Time Series Analysis

Daniele Piloni

February 2021

1 Introduction

The goal of this project is to capture different features of financial data, trying to find the time series generator process. In particular, I adapt ARMA and GARCH models to see if financial returns, represented by Lufthansa stocks historical performance, show that the conditional mean and the conditional variance depend on the past. Being able to identify the generator process can potentially be used for time series forecasting, in case we can assume stationarity. An autoregressive-moving-average model ARMA(p,q) is defined by the equation:

$$X_t = c + \phi_1 X_{t-1} + ... + \phi_p X_{t-p} - \theta_1 a_{t-1} - ... - \theta_q a_{t-q} + a_t,$$
$$\text{where } a_t \sim wn(0, \sigma_a^2).$$

The key features of the ARMA model are:

- Unconditional mean is constant,
- Unconditional variance is constant,
- Conditional mean depends on the past,
- Conditional variance is constant,
- Autocorrelation function (ACF) decay quickly to zero,
- Partial autocorrelation function (PACF) decay quickly to zero.

A generalized autoregressive conditional heteroschedasticity process GARCH(p,q) is defined by the equation:

$$X_t = \sigma_t \epsilon_t,$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 X_t^2 - 1 + ... + \alpha_p X_t^2 - p + \beta_1 \sigma_t^2 - 1 + ... + \beta_q \sigma_t^2 - q,$$

$$\text{where } \epsilon_t \sim wn(0,1), \alpha_0 > 0, \text{ and } \beta_j >= 0.$$

The key features of the GARCH model are:

- Unconditional mean is constant,
- Unconditional variance is constant,
- Conditional mean is constant,
- Conditional variance depends on the past,
- Autocorrelation function (ACF) decay quickly to zero,
- Partial autocorrelation function (PACF) decay quickly to zero.

2 Graphical Representation

As shown in Figure 1, we can easily see that the price series is non-stationary (in facts, the long-term trend is decreasing). For what concerns the return series, we can assume the stationarity, and it's possible to see volatility clusters.

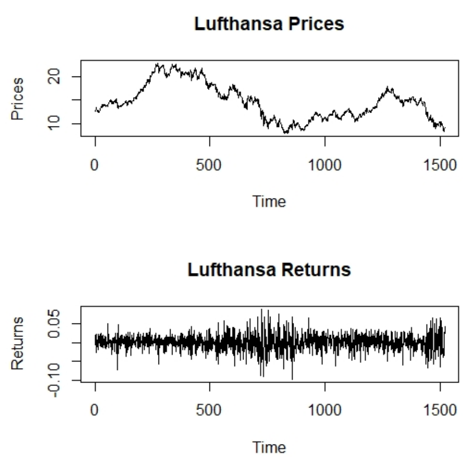

Figure 1: Lufthansa prices and returns time series

3 Descriptive Analysis

In Table 1 are shown the descriptive statistics of the returns distribution. We can observe two common characteristics of financial returns: the negative skewness and the excess kurtosis.

In Figure 2 the observed distribution is compared to the normal, and it's easier to see the excess kurtosis. Using the Jarque-Bera test to confirm the statement, for a p-value of 0,00 we can reject the null hypothesis and confirm that Lufthansa returns are not normally distributed.

Lufthansa Returns	
N.obs	1520
Min	-0,09
Max	0,08
Mean	-0,01
Median	0,00
Variance	0,01
Std.dev	0,02
Skewness	-0,16
Kurtosis	4,75

Table 1: Descriptive statistics

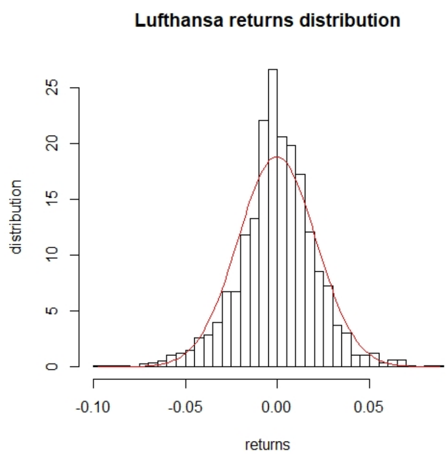

Figure 2: Lufthansa returns distribution

4 Arma Model Fitting

To verify if autoregressive models can fit the data, we need to look at the autocorrelation (ACF) and partial autocorrelation functions (PACF). From Figure 3 we can see that only the coefficient corresponding to lag 5 is significantly different from zero, so we can assume that the conditional mean doesn't depend on the past.

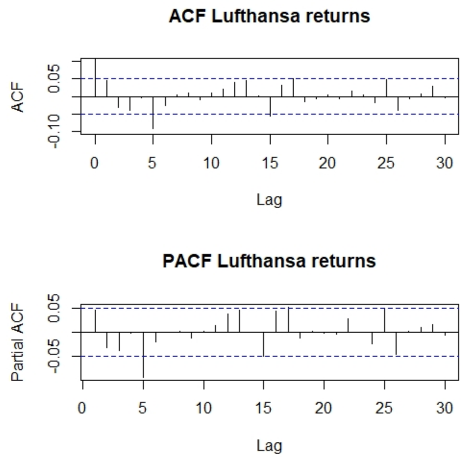

Figure 3: Lufthansa returns ACF and PACF

5 GARCH Model Fitting

In order to check heteroschedasticity in the time series, we have to look at ACF and PACF of squared returns that are shown in Figure 4. In this case, we can see the decay to zero of both functions (even if the decay of ACF is not quickly).

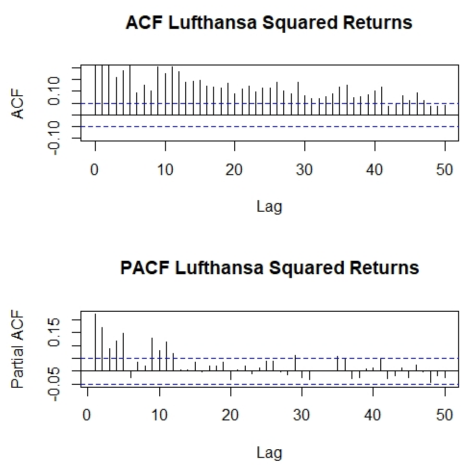

Figure 4: Lufthansa squared returns ACF and PACF

In Table 2 are shown coefficient estimates of a GARCH(1,1) model. We can see that all the coefficients are statistically significant for a chosen level of 5%, despite the intercept value is very close to zero.

Coefficient	Estimate	Std. Error	P-value
Alfa 0	0,000	0,000	0,045
Alfa 1	0,058	0,015	0,000
Beta 1	0,924	0,023	0,000

Table 2: GARCH(1,1) coefficients estimates

In order to check if the GARCH(1,1) model is able to to fit the model, we need to see if the residual don't show heteroschedasticity. Ljung-Box and Engle's ARCH tests, whose results are shown in Table 3, confirm that there aren't heteroschedasticity effects in the residuals: for a given significance level of 5%: in facts, we can accept the null hypothesis of the two tests.

Test	Statistic	P-value
LB test	5,498	0,855
ARCH test	6,461	0,692

Table 3: Ljung-Box and Engle's ARCH tests results

Finally, we need to check the assumpion that residuals are normally distributed. Figure 5 shows the Q-Q plot, comparing residuals distribution quantiles with a normal distribution. We can clearly see differences in the tails.

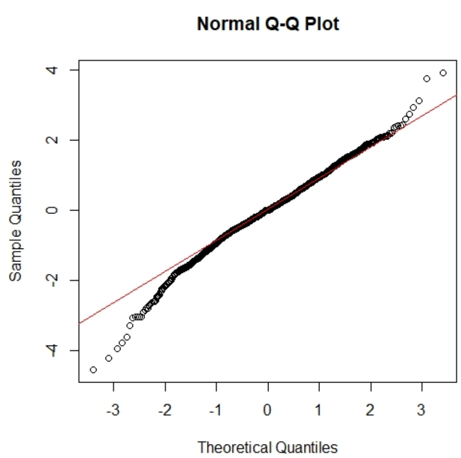

Figure 5: Residuals Q-Q plot with normal distribution

We can try to estimate again a GARCH(1,1), this time assuming that the white noise is distributed like a Student's T distribution. Table 5 shows the new coefficients' estimates; notice that for this model the intercept is not statistically significant. Again, Ljung-Box and ARCH tests both exclude the presence of conditional heteroschedasticity in the residuals.

Coefficient	Estimate	Std. Error	P-value
Alfa 0	0,000	0,000	0,104
Alfa 1	0,065	0,019	0,000
Beta 1	0,921	0,026	0,000

Table 4: GARCH(1,1) with Student's T white noise coefficients estimates

Finally, Figure 6 the Q-Q plot of the residuals; this time the assumption of the white noise distribution seems acceptable.

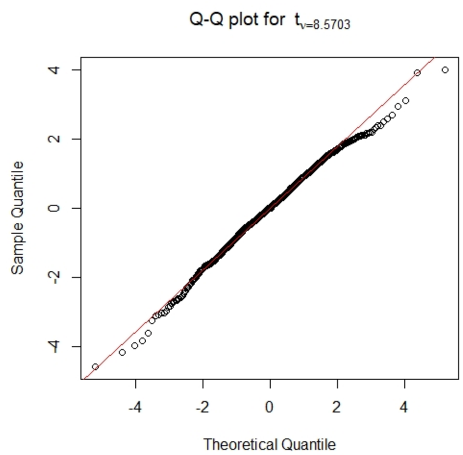

Figure 6: Residuals Q-Q plot with Student's T distribution

6 Leverage Effect

Finally, I want to check one last important behaviour financial returns: leverage effect, i.e. asymmetric volatility. In fact, we can often observe in financial markets that negative shocks are followed by higher volatility than positive shocks. In order to check for leverage effect, we can use Sign Bias Test. Due to the fact we reject the null hypothesis, we have to find a model that capture leverage effect. GARCH model is not able to represent the generator process of the series, because it can't get asymmetric effects. That's why I try to fit an exponential GARCH (or EGARCH) model to the data. An EGARCH(1,1) model is defined by the equation:

$$X_t = \sigma_t \epsilon_t,$$
$$\log\sigma_t^2 = \alpha_0(1-\alpha_1) + \alpha_1 \ \log\sigma_{t-1}^2 + f(\epsilon_{t-1})$$
$$\text{where } f(\epsilon_t) = \theta\epsilon_t + \gamma[|\epsilon_t| - E(|\epsilon_t|)] \text{ and } \epsilon_t \sim iid(0,1).$$

The $f(\epsilon_t)$ function term in the second equation is able to capture leverage effect. The second equation can be rewritten as:

$$\log\sigma_t^2 = \omega + \beta_1 \ \log\sigma_{t-1}^2 + \alpha_1\epsilon_{t-1} + \gamma_1(|\epsilon_{t-1}| - E[|\epsilon_t|]).$$

This is the notation used in the *rugarch* R package. Table **??** shows the estimates of the parameters. The assumption on the error term distribution is Student's T.

Coefficient	Estimate	Std. Error	P-value
Omega	-0,131	0,001	0,000
Alfa 1	-0,055	0,012	0,000
Beta 1	0,983	0,000	0,000
Gamma 1	0,101	0,008	0,000

Table 5: EGARCH(1,1) coefficients estimates

Q-Q plot in Figure 7 confirms that the assumption on the erratic component is acceptable.

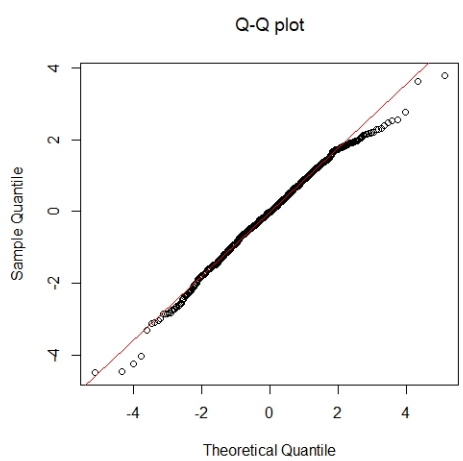

Figure 7: EGARCH error term Q-Q plot with Student's T distribution

7 Conclusions

The generator process of Lufthansa returns time series can be described by the following equations:

$$X_t = \sigma_t \epsilon_t,$$
$$\log \sigma_t^2 = \alpha_0(1-\alpha_1) + \alpha_1 \log \sigma_{t-1}^2 + f(\epsilon_{t-1})$$
$$\text{where } f(\epsilon_t) = \theta\epsilon_t + \gamma[|\epsilon_t| - E(|\epsilon_t|)] \text{ and } \epsilon_t \sim std(0,1).$$

We can finally plot in Figure 8 the estimated volatility series.

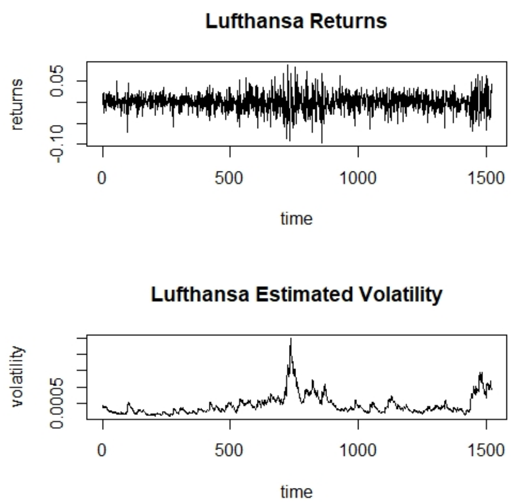

Figure 8: Lufthansa returns and estimated volatility series