École polytechnique

Third-year Research Project

Statistical Modeling

Authors:
Imane Farhat
Hussein Fellahi

Supervisor:
Mathieu Rosenbaum

April 5, 2020

Contents

1 Preliminary analysis of the financial returns 2
1.1 Validity of the Black Scholes model . . . 2
1.2 Stationarity of the returns . . . 3
1.2.1 Definition . . . 3
1.2.2 Verification of the stationarity of the returns . . . 4
1.3 Independence default and volatility clustering . . . 4
1.4 Fat tails . . . 4

2 Time Series Analysis - ARIMA Models 5
2.1 Definition . . . 5
2.2 Deriving optimal parameters of the ARIMA . . . 5
2.2.1 The order of differenciation (d) . . . 5
2.2.2 The order of the moving average (q) . . . 6
2.2.3 The order of autocorrelation (p) . . . 6
2.3 The ARIMA Model . . . 7

3 GARCH models 7
3.1 Definition . . . 8
3.2 Deriving optimal parameters of the GARCH . . . 8
3.2.1 Finding optimal p and q parameters . . . 8
3.2.2 Calibrating the volatility . . . 8
3.3 Empirical study of GARCH model . . . 8
3.3.1 GARCH(1,1) with Gaussian white noise . . . 8
3.3.2 More complex structure with Laplace white noise . . . 9

4 Volatility clustering modeling with Hawkes processes 11
4.1 Context of the study . . . 11
4.2 Homogeneous Poisson Process modeling . . . 12
4.3 Hawkes Process modeling . . . 13
4.3.1 Theoretical framework . . . 13
4.3.2 Results . . . 14

Introduction

The aim of this project is to use some statistical modeling tools to a financial dataset. The data used is the daily open prices of the Bitcoin from 01/01/2018 to 17/02/2020. The data can be found in the archive below:
https://drive.google.com/file/d/1Yo7KJCpeuiQs-IpEvhIz5X8Yf-5ZGHL6/view?usp=sharing

We first analyzed some preliminary properties of the Bitcoin price time series such as the normality and stationarity of the returns. Then we implemented an ARIMA model as a first attempt to forecast the prices of Bitcoin. Following this attempt, we moved to GARCH models in order to improve the modelling of the time series, as GARCH process are known to be more suited for this kind of financial data. Finally, we studied the volatility clustering effect by studying the arrival times of large changes in returns, and modeling them first as Poisson process and then a Hawkes process and calibrating it with a maximum likelihood estimation.

1 Preliminary analysis of the financial returns

We observe the Bitcoin price time series, and in this section we focus on the returns of this asset $(r_t)_t$ defined as $r_t = \log \frac{p_t}{p_{t-1}}$ where p_t is the daily price of the Bitcoin.

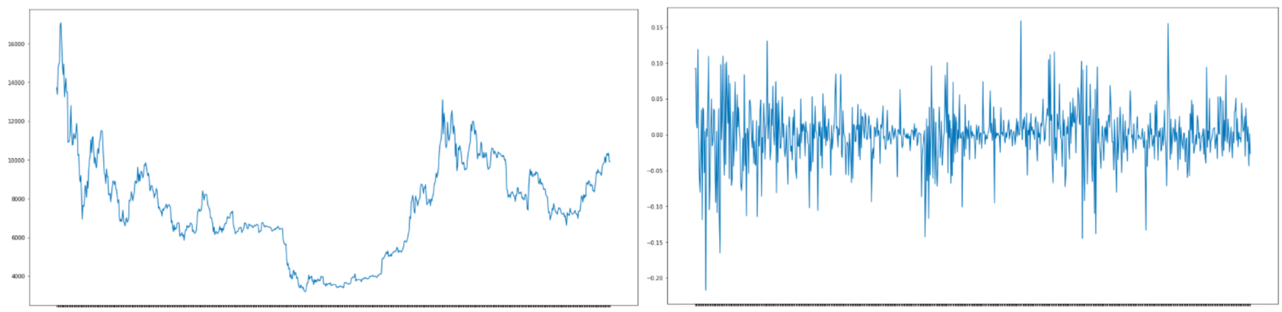


Figure 1: Bitcoin price over the time (left) and Bitcoin returns (right)

1.1 Validity of the Black Scholes model

In the Black Scholes model, the returns (r_t) are i.i.d and follow a Gaussian distribution $\mathcal{N}(m, s^2)$. We can verify if this property holds by first plotting the histogram of the returns of the Bitcoin prices and comparing this histogram to that of a normal Gaussian distribution with the same mean and standard deviation.

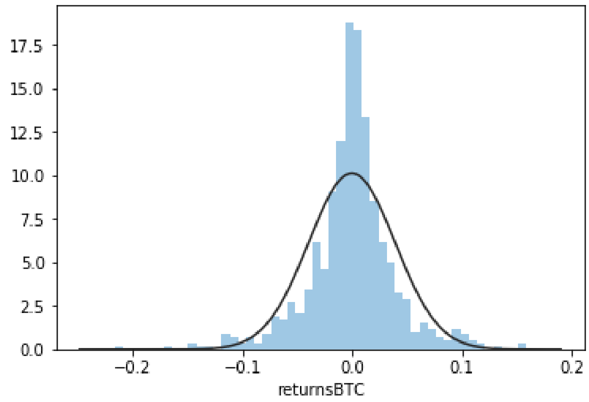


Figure 2: Fitting a Gaussian distribution

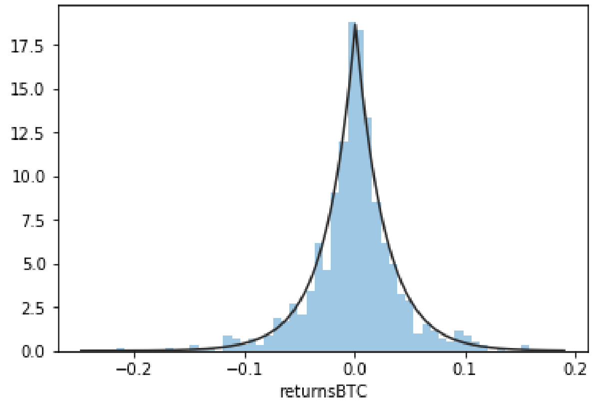


Figure 3: Fitting a Laplace distribution

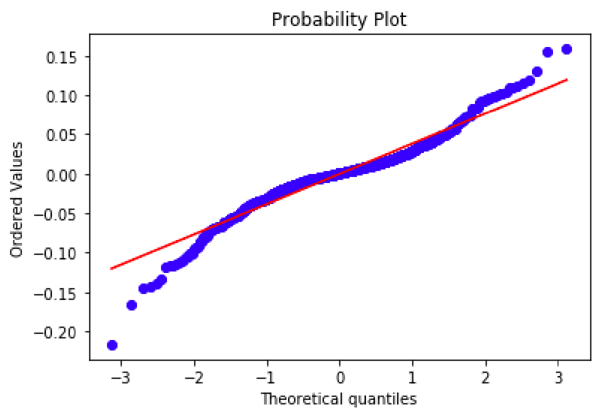


Figure 4: Quantile-Quantile plot against a normal distribution

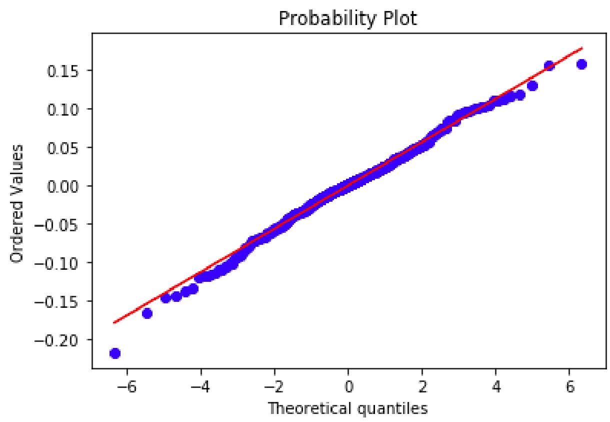


Figure 5: Quantile-Quantile plot against a Laplace distribution

In this case, it seems clear that the returns of the Bitcoin time series don't follow a Gaussian distribution as stipulated in the Black Scholes model, and rather follow a Laplace distribution.

1.2 Stationarity of the returns

A good model to describe the returns of a financial time series should incorporate the stationarity property of the returns. Indeed, contrary to (p_t), (r_t) is usually stationary at least in the second order. This property generalizes the i.i.d hypothesis and is indispensable for estimation and prediction.

1.2.1 Definition

A process $(X_t)_{t \in N}$ is called strictly stationary if $\forall k, \forall h$: $(X_1, ..., X_k)$ and $(X_{1+h}, ..., X_{k+h})$ have the same distribution.

A process $(X_t)_{t \in N}$ is second-order stationary if:

$$\forall t \in N, E[X_t^2] < +\infty$$

$$\forall t \in N, E[X_t] = m$$

$$\forall t, h, Cov(X_t, X_{t+h}) = \gamma(h)$$

where γ is the autocovariance function.

1.2.2 Verification of the stationarity of the returns

To verify the stationarity of the returns (r_t), we apply an augmented Dickey-Fuller test. It tests the null hypothesis that a unit root is present in the time series and it is therefore non-stationary ad has a time-dependent structure. When applied to the returns time series, the ADF test returns a p-value of 0.00, meaning that (r_t) is indeed stationary.

1.3 Independence default and volatility clustering

The time series of the returns (r_t) is usually close to a white noise. However, we tend to have a significant autocorrelation of (r_t^2), which is incompatible with an i.i.d white noise.

We compute the Autocorrelation Plot of the returns time series (r_t), as well as that of the square of returns (r_t^2) with regards to lags that range from 0 to 30.

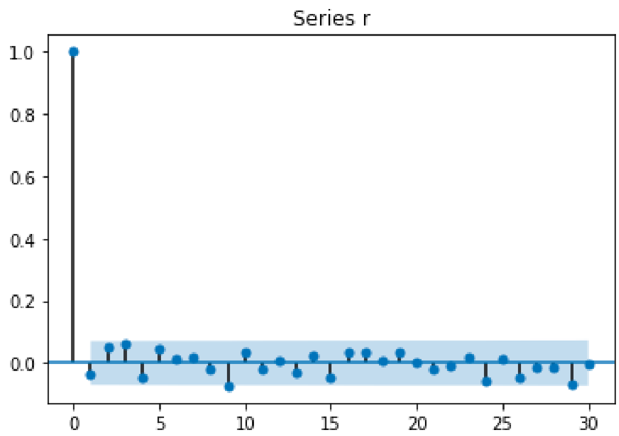


Figure 6: Autocorrelation of (r_t)

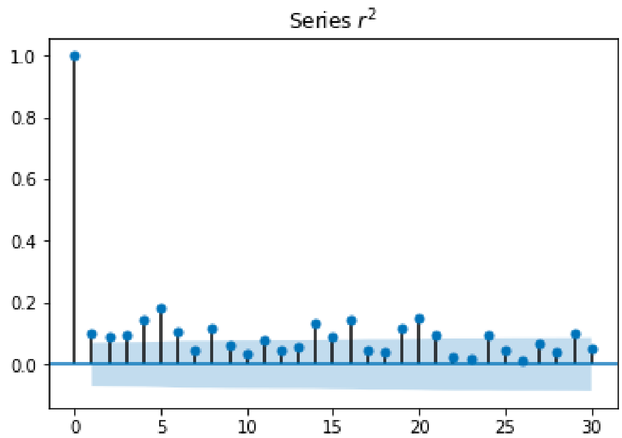


Figure 7: Autocorrelation of (r_t^2)

This autocorrelation of (r_t^2) means that large values of (r_t^2) will likely be followed by more large values of (r_t^2), and thus creating a volatility clustering phenomenon. This autocorrelation of (r_t^2) is not incompatible with stationary returns and in particular a constant variance for the returns. However, the variance of (r_t) conditionally to $(r_{t-1}, r_{t-2}, ...)$ seems to be non-constant. We recapture the conditional heteroscedasticity and inconditional homoscedasticity phenomenon.

1.4 Fat tails

As shown by the quantile-quantile plot of the figure (4), the tails of the distribution of the returns (r_t) differ a lot from those of a normal distribution. The tails of the returns seem to be fatter, which means that there are more extreme events than it is predicted by a Gaussian distribution. This is confirmed by an excess of kurtosis of the returns $\gamma_2 = \beta_2 - 3 = 0.325$.

2 Time Series Analysis - ARIMA Models

We will start deepening our analysis with studying how would ARIMA models perform on the series.

2.1 Definition

An ARIMA (Auto Regressive Integrated Moving Average) is a stochastic process depending on time, that defines as follow:

$$(1-L)^d X_t = \mu + \phi_1 (1-L)^d X_{t-1} + ... + \phi_p (1-L)^d X_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - ... - \theta_q \epsilon_{t-q}$$

It has both an Autoregressive part with parameters ϕ_i, a moving average part with ϵ_i a weak white noise and parameters θ_i, as well as a differentiated part of order d with L the lag parameter, aiming to eliminate any non-stationarity of the series.

The objective of this part is therefore to try and calibrate the ARIMA model, i.e. finding the best p, d, q as well as the ϕ_i and θ_j parameters in order to best capture the behaviour of our series. To do so, we fit the model on the data and try to forecast the series on a test set.

2.2 Deriving optimal parameters of the ARIMA

As shown in the previous section, an ARIMA model relies on three main components: Autoregression, Moving Average, Differenciation. In order to provide the most effective modelling via an ARIMA model, we proceed to tune each of the three components, corresponding to the three parameters we can work on, in order to maximize a metric. For this entire section, we will choose to work with the AIC, i.e. we try to find the parameters that maximize the AIC of the model.

2.2.1 The order of differenciation (d)

The first step in our study is to work with the stationarity of thr series. To do so, we rely on the Augmented Dickey-Fuller (ADF) test: the null hypothesis is that the series has a unit root, which is equivalent to non-stationarity in the context of of ARIMA models.

Starting with applying the ADF test on the series of Bitcoin Prices, it appears obvious that the series in not stationary.

```
ADF Statistic: -3.087580
p-value: 0.027486
```

We therefore try to differentiate the series to make it stationary. The first order of differentiation $Y_t = X_t - X_{t-1}$ is equivalent to studying the returns of the series.

We reapply the ADF test on the series of the returns:

```
ADF Statistic: -14.819078
p-value: 0.000000
```

Therefore, the returns are indeed stationary which makes it possible to work with them. To assert this first order of differentiation, we plot the ACF of the two-times differentiated series $Z_t = Y_t - Y_{t-1}$

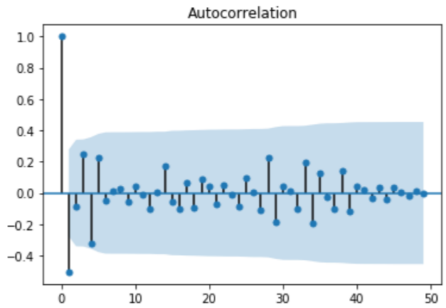


Figure 8: Autocorrelation plot of differentiated returns

This plot tells us that the series might be overdifferentiated, as hinted by the autocorrelations that go in the negative values rather quickly.

2.2.2 The order of the moving average (q)

The next step of study is to find the moving average order, relying on an ACF of the returns series seen in the previous section. Indeed, one can say that the ACF tells how many MA terms are needed to remove any autocorrelation in the stationarized series.

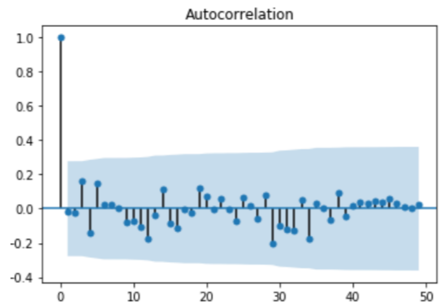


Figure 9: Autocorrelation plot of returns

We see that this plot is not that reliable, as hardly any lags go outside the confidence interval. This parameter may need some tuning, and for the moment, we fix it at 2.

2.2.3 The order of autocorrelation (p)

The third step of study is determining the order of autocorrelation, i.e. figuring the AR part of our model. To do so, we rely on studying the Partial Autocorrelation function (PACF) of our stationarized series, that will show us the most significant lags to takes into account is the analysis of this series.

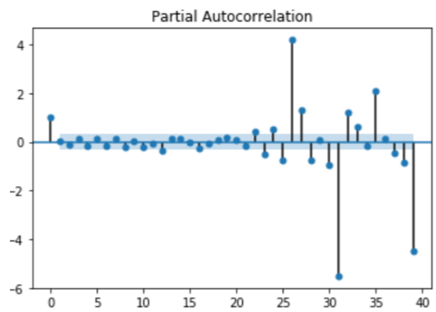


Figure 10: Partial Autocorrelation plot of returns

The PACF plot gives rather unclear results, as the correlations seem to be weak. This means that this parameter may also need further tuning empirically. For the moment, we fix the parameter p to 6 as it appears to be the first order that gives significant correlation (i.e. outside the confidence interval).

2.3 The ARIMA Model

Following the previous section, we start with an ARIMA(6, 1, 2). The behavior of such model is as follow:

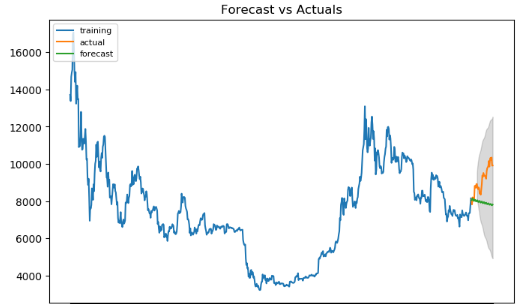


Figure 11: ARIMA(6,1,2)

Yet, we find that this results are not very convincing. Thus, we try to empirically tune the ARIMA model, and find the following setup to be optimal (to do so, we use the *auto.arima*() function of R that provides the parameters that maximize the AIC) : ARIMA(1,1,1)

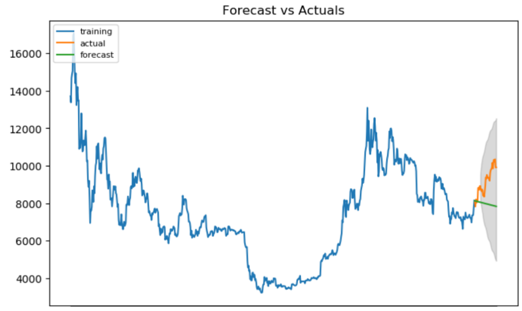


Figure 12: ARIMA(1,1,1)

Therefore, we find that the ARIMA model might have its limits in our study case, given the highly volatile structure of the Bitcoin prices. Another explanation would be that ARMA and ARIMA processes are suitable for modelling non-trivial structures of autocovariance, which is usually not the case in financial data.

In the next section, we will try an approach using GARCH stochastic volatility models, that may find themselves more suitable to our data.

3 GARCH models

Generalized Auto Regressive Conditional Heteroskedasticity models are models that aim to describe the variance of a time series. The basic assumption of such models is that it models the error terms (i.e. the returns in our setup) as a function of the standard deviation and a known stochastic process. If the model of the error variance is autoregressive, we use an ARCH model; if we add a moving average component, it becomes a GARCH model.

3.1 Definition

Let $(\eta_t)_{t\in Z}$ be a sequence of i.i.d random variables with a density such that $E(\eta_t) = 0$ and $Var(\eta_t) = 1$. (ϵ_t) is a GARCH(p,q) process with regards to (η_t) if:

$$\forall t \in Z : \epsilon_t = \sigma_t \eta_t$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

with σ_t^2 measurable with regards to the filtration of $\{\epsilon_u | u < t\}$, and η_t is independent from $\{\epsilon_u | u < t\}$, and $\omega > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$.

3.2 Deriving optimal parameters of the GARCH

Similarly to the previous section on the ARIMA Models, we will try to derive the optimal p and q parameters of the GARCH model using the AIC as a metric. Likewise, we derive the optimal α_i and β_j parameters using the maximum likelihood method.

3.2.1 Finding optimal p and q parameters

We first try to derive the optimal orders of the GARCH model. To do so, we use a method based on the ARMA representation of ϵ_t^2.
After calibrating, we find that the optimal setup for the ϵ_t^2 series is an ARMA(1,1), therefore we go for a GARCH(1,1) model.

3.2.2 Calibrating the volatility

After retrieving the parameters p and q, we are able to fit a GARCH(1,1) model on our data, first in order to get optimal α_1 and β_1 parameters.
Using the R function *garchFit* we find

```
          mu        omega       alpha1        beta1
1.308208e-04 9.780125e-05 7.934812e-02 8.538452e-01
```

Therefore, our model is as follow:

$$\epsilon_t = 1.308208\text{e-}04 + \sigma_t \eta_t$$
$$\sigma_t^2 = 9.780125\text{e-}05 + 7.934812\text{e-}02\ \epsilon_{t-1}^2 + 8.538452\text{e-}01\ \sigma_{t-1}^2$$
$$\eta_t \sim i.i.d.(0,1)$$

3.3 Empirical study of GARCH model

3.3.1 GARCH(1,1) with Gaussian white noise

We try to implement the setup found in the previous section to generate a time series of returns.
First, we use a Gaussian white noise for the η_t series

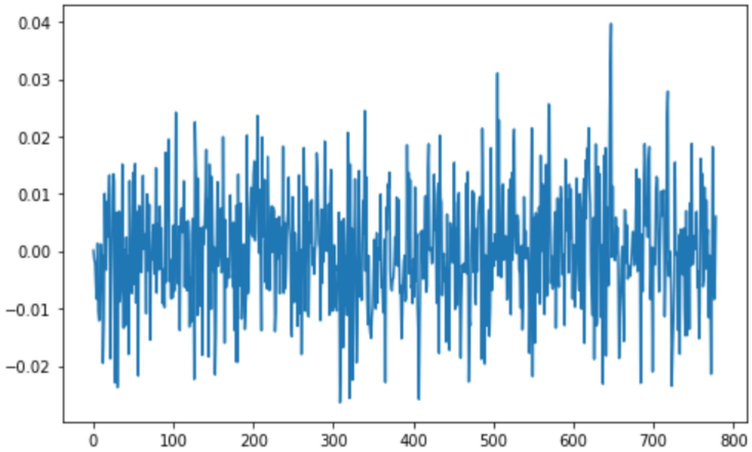
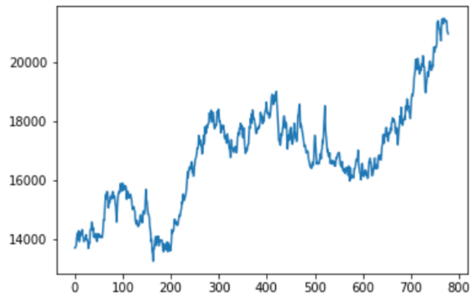


Figure 13: Generated returns using GARCH(1,1) with Gaussian white noise

We then try to generate a series of prices form the returns we just constructed, using the following formula: $p_t = exp(log(p_0) + \sum_{s=1}^{t} \epsilon_s)$

We also do a Monte-Carlo sampling of these series to have a more precise idea of the general behavior of our samples.



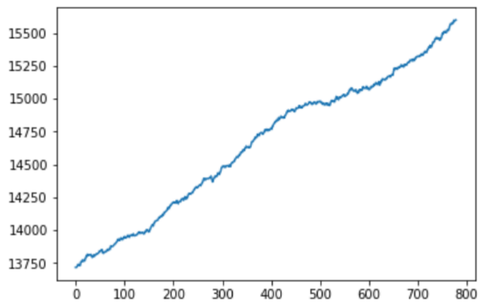


Figure 14: Generated Bitcoin prices using GARCH(1,1) (left) and Monte Carlo sampling of the series (right)

Therefore we see that our series are far from behaving like the original series of Bitcoin prices. Possible explanations of this results may be:

- The returns have a Gaussian distribution and not a Laplace distribution as we empirically observed on our data
- The orders of the GARCH modelling may not be optimal, as we may need a more complex structure of σ_t^2.
- A GARCH(1,1) model supposes a constant autocovariance of the returns, which is a wrong assumption as we've previously seen.

3.3.2 More complex structure with Laplace white noise

We start with generating samples using a Laplace white noise instead of a Gaussian white noise. The distribution of our generated returns thus becomes

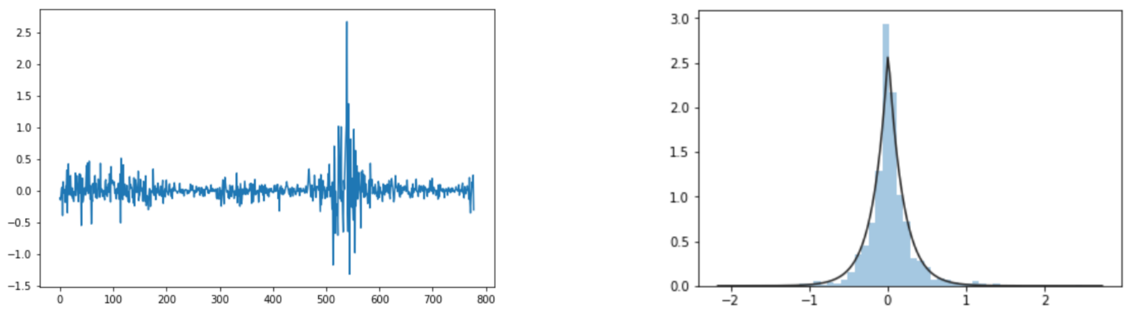


Figure 15: Generated returns prices using GARCH(1,1) and Laplace white noise (left) and distribution of the returns (right)

Yet, this change in the distribution of the returns seems giving wrong results: the Laplace distribution provides more possibilities for rares events, which explains the extreme values that occure a lot more often, as it can be seen on a generated sample of Bitcoin prices and a Monte Carlo simulation of the series

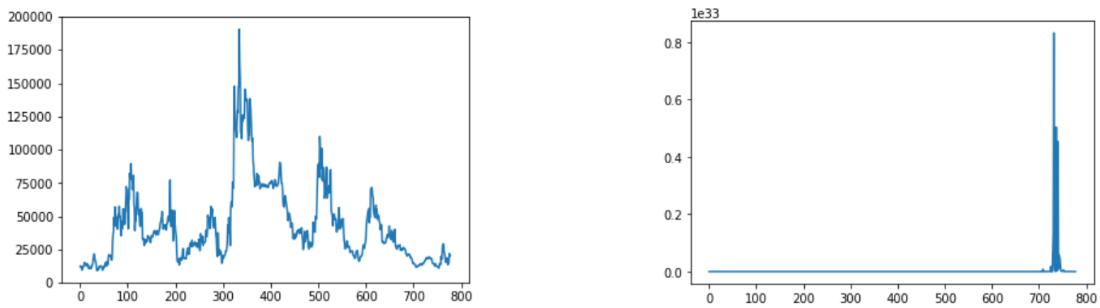


Figure 16: Bitcoin price generated with GARCH(1,1) and Laplace white noise (left) and Monte Carlo simulation of the series (right)

Thus, we choose to stay with a Gaussian white noise for η_t and go for higher orders in the GARCH parameters.

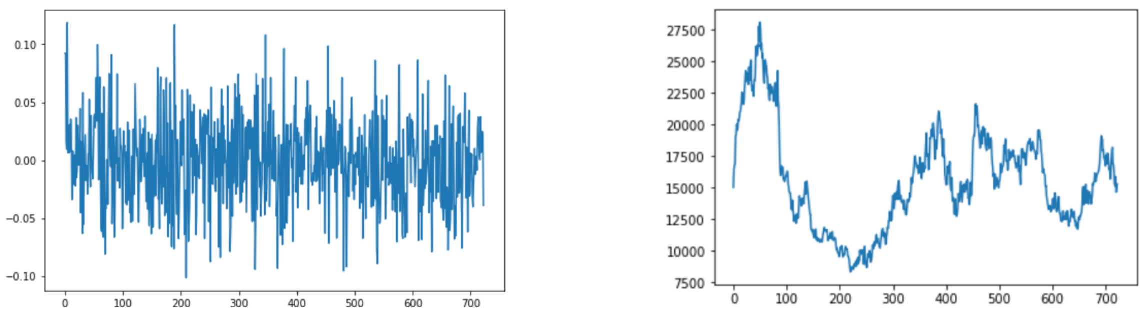


Figure 17: Bitcoin returns generated with GARCH(6,6) and Gaussian white noise (left) and prices genenerated from the returns (right)

Finally, we compare the returns we have generated and the historic returns of the Bitcoin on the same plot

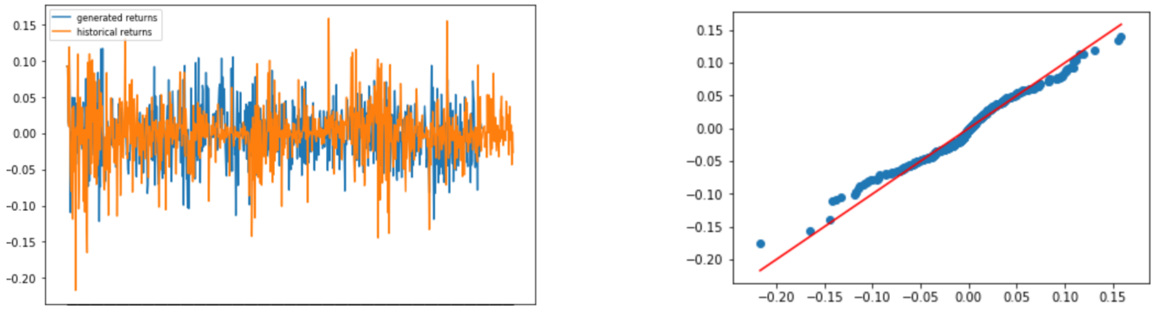


Figure 18: Generated Gaussian returns using GARCH(6,6) vs actual historical returns of the Bitcoin (left) and QQ plot of Generated vs historical returns (x-axis) (right)

4 Volatility clustering modeling with Hawkes processes

The price time series exhibits several autocorrelation properties. First, the frequency of up movements tends to increase with increasing frequency of the past down movements and vice versa. This causes a mean reverting property in the price dynamics. Second, there are also autocorrelations between the movements of the same direction. This causes volatility clustering that is different from the clustering on the macro level, which is typically modeled by GARCH (Bollerslev, 1986) or the stochastic volatility model (Heston, 1993). These properties are well incorporated into the Hawkes model, which belongs to the class of point processes and is introduced by Hawkes (1971). Therefore, there has been an increase in the related work of modeling price dynamics based on the Hawkes process.

4.1 Context of the study

In this section we choose to study the volatility clustering effect and its modeling with a Hawkes process. This effect implies that large changes of volatility are usually followed by larges changes, of either sign. First we define a criterion for a large value of the volatility as a point where the daily return verifies the following condition:

$$|r_t - m| > 2\sigma$$

where m is the historic mean of the returns time series and σ its standard deviation. To justify this choice, we have seen in the sections above that the returns time series is stationary and therefore its mean and standard deviation don't vary with time. Second, the factor of 2 was arbitrary and was chosen so that this definition would give visually points that are in accordance with points that we would consider as large values intuitively. This definition allows us to flag the following points as large volatility points.

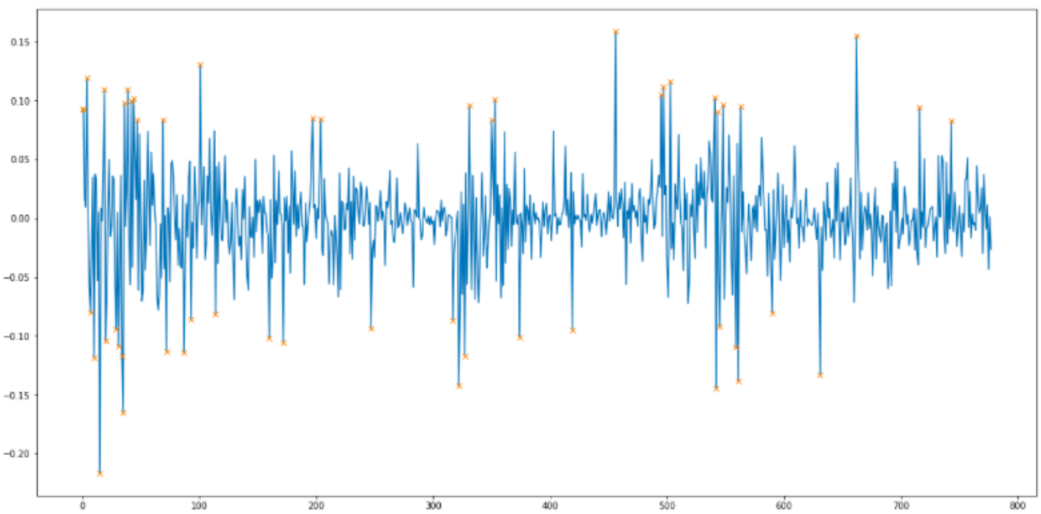


Figure 19: The points marked x correspond to the points of the returns time series that fit the previous criterion (and therefore are considered as large volatility change points)

In the following sections, the process that will be studied concerns the arrival times of these large volatility change points by modeling with a Poisson process and then by a Hawkes process.

4.2 Homogeneous Poisson Process modeling

Let (T_n) be a sequence of random variables taking values in $\bar{R_+}$ such that $T_n < T_{n+1}$. We define the corresponding count process $N_t := \sum_{n \geq 1} 1_{\{T_n \leq t\}}$.

Let (λ_s) be a progressively measurable process such that $\int_0^{T_n} \lambda_s < +\infty, \forall n$, there exists at most one measure P such that $\forall n, N_{t \wedge T_n} - \int_0^{t \wedge T_n} \lambda_s ds$ is a martingale. P is called the law of the point process (T_n) et is characterized by the intensity of the process (λ_s).

In the case of homogeneous Poisson process, the intensity $\lambda_t = \lambda < 0$ is constant.

Let $N(T)$ be a Poisson process with rate λ. Let X_1 be the time of the first arrival. Then,

$$P(X_1 > t) = P(\text{no arrival in}(0,t]) = e^{-\lambda t}$$

Therefore $X_1 \sim Exp(\lambda)$. Let X_2 be the elapsed time between the first and the second arrival.

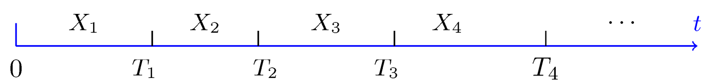


Figure 20: The random variables $X_1, X_2, \cdots$ are called the interarrival times of the counting process N(t)

Let $s > 0$ and $t > 0$. The intervals $(0, s]$ and $(s, s+t]$ are disjoint. Therefore:

$$\begin{aligned}
P(X_2 > t | X_1 = s) &= P(\text{no arrival in } (s, s+t] | X_1 = s) \\
&= P(\text{no arrival in } (s, s+t])(\text{independent increments}) \\
&= e^{-\lambda t}
\end{aligned}$$

Therefore $X_2 \sim Exp(\lambda)$. We can conclude by induction that the interarrival times $X_1, X_2, ...$ are independent and follow an exponential distribution $Exp(\lambda)$.

We can verify if the point process defined in the section above by the arrival times of large volatility changes can be modeled by a homogeneous Poisson process. For this purpose, we verify if the interarrival times $X_1, X_2, ...$ follow the same exponential distribution $Exp(\lambda)$.

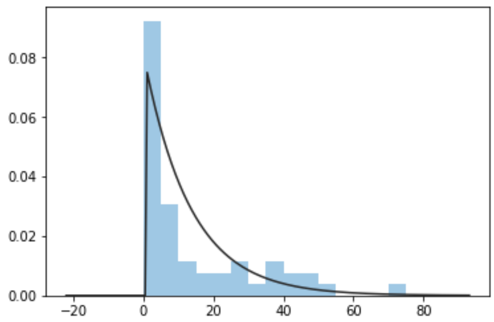


Figure 21: Histogram of the interarrival times fitted with an Exponential distribution

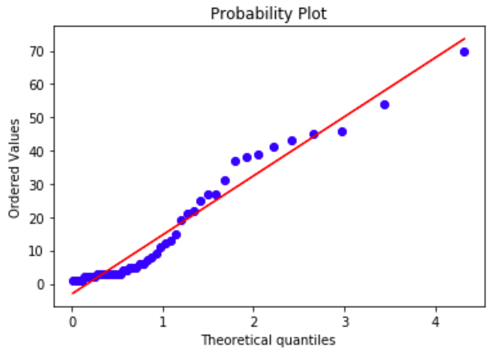


Figure 22: QQ-plot of the interarrival times against an exponential distribution

The figures above show that the interarrival times don't exactly follow an Exponential distribution $Exp(\lambda)$. This is further confirmed by a Kolmogorov-Smirnov equality test which returns a p-value of 0.0.

4.3 Hawkes Process modeling

4.3.1 Theoretical framework

In the case of a Hawkes process, the intensity (λ_t) is defined as:

$$\lambda_t = \mu + \int_0^t \phi(t-s)dN_s = \mu + \sum_{T_i<t} \phi(t-T_i)$$

where μ is the base intensity and ϕ is a measurable function with positive values such that $\int_0^T \phi(s)ds < +\infty, \forall t$.

The Hawkes process is self-exciting: with ϕ a non-increasing function (typical), the more events there are in a recent past, the more probable it is that a new event occurs.

In order to estimate the Hawkes process, we have to estimate the parameters μ and ϕ. A classical parametrization of ϕ consists in choosing:

$$\phi(t) = \sum_{j=1}^{p} \alpha_j e^{-\beta_j t} 1_{t>0}$$

We can choose to set $p = 4$. The parameters to estimate in this case are μ, α_j and β_j.

The likelihood on $[0, T]$ of an observation of a process $T_1 = t_1, ..., T_n = t_n$ is :

$$
\begin{aligned}
\mathcal{L} &= \frac{1}{dt_1...dt_n} P\big(T_1 \in [t_1 - dt_1, t_1], ..., T_n \in [t_n - dt_n, t_n], T_{n+1} > T\big) \frac{1}{dt_1...dt_n} \\
&= E\big[1_{Opt\in[0,t_1-dt_1],1pt\in[t_1-dt_1,t_1],...,1pt\in[t_n-dt_n,t_n],0pt\in[t_n,T]}\big] \frac{1}{dt_1...dt_n} \\
&= E\big[1_{Opt\in[0,t_1-dt_1],1pt\in[t_1-dt_1,t_1],...,1pt\in[t_n-dt_n,t_n]} E[1_{0pt\in[t_n,T]}|F_{t_n}]\big] \frac{1}{dt_1...dt_n} \\
&= \exp\big(-\int_{t_n}^{T} \lambda_s ds\big) E\big[1_{Opt\in[0,t_1-dt_1],1pt\in[t_1-dt_1,t_1],...,1pt\in[t_n-dt_n,t_n]}\big] \frac{1}{dt_1...dt_n} \\
&= \exp\big(-\int_{t_n}^{T} \lambda_s ds\big) E\big[1_{Opt\in[0,t_1-dt_1],1pt\in[t_1-dt_1,t_1],...,1pt\in[t_{n-1}-dt_{n-1},t_{n-1}]} E\big[1_{1pt\in[t_n-dt_n,t_n]}|F_{t_n-dt_n}\big]\big] \frac{1}{dt_1...dt_n} \\
&= \exp\big(-\int_{t_n}^{T} \lambda_s ds\big) \lambda(t_n) dt_n E\big[1_{Opt\in[0,t_1-dt_1],1pt\in[t_1-dt_1,t_1],...,1pt\in[t_{n-1}-dt_{n-1},t_{n-1}]}\big] \frac{1}{dt_1...dt_n} \\
&\dots \\
&= \prod_{i=1}^{N(T)} \lambda(t_i) \exp\big(-\int_0^T \lambda_s ds\big)
\end{aligned}
$$

Given the previous parametrization of the function ϕ, the log-likelihood can be written as follows:

$$l(\mu, \alpha, \beta) = \mu T + \sum_{j=1}^{4} \int_0^T \sum_{T_i} e^{-\beta_j(t-T_i)} 1_{t>T_i} dt$$

4.3.2 Results

The objective is to maximize the previous log-likelihood function with regards to the parameters μ, α, β with the constraints $\beta_i \geq 0, i = 1, ..., 4$. An optimization algorithm on Python using the Sequential Least Squares Programming algorithm allows to have the results of the optimization in the table below.

μ	α_1	α_2	α_3	α_4	β_1	β_2	β_3	β_4
61.18	1.96e+03	3.30e+02	1.96e+03	3.30e+02	1.25e-02	1.54e-02	1.25e-02	9.84e-03

Table 1: Results Maximum of Likelihood estimation. Initial parameters [0.2,0.01,0.1,0.01,0.01, 0.,0.01,0.,0.01]

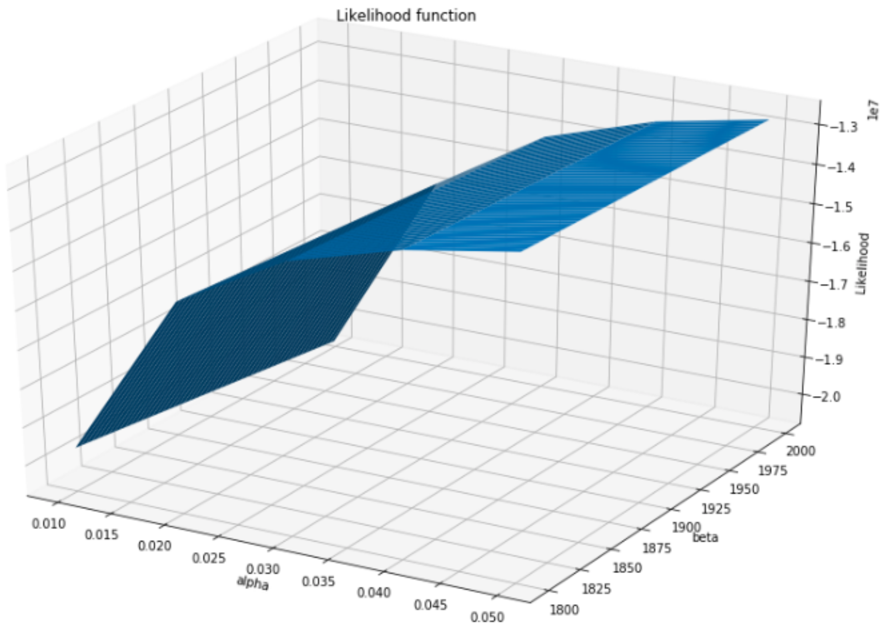


Figure 23: Two-dimensional plot of the log-likelihood function with regards to α_1 and β_1 that shows its concave nature and therefore the potential existence of a solution to the maximization problem

Given those parameters, we can plot the intensity of the process

$$\lambda_t = \mu + \sum_{T_i < t} \sum_{j=1}^{4} \alpha_j e^{-\beta_j (t - T_i)} 1_{T_i < t}$$

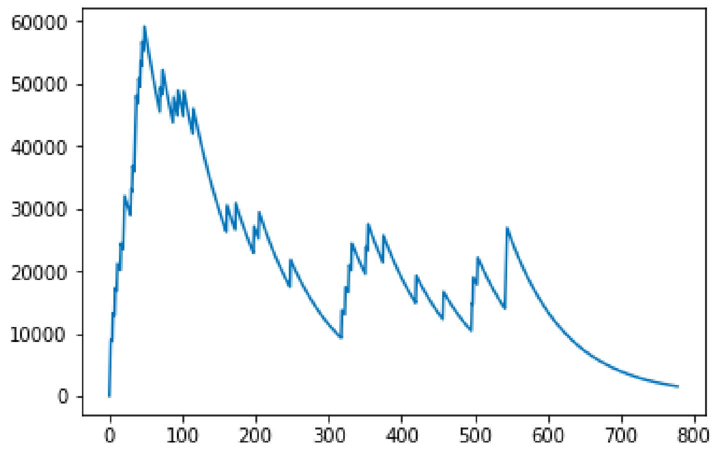


Figure 24: Intensity λ_t of the Hawkes process with regards to time t after maximum likelihood optimization of the parameters

In the figure above, we can see that the intensity λ_t has a discontinuous jump at arrival times T_i of a large volatility change, and then decreases until the next jump in volatility.

Conclusion

In conclusion, we found that the Bitcoin prices do not verify classical assumptions of financial data such as the one underlying the Black-Scholes model (normality and independence of the returns etc.). This has lead us to try different statistical approaches in order to analyze this complex structure.
The ARIMA model, although a powerful tool for forecasting, was not very effective in our setup as the assumptions for its validity were not verified (namely assumptions on autocorrelation). Conversely, GARCH models have proven very useful as we managed to replicate rather effectively the dynamic of Bitcoin returns. We were also able to capture some stylistic effects of the returns time series that were missing in the ARIMA model, such as fat tails and conditional heteroscedasticity.
Finally, we used a Poisson point process to model the arrival times of large changes in volatility to try to understand the volatility clustering phenomenon. Then, we moved on to a Hawkes model which we estimated by a Maximum-likelihood estimation to find the best parameters of the Hawkes model using an exponential kernel. This model was successfully capturing the self-exciting nature of the point process of large volatility changes, which was not possible by using a homogeneous Poisson process.