice for the distribution of η is given in [6].

It is often useful when dealing with market data that are by nature discrete, to consider a *discretized* version of (3.14) in which the values are sampled at intervals of Δt in $[0, T]$. We thus get that for $X_i = \log(\frac{S_{i+1}}{S_i})$:

$$X_i = \mu \Delta t + \sigma \sqrt{\Delta t} z + \sum_{k=1}^{N_{i+1} - N_i} Y_k \quad (3.15)$$

where we denote $X_i = X_{t_i}$, $N_i = N(t_i)$ and $t_i = i\Delta t$ with $i = 0 \dots N$, $t_N = N\Delta t = T$, z is distributed as a standard Gaussian $z \sim \mathcal{N}(0, 1)$.

The Poisson process $N(t)$ in (3.15) is computed at times t_{i+1} and t_i and these quantities are subtracted. Following basic stochastic analysis, one can prove that the resulting value $N_{i+1} - N_i$, is distributed as a Poisson random

variable N of parameter $\lambda\Delta t$. This allows us to provide an explicit formulation for the transition density of the returns using the theorem of total probability:

$$f_{\Delta X}(x) = \sum_{k=0}^{\infty} \mathbb{P}(N = k) f_{\Delta X|N=k}(x) \quad (3.16)$$

This is an infinite mixture of Gaussian distributions, due to the infinite possible realization of the Poisson variable, and renders the estimation of the model through MLE technique intractable, see [5]. To solve this problem we introduce a first order approximation, as it's been proposed in [2]. Considering small Δt , so that also $\lambda\Delta t$ is small, we obtain that the only relevant terms in (3.16) are the one for $k = 0, 1$. The formula for the transition density becomes:

$$f_{\Delta X}(x) = \mathbb{P}(N = 0) f_{\Delta X|N=0}(x) + \mathbb{P}(N = 1) f_{\Delta X|N=1}(x)$$

expressing it explicitly:

$$f_{\Delta X}(x) = (1 - \lambda\Delta t) f_{\mathcal{N}}(x; \mu, \sigma^2) + (\lambda\Delta t) f_{\mathcal{N}}(x; \mu + \theta, \sigma^2 + \delta^2) \quad (3.17)$$

where $f_{\mathcal{N}}(x; \mu, \sigma^2)$ is the density of a Gaussian with parameters $\mathcal{N}(\mu, \sigma^2)$.

3.2.2 Multivariate Model

Starting from the univariate model introduced in [8], we developed a generalization to n assets including only idiosyncratic jumps:

$$\frac{dS_t^{(j)}}{S_t^{(j)}} = \alpha_j dt + \sigma_j dW_t^{(j)} + (Y_t^{(j)} - 1) dN_t^{(j)} \quad (3.18)$$

where \mathbf{S}_t are the prices of the assets, $j = 1 \dots n$ represents the asset, α_j are the drifts, σ_j are the diffusion coefficients, $W_t^{(j)}$ are the components of an n -dimensional Wiener process \mathbf{W}_t with $dW^{(j)} dW^{(i)} = \rho_{j,i}$, η_j represent the intensities of the jumps and are distributed as Gaussian: $\eta_j \sim \mathcal{N}(\theta_j, \delta_j^2)$. Finally, $N^{(j)}(t)$ are Poisson processes with parameters λ_j , which are independent of \mathbf{W}_t and of one another.

In order to calibrate the parameters to the value of the market log-returns, we used a Maximum Likelihood approach. We thus maximize:

$$\mathcal{L}(\psi | \Delta \mathbf{x}_{t_1}, \Delta \mathbf{x}_{t_2}, \dots, \Delta \mathbf{x}_{t_N}) = \sum_{i=1}^N f_{\Delta \mathbf{x}}(\Delta \mathbf{x}_{t_i} | \psi) \quad (3.19)$$

where $\psi = \{\{\mu_j\}, \{\sigma_j\}, \{\rho_{i,j}\}, \{\theta_j\}, \{\delta\}_j, \{\lambda_j\}\}$ are the model parameters, $f_{\Delta \mathbf{x}}$ is the transitional density of the log-returns which is computed approximately using the theorem of total probability. For a full insight on the model and the calibration procedure, please refer to the [*APPENDIX LINK*](#)

3.3 Heston Model

The Heston model was presented in 1993 in [4] as a new framework to model stochastic volatility in the asset price dynamics, which allows to better fit the skewness and kurtosis of the log-return distribution.

3.3.1 Univariate Heston Model

The Heston model belongs to the family of SD processes, generalizations of B&S model in which the volatility is no more constant but is itself stochastic.

add examples?

Starting with a GBM as in the B&S framework, we obtain the dynamics of the price process simply by allowing the volatility in (3.1) to evolve over time and specifying how this evolution takes place. In particular, the dynamics of the *instantaneous* variance $V_t = \sigma_t^2$ for the Heston model is described by a CIR process :

$$\frac{dS_t}{S_t} = \mu dt + \sqrt{V_t} dW_t^S \quad (3.20)$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma_V \sqrt{V_t} dW_t^V \quad (3.21)$$

where μ represents the drift in the asset prices, $\kappa > 0$ and $\theta > 0$ are respectively the *mean-reversion* rate and the long-run level for the V_t process, $\sigma_V > 0$ is often referred to as the *volatility of volatility* parameter, in short vol-of-vol. The two brownian motions W_t^S and W_t^V are correlated with correlation coefficient equal to ρ . The variance process is always strictly positive if the Feller condition $2\kappa\theta > \sigma_V$ is satisfied.

Let us now consider the dynamics of the log-return $x_t = \log(S_t)$, as we did in the Merton case. Unfortunately, given the increased complexity of the model due to the SV part, an explicit formula for the density of the log-return is not available and it has to be computed from the Fourier inversion of the characteristic function:

$$f_{x_t}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi_{x_t}(iu) e^{iux} du \quad (3.22)$$

In (3.22) we omitted the dependence of $f_{x_t}(x)$ and $\phi_{x_t}(u)$ on model parameters and on the initial values of the log-returns x_0 and of the variance process V_0 .

To derive the expression of the characteristic function $\phi_{x_t}(u)$, one has to solve a couple of *Fokker-Planck* partial differential equation as is shown in the Appendix of the reference paper by Heston [4]. This procedure is beyond the scope of this thesis and thus we will only report the final result, which is a log-affine equation on the information at time $t = 0$:

$$\phi_{x_t}(u|x_0, V_0) = \exp\{A(t, u) + B(t, u)x_0 + C(t, u)V_0\}$$

$$A(t, u) = \mu u i t + \frac{\kappa \theta}{\sigma_V^2} \left((\kappa - \rho \sigma_V u i + d)t - 2 \log \left[\frac{1 - g e^{dt}}{1 - g} \right] \right) \quad (3.23a)$$

$$B(t, u) = i u \quad (3.23b)$$

$$C(t, u) = \frac{\kappa - \rho \sigma_V u i + d}{\sigma_V^2} \left[\frac{1 - e^{dt}}{1 + g e^{dt}} \right] \quad (3.23c)$$

where :

$$d = \sqrt{(\rho \sigma_V u i - \kappa)^2 + \sigma_V^2 (u i + u^2)}$$

$$g = \frac{\kappa - \rho \sigma_V u i + d}{\kappa - \rho \sigma_V u i - d}$$

This formulation is the one proposed by Heston in his original paper [4] but it is shown in [1] that it has numerical issues when pricing Vanilla options using Fourier methods. We will not be pricing any instrument, however we may incur in the same errors when calibrating our model through the techniques explained in Chapter 4. For this reason, we will be using the alternative formulation presented in [1]:

$$\phi_{x_t}^*(u|x_0, V_0) = \exp\{A(t, u) + B(t, u)x_0 + C(t, u)V_0\}$$

$$A(t, u) = \mu u i t + \frac{\kappa \theta}{\sigma_V^2} \left((\kappa - \rho \sigma_V u i - d)t - 2 \log \left[\frac{1 - g^* e^{-dt}}{1 - g^*} \right] \right) \quad (3.24)$$

$$B(t, u) = i u$$

$$C(t, u) = \frac{\kappa - \rho \sigma_V u i - d}{\sigma_V^2} \left[\frac{1 - e^{-dt}}{1 - g^* e^{-dt}} \right]$$

where :

$$d = \sqrt{(\rho\sigma_V ui - \kappa)^2 + \sigma_V^2(ui + u^2)}$$

$$g^* = \frac{\kappa - \rho\sigma_V ui - d}{\kappa - \rho\sigma_V ui + d} = \frac{1}{g}$$

The only difference is that the signs of the d terms are all flipped: the origin of the two representations for the characteristic function lies in the fact that the complex root d has two possible values and the second value is exactly minus the first value. A uniform choice between the two possibilities cannot be made over all the complex plane to obtain a continuous function due to the presence of a branch cut in the graph of \sqrt{z} . Selecting the sign of d as in $\phi_{x_t}^*$ and g^* avoids the discontinuity affecting our future computations.

Our characteristic function still depends on the initial values x_0 and V_0 and is thus often called *conditional* characteristic function on x_0 and V_0 . We can easily remove the dependence on x_0 by considering the distribution of incremental returns, namely $\Delta x_t = \log(S_t/S_0) = x_t - x_0$

Applying the definition of characteristic function:

$$\begin{aligned}\phi_{\Delta x_t}(u) &= \mathbb{E}[e^{iu\Delta x_t}] \\ &= \mathbb{E}[e^{iu(x_t - x_0)}] \\ &= \mathbb{E}[e^{iux_t}]e^{-iux_0} \\ &= \phi_{x_t}(u)e^{-iux_0} \\ &= \exp\{A(t, u) + (B(t, u) - iu)x_0 + C(t, u)V_0\}\end{aligned}$$

and since $B(t, u) = iu$, the second term in the exponential is equal to zero and we are left with:

$$\phi_{\Delta x_t}(u) = \exp\{A(t, u) + C(t, u)V_0\} \quad (3.25)$$

This equation is however still dependent on the initial value of the variance process. In a simulation framework, this would not be an issue, since we can define the level of V_0 ourselves and then generate all the different scenarios. However, if we need to calibrate the model parameters from asset prices, the market data for the variance process are not available. We will address more in depth different ways to solve this problem later on in Chapter 4. As for now, we show that we can obtain an *unconditional* expression for the characteristic function of a Heston process by approximating the distribution of V_t with its long-run distribution.

***** add unconditional chf *****

3.4 Bates Model

Chapter 4

Calibration of the Models

In this chapter we will explain how the different models were calibrated and what difficulties were overcome. Empirical results are included for each section.

Chapter 5

Markowitz Portfolio Optimization

Chapter 6

Conclusions

Appendix A

Characteristic Function of a Compound Poisson Process

By definition, the characteristic function of a random variable X is given by:

$$\phi_X(u) = \mathbb{E}[e^{iuX}] \quad (\text{A.1})$$

Let's first consider a simple Poisson process N_t of parameter λ . At each time $t > 0$, N_t has a discrete distribution that follows:

$$\mathbb{P}(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \dots \quad (\text{A.2})$$

Since N_t is a *discrete* random variable, the expectation in A.1 amounts to a sum over all the possible values of N_t :

$$\begin{aligned} \phi_{N_t}(u) &= \mathbb{E}[e^{iuN_t}] = \sum_{n=0}^{\infty} e^{iun} \mathbb{P}(N_t = n) \\ &= \sum_{n=0}^{\infty} e^{iun} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \\ &= e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t e^{iu})^n}{n!} \end{aligned}$$

Using the definition of exponential $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ we then get the final result:

$$\begin{aligned} \phi_{N_t}(u) &= e^{-\lambda t} e^{\lambda t e^{iu}} \\ &= e^{\lambda t(e^{iu} - 1)} \end{aligned}$$

Let's consider now a compound Poisson process defined by

$$X_t = \sum_{i=1}^{N_t} Y_i \quad (\text{A.3})$$

where Y_i are i.i.d. and have density expressed by the function $f_Y(y)$.

To compute the characteristic function of X_t we can follow the same steps as in the simple Poisson case, but in addition we use the theorem of total expectation to first simplify the expression and then proceed exploiting the i.i.d property:

$$\begin{aligned} \phi_{X_t}(u) &= \mathbb{E}[e^{iuX_t}] \\ &= \sum_{n=0}^{\infty} \mathbb{E}[e^{iuX_t} | N_t = n] \mathbb{P}(N_t = n) \\ &= \sum_{n=0}^{\infty} \mathbb{E}\left[\prod_{i=1}^{N_t} e^{iuY_i} | N_t = n\right] \mathbb{P}(N_t = n) \\ &= \sum_{n=0}^{\infty} \prod_{i=1}^n \mathbb{E}[e^{iuY_i} | N_t = n] \mathbb{P}(N_t = n) \\ &= \sum_{n=0}^{\infty} (\mathbb{E}[e^{iuY}])^n \mathbb{P}(N_t = n) \\ &= \sum_{n=0}^{\infty} (\phi_Y(u))^n e^{-\lambda t} \frac{(\lambda t)^n}{n!} \\ &= e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t \phi_Y(u))^n}{n!} \\ &= e^{\lambda t(\phi_Y(u)-1)} \end{aligned}$$

Bibliography

- [1] ALBRECHER, H., MAYER, P., SCHOUTENS, W., AND TISTAERT, J. The little heston trap. *Wilmott* (03 2007), 83–92.
- [2] BALL, C. A., AND TOROUS, W. N. A simplified jump process for common stock returns. *Journal of Financial and Quantitative Analysis* 18, 01 (1983), 53–65.
- [3] COX, J. C., INGERSOLL, J. E., AND ROSS, S. A. A theory of the term structure of interest rates. *Econometrica* 53, 2 (1985), 385–407.
- [4] HESTON, S. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* 6 (02 1993), 327–43.
- [5] HONORÉ, P. Pitfalls in estimating jump-diffusion models. *SSRN Electronic Journal* (01 1998).
- [6] KOU, S. G. A jump-diffusion model for option pricing. *Management Science* 48, 8 (2002), 1086–1101.
- [7] MARTIN, M. A two-asset jump diffusion model with correlation. Master’s thesis, University of Oxford, 2007.
- [8] MERTON, R. Option prices when underlying stock returns are discontinuous. *Journal of Financial Economics* 3 (01 1976), 125–144.
- [9] TANKOV, P., AND CONT, R. *Financial Modelling with Jump Processes, Second Edition*. Chapman and Hall/CRC Financial Mathematics Series. Taylor & Francis, 2015.