

GARCH TIME SERIES PROCESS

Econometrics 7590

Projects 2 and 3

Zsuzsanna HORVÁTH and Ryan JOHNSTON

Abstract: We examine in detail the various attributes of a time series data set for financial returns. We consider multiple time series models in order to determine which will provide the best fit. Particular attention is placed on the mathematical properties of the most common models. Topics addressed include: differencing, stationarity, autocorrelation, lag, acf, partial acf, independence, AR process, MA models, GARCH process, and integrated models, among others. We perform rigorous analysis and statistical tests in order to support our conclusions concerning the results.

Contents

1	The Data	3
2	Fitting the Data	4
2.1	AR(1)	7
3	Fitting GARCH	14
4	Introduction to GARCH processes	14
5	Estimating the Parameters	18
5.1	Further Analysis	19
5.2	Chi-Square Test for Goodness-of-Fit	20
6	Conclusion	24
7	Bibliography	25

1 The Data

Our peak interest has been in stock market analysis; which is considered a time series regardless of which index is used, so we had a lot to choose from. We looked at a few of the data sets found on the *EconStats* website and decided to go with the first one listed, which happened to be S&P500. We decided to go with the daily index since that would produce the most number of observations. We also had to choose exactly what set of observations we wanted to use: open price for the day, high price for the day, low price for the day, close price for the day, and percent change (returns) of the close price for the day from the previous day. We chose to take a look at the latter two and try to decide which one of the two might fit what we are trying to accomplish, in particular we would like to have a stationary data set. Thus our set consists of 4,231 daily closing prices and returns from Standard and Poor's index of large firms ranging from 1/2/1990 to 10/10/2006.

We now plot the two data sets shown by the following two figures: Figure 1.1 and Figure 1.2. Notice that the plot of the closing prices are not stationary (i.e. the data does not fluctuate around some common mean or location, it changes throughout the time series), whereas the plot for the returns does appear to be stationary (the data does fluctuate around a common mean or location). When dealing with time series it is desirable to have a stationary data set primarily because the traits of a stationary data set allow one to assume models that are independent of a particular starting point, which may be difficult to obtain in practice.

We were fortunate to find a data set that has stationary data points (returns), although we could have easily performed a transformation on a non-stationary data set to make it stationary. This process is called differencing. The most basic method of differencing consists of simply taking the difference between consecutive observations. For example, suppose we have the data set consisting of $Y_{i=1}^N$. The differenced data set would consist of $X_i = Y_{i+1} - Y_i$. Notice that the differenced data set has one less observation than the original data set. If this new data set is still not stationary, the process can be repeated until a stationary data set is obtained; typically a non-stationary data set only needs to be differenced one or two times to obtain a stationary data set. Another common type of differencing is called taking the log returns. This consists of taking the natural log of the quotient of consecutive observations; in other words

$$X_i = \log \left[\frac{Y_{i+1}}{Y_i} \right].$$

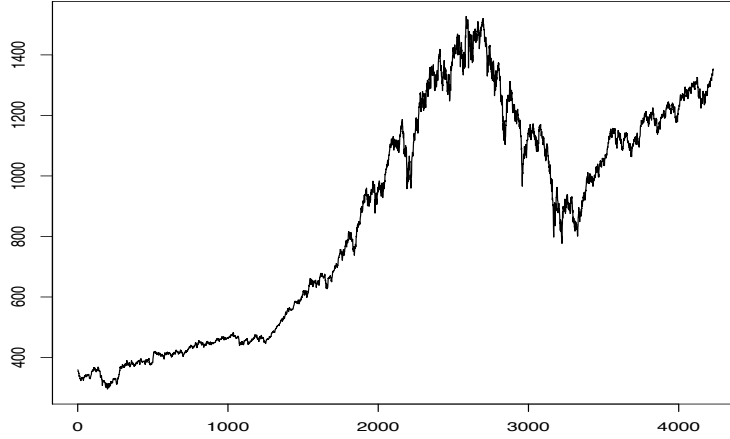


FIGURE 1.1 Daily observations of Close Price from S&P500 from 1/2/1990 to 10/10/2006

This equation can be rewritten as

$$X_i = \log \left[\frac{Y_{i+1}}{Y_i} \right] = \log(Y_{i+1}) - \log(Y_i).$$

This equation is used for continuous time series data sets. The analog of the above equation for discrete time series data sets has also been developed and is represented by the following equation

$$X_i = \frac{(Y_{i+1} - Y_i)}{Y_i}.$$

This is in fact the equation used to obtain the returns from the daily closing prices, so we will be considering the returns throughout the remainder of this paper since it is stationary.

2 Fitting the Data

Observing the time series data set of returns, we see that there exists heteroskedasticity in the model. However, we cannot determine whether this is enough to warrant consideration. We will begin with a simple model and work our way to more complex models when and if the simple models are deemed poor fits to the data. Thus, we begin by considering the most common time series models: the autoregressive (AR), moving average (MA), and integrated (I) models. To do this we take a look at the autocorrelation function (acf)

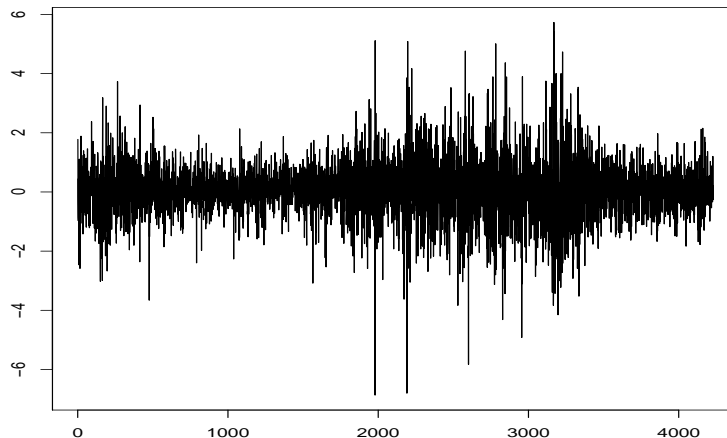


FIGURE 1.2 Daily observations of Returns from S&P500 from 1/2/1990 to 10/10/2006

of the data set, which is shown in Figure 2.1 below. The acf, as the name implies, shows a self (auto) correlation or relationship among observations. What influence or effect does one observation have on successive observations and how far does that effect extend down to succeeding observations? The acf plot can be used to determine what sort of time series might be appropriate, if any. The plot shows the lags from one observation to the next where the value of the observation at lag zero is always one. A geometrically decaying acf plot would indicate that an AR model would be ideal or possibly a combination of an AR and MA model. One that shows no discernible pattern would most likely be considered random or white noise. Notice that the first lag of the acf plot is close to zero, indicating that our data set does not appear to have much correlation between observations.

The most we could expect from an AR model would be a first-order model $AR(1)$. We can check this by looking at the plot of the partial autocorrelation function (pacf). The pacf is used to determine the appropriate order of a fitted AR data set. The pacf of our data set is shown in Figure 2.2. The order is determined by viewing the lines that fall outside the confidence bounds (the dotted lines) and counting how many lags it takes for the data to fall inside the confidence bounds. In this case it appears that it only takes one lag. By viewing the acf and pacf, the evidence is weak towards finding a good fitting AR model for the data. According to the acf and pacf the data looks random and certainly shows no easily discernible pattern. This would support the appearance of the time series plot since the plot looks a lot like white noise except for the change in spread (variation) of observations. Such heteroskedasticity would most likely not be

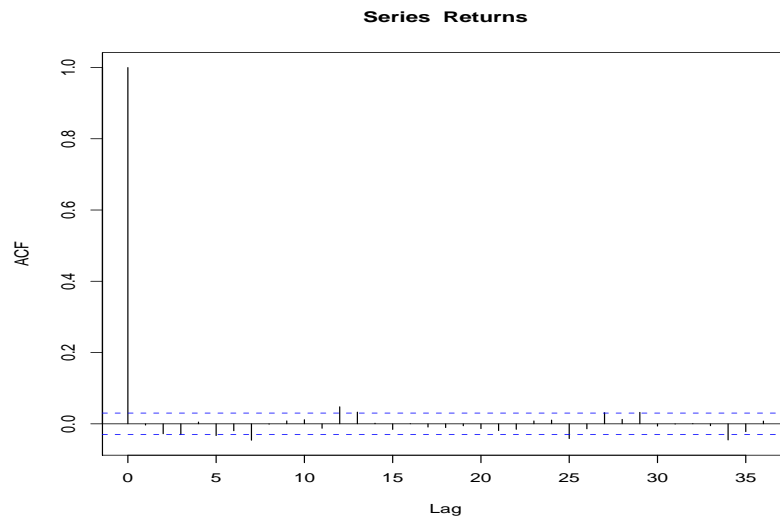


FIGURE 2.1 Autocorrelation function for Returns

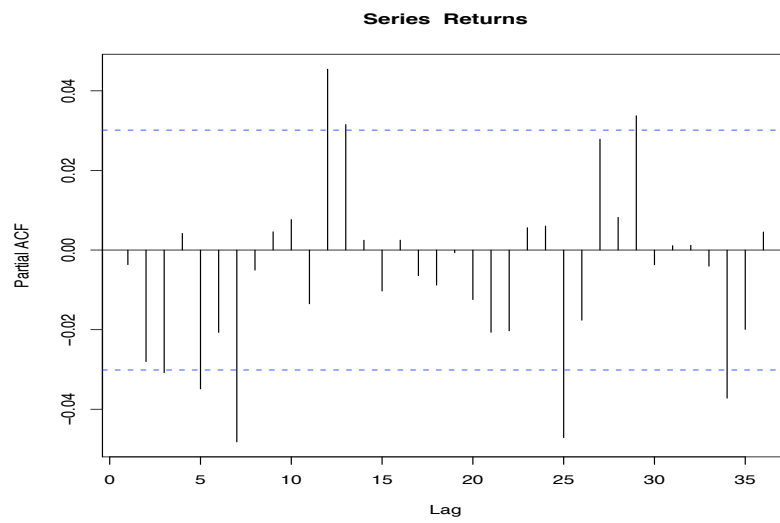


FIGURE 2.2 Partial autocorrelation function for Returns

evident in a truly random data set. This, however, does not mean we should rule out the possibility that the data fits an AR(1) model with weak autocorrelation or possibly an ARIMA model. Before proceeding, we review a few of the critical properties of an AR(1) process in the following subsection.

2.1 AR(1)

Recall the following equations for an AR(1) process,

$$X_k = \rho X_{k-1} + \varepsilon_k. \quad k = 0, \pm 1, \pm 2, \dots$$

The value ρ represents the correlation between observations. Typically we think of correlation as being a measured relationship between two or more variables, however, in this case we have only one variable. The variable is simply viewed at different time points so the correlation between observations is called autocorrelation (self-correlation) since it is measuring the correlation between the same variable at different time periods. Furthermore, ρ can also be thought of as the percent relationship between two consecutive observations. Thus, the value of the observation at time k is $\rho\%$ of the value of the observation at time $k - 1$ plus the value of the assumed error term. To obtain the relationship between the time zero observation and subsequent lags we simply count the number of lags and multiply ρ that many times. For example say we wanted to know the effect of the third lag for the AR(1) model. We would calculate ρ^3 . It is worth mentioning that modelling a stationary time series assumes the value of ρ to be between -1 and 1 . Note that if $|\rho| = 1$ the model would be completely deterministic and for $|\rho| = 0$ the model takes on the values of the error terms, which are typically assumed to be random. Thus it can be concluded that the closer to zero that the value of $|\rho|$ takes on, the more random the model is. To perhaps gain a more intuitive feel for the effect of ρ , compare Figures 2.3 through 2.12 consisting of simulated AR(1) models with varying parameters of ρ and varying distributions for the error terms. We note the lag effect that ρ has on the plots. For larger values of ρ the number of subsequent observations required for the data to return to the mean is greater than for smaller values of ρ . In the case for $\rho > 1$, we get an unstable non-stationary process.

We turn our attention back to trying to fit our data to an AR model. To accomplish this we use the programming language R to explore these possibilities. We first try a function that calculates the AR order of the data set. The following is a printout of the results for the order and the estimated coefficients:

Coefficients:

1	2	3	4	5	6	7	8
-0.0071	-0.0304	-0.0325	0.0016	-0.0342	-0.0190	-0.0455	-0.0042
9	10	11	12	13			
0.0056	0.0100	-0.0123	0.0456	0.0315			

Order selected 13 σ^2 estimated as 0.9943.

The order calculated is 13. For practical purposes, an order of this degree would generally not be considered a good fit since it would be inconvenient and difficult to work with such a high order process. The order can be restricted and used to estimate the coefficients of a desired process of lower order. We take this approach to estimate the coefficient for an AR(1) model and obtain

Coefficients:

1
-0.0037

Order selected 1 σ^2 estimated as 1.001.

We note the difference in the estimated parameter for the first coefficient of the previous two models; -.0071 for the former and -.0037 for the latter. To keep notation, we will call this value ρ . For our obtained values, -.0071 and -.0037, the negative sign implies an inverse relationship between observations and a very small effect in both cases $-.71\%$ and $-.37\%$, respectively. These estimated parameters imply that our data is a random time series. Hence we can conclude that a simple AR model is not a good fit for our data set.

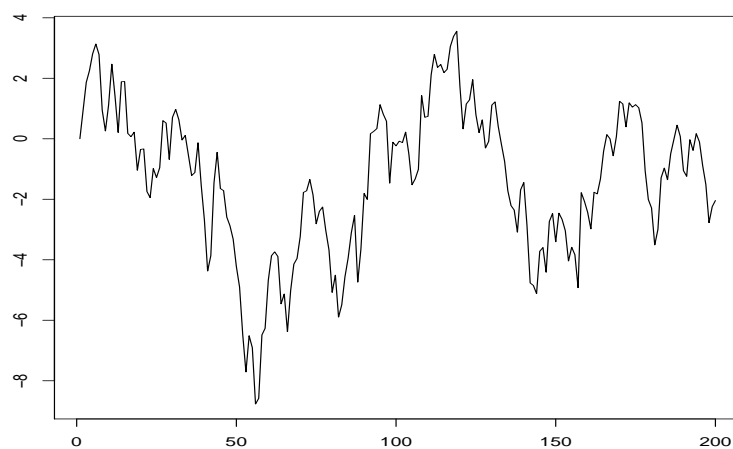


FIGURE 2.3 200 simulated observations from an AR(1) process with $\rho = .95$ and standard normal error terms

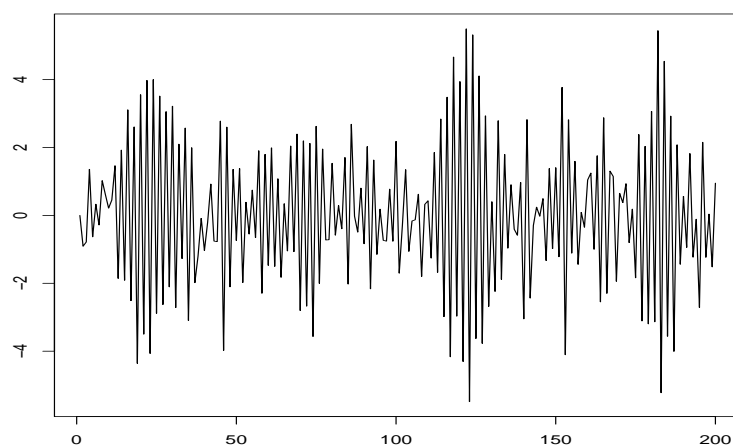


FIGURE 2.4 200 simulated observations from an AR(1) process with $\rho = -.9$ and standard normal error terms

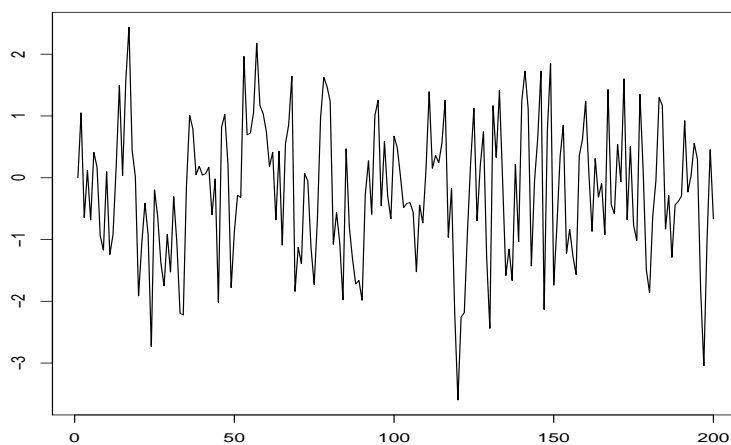


FIGURE 2.5 200 simulated observations from an AR(1) process with $\rho = .5$ and standard normal error terms

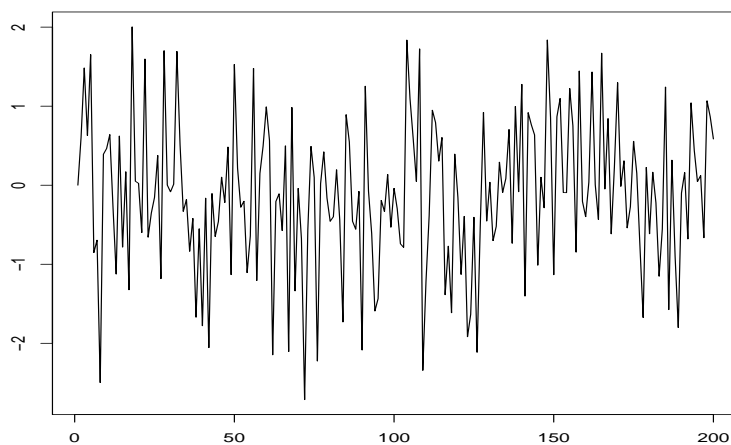


FIGURE 2.6 200 simulated observations from an AR(1) process with $\rho = .01$ and standard normal error terms

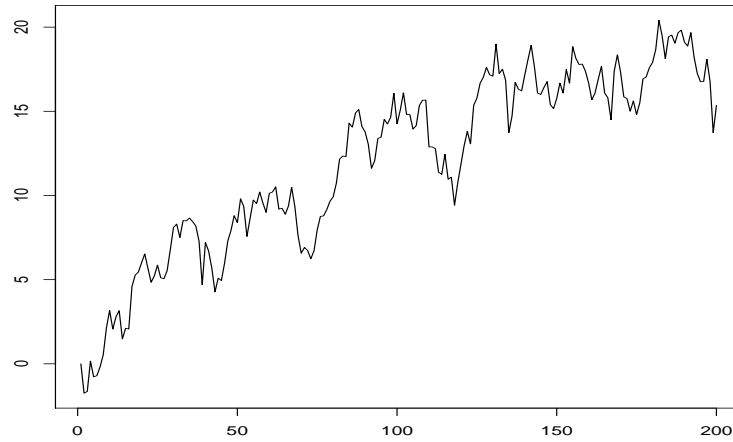


FIGURE 2.7 200 simulated observations from an AR(1) process with $\rho = 1.01$ and standard normal error terms

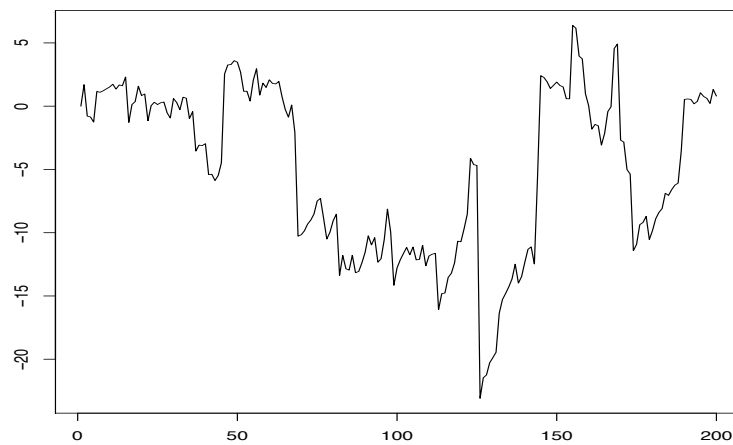


FIGURE 2.8 200 simulated observations from an AR(1) process with $\rho = .95$ and Cauchy distributed error terms with parameters (location = 0, scale = .5)

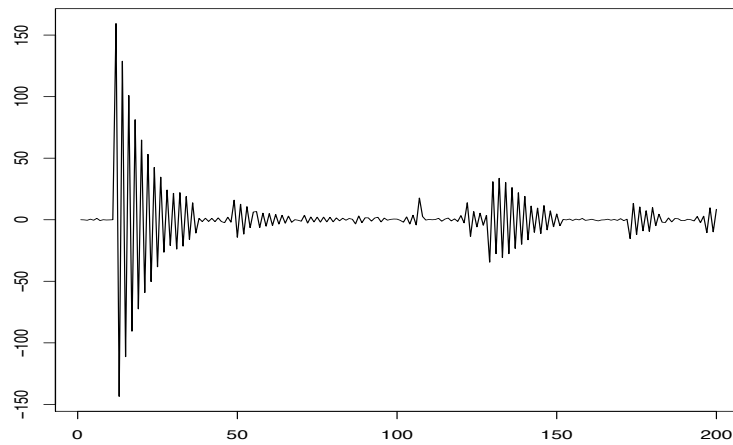


FIGURE 2.9 200 simulated observations from an AR(1) process with $\rho = -.9$ and Cauchy distributed error terms with parameters (location = 0, scale = .5)

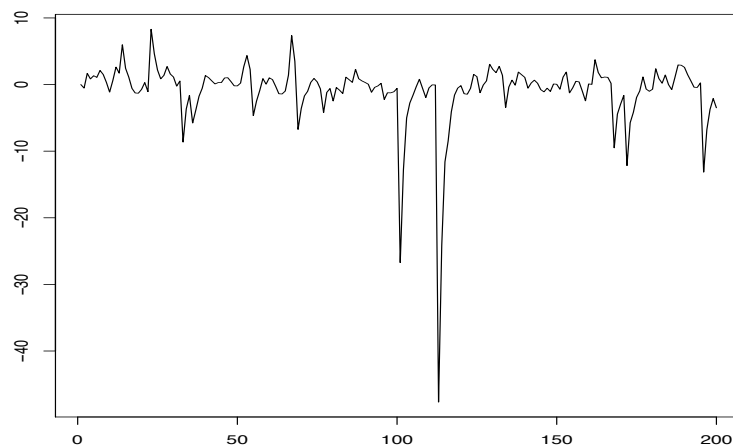


FIGURE 2.10 200 simulated observations from an AR(1) process with $\rho = .5$ and Cauchy distributed error terms with parameters (location = 0, scale = .5)

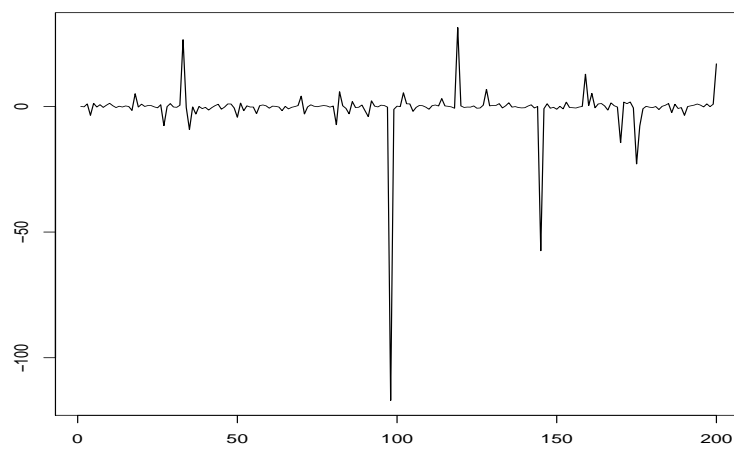


FIGURE 2.11 200 simulated observations from an AR(1) process with $\rho = .01$ and Cauchy distributed error terms with parameters (location = 0, scale = .5)

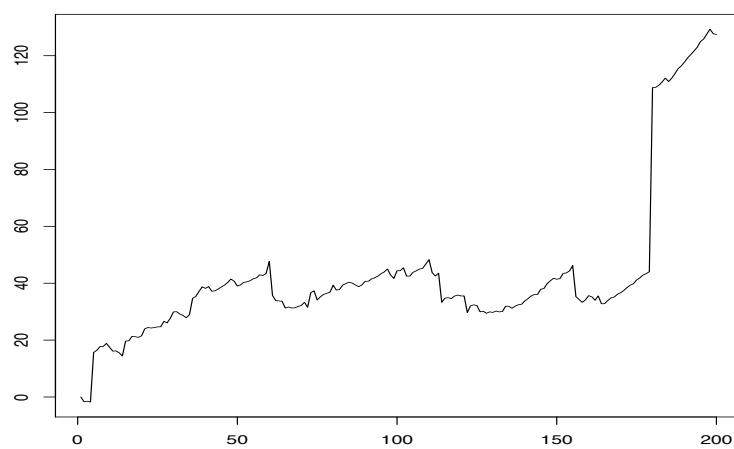


FIGURE 2.12 200 simulated observations from an AR(1) process with $\rho = 1.01$ and Cauchy distributed error terms with parameters (location = 0, scale = .5)

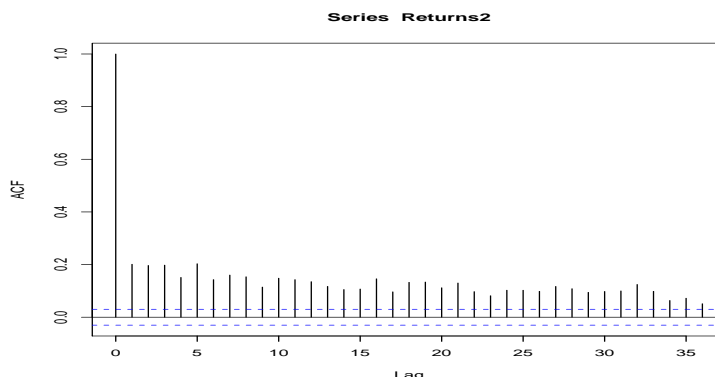


FIGURE 3.1 Autocorrelation function of squared Returns.

3 Fitting GARCH

We now shift gears in an attempt to find some sort of pattern in the data that would suggest the use of a different model. Recall from the previous section the acf showed that the returns appeared to be random. It even appeared that they were equally likely to be positive or negative from one observation to the next. This implies that the sign of the value of observations is independent of the past, however, the magnitude of change in observations may show correlation. To see this, we square the observations and run a new acf plot on this new data set of squared returns. The result of the acf plot is shown in Figure 3.1. Notice the slow decay of the lag plots, this is great news for us; this indicates there is correlation between the magnitude of change in the returns. In other words, there is serial dependence in the variance of the data. We run a plot of the pacf to further strengthen this proposal. Figure 3.2 shows the pacf of the squared returns. These results indicate that a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model would be a good model to try fitting the data to.

4 Introduction to GARCH processes

In order to assess the risk of many financial transactions, estimates of asset return volatility are required. Accurate measures and reliable forecasts of volatility are essential to numerous aspect of finance and economics. An example of the aforementioned includes

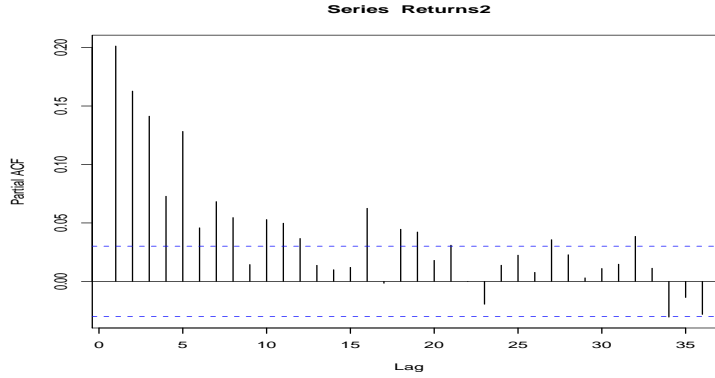


FIGURE 3.2 Partial autocorrelation function of squared Returns.

the importance of correct derivative pricing strategies, as well as the trading and hedging techniques that occur in portfolio allocation. Financial return volatility data is highly influenced by time dependence, which can cause volatility clustering. Time series such as this can be parameterised using the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, which can then be used to forecast volatility.

The GARCH(1,1) model was originally introduced by Bollerslev (1986) and Engle (1982) in order to model the fluctuations of the variances of the time series data. They observed that there is significant variation between the variances of log returns of stocks. Since the aforementioned stock values do not have constant variance, the volatility of the stock market can be modeled using the GARCH(1,1) model. It is assumed that the log returns will satisfy the following equations

$$(4.1) \quad y_k = \sigma \epsilon_k, \quad -\infty < k < \infty$$

and

$$(4.2) \quad \sigma_k^2 = \omega + \alpha y_{k-1}^2 + \beta \sigma_{k-1}^2 \quad -\infty < k < \infty,$$

where (ω, α, β) is the parameter of the process. It is assumed that the errors (innovations) $\epsilon_k, -\infty < k < \infty$ are independent identically distributed random variables. Nelson (1990) found the necessary and sufficient condition for the existence of $(y_k, \sigma_k^2), -\infty < k < \infty$.

Lumsdaine (1996) used the quasi-maximum likelihood method to estimate the parameters.

Financial data typically show the spread and clustering of the volatility of the data. The ARCH (Engle, 1982) and GARCH (Bollerslev, 1986) models were designed as a model by which to capture such features of volatility of data. More specifically, the time-series technique, GARCH allows for modeling the serial dependence of the volatility. Due to the conditional property of GARCH, the mechanism depends on the observations of the immediate past, thus including past variances into explanation of future variances. Hence, GARCH models rely on time-varying variance. In other words, the model considers the variance of the current error term to be a function of the variances of the previous time period's error terms.

The GARCH model also takes into account volatility clustering and tail behavior; important characteristics of financial time series. It provides an accurate assessment of variances and covariances through its ability to model time-varying conditional variances. Bollerslev (1986) developed GARCH based on Engle's original ARCH volatility modeling technique. Bollerslev (1986) designed GARCH using fewer parameters, which lessens computations and may provide a better fit for the data.

In this paper we use a large sample size (4231 observations) in order to best analyze the results. Furthermore, we examine changes which occur using different values for the parameters. We will investigate if our large set of financial data can be fitted to a time series model, and which model will provide the best fit.

We now examine the behavior of the GARCH(1,1) model as we vary the parameters. Figure 4.1 contains the graph of $y_k, 1 \leq k \leq 500$, when $\omega = 1$, $\alpha = .02$, and $\beta = .3$. By comparison, Figure 4.2 contains the graph of $y_k, 1 \leq k \leq 500$, when $\omega = 1$, $\alpha = .02$, and $\beta = .9$. We note that in Figure 4.1 the values that y_k takes ranges between $[-3,3]$. By slightly changing the value of β , the y_k 's now range between $[-10,10]$. To gain further inference concerning our results, we examine the Figure 4.3.

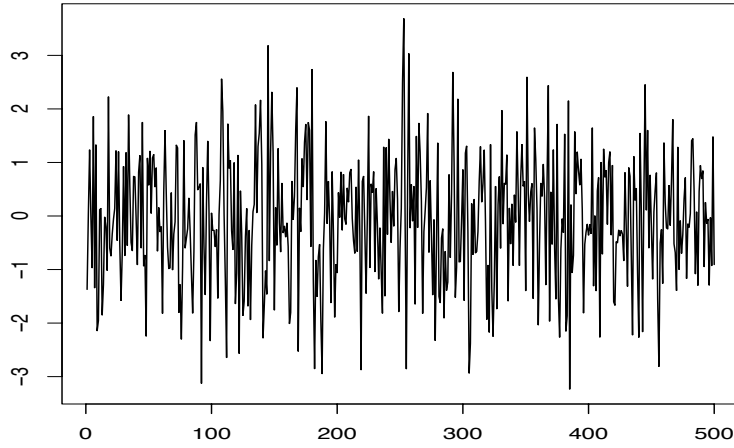


FIGURE 4.1 $n = 500$, $\omega = 1$, $\alpha = .02$, $\beta = .3$

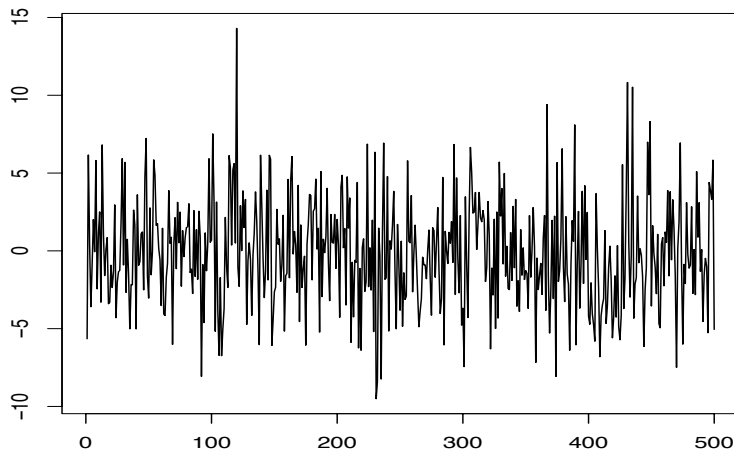


FIGURE 4.2 $n = 500$, $\omega = 1$, $\alpha = .02$, $\beta = .9$

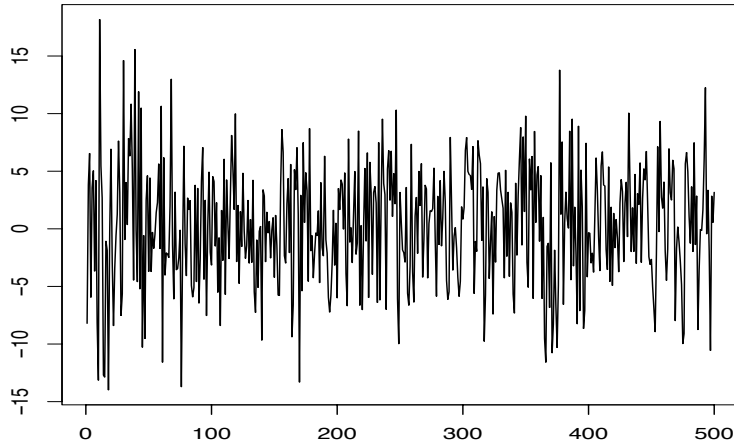


FIGURE 4.3 $n = 500$, $\omega = 1$, $\alpha = .07$, $\beta = .9$

5 Estimating the Parameters

Using computer simulation, we obtain estimates for the parameters of the GARCH(1,1) process. The best fit for our data is a GARCH(1,1) model with the following parameters: $\omega = .0005306$, $\alpha = 0.054962$, $\beta = 0.940181$. The computer output that produced these results can be found below:

Call:

```
garch(x = returns)
```

Model:

```
GARCH(1,1)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-6.72635	-0.53679	0.05432	0.64415	4.30692

Coefficient(s):

	Estimate	Std. Error	t value	Pr(> t)	
a0	0.0005306	0.001123	4.724	2.31e-06	***
a1	0.054962	0.004357	12.615	< 2e-16	***
b1	0.940181	0.004747	198.042	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 '.'

We note that the produced p-values provide extremely strong evidence that the GARCH(1,1) model with the specified parameters provide the best fit for our data. All p-values are significant at all specified levels, as shown on the output.

The simulation using these parameters can be found in Figure 5.2. We superimpose the plot of the original and simulated data. This can be found in Figure 5.3. We note that both graphs contain 4231 observations, and are plotted on the same axes in order to best compare the results. Figure 5.3 provides further evidence that our data can be modeled by a GARCH(1,1) model with parameters: $\omega = .0005306$, $\alpha = 0.054962$, $\beta = 0.940181$. Although both Figure 5.3 and our produced p-values concerning the fit of our coefficients indicate a good fit, we perform diagnostics to rigorously prove our conclusions concerning goodness-of-fit.

5.1 Further Analysis

For further analysis, we test to see how probable it is that our data came from a GARCH(1,1) process with the specified parameters. We fit the GARCH(1,1) time series model to the data by computing the maximum-likelihood estimates of the conditionally normal model. Furthermore, we test the residuals for normality and remaining ARCH effects. We analyze the residuals and use the Ljung-Box test to test for independence.

The Ljung-Box test is used to test the null hypothesis of independence using the Ljung-Box statistic. We note that the Ljung-Box test is applied to the residual of a fitted GARCH model, not the original data. Below are the results from the computer output:

Diagnostic Test:

Box-Ljung test

data: Squared.Residuals

X-squared = 0, df = 1, p-value = 0.9947.

We note the extremely large p-value of 0.9947. Recall, the p-value of a statistical hypothesis test is the probability of getting a value of the test statistic as extreme as or more extreme than that observed by chance alone, if the null hypothesis is true. It

is equal to the significance level of the test for which we would only just reject the null hypothesis. The p-value is compared with the significance level and, if it is smaller, the result is significant. Small p-values suggest that the null hypothesis is unlikely to be true. The smaller it is, the more convincing is the rejection of the null hypothesis. A p-value as large as 0.9947, indicates that the null hypothesis cannot be rejected at any reasonable level of significance.

5.2 Chi-Square Test for Goodness-of-Fit

The histogram of the residuals can be found in Figure 5.1. We superimpose this with a normal density curve with parameters $\mu = .05441$ and $\sigma = .85$. In order to obtain a detailed picture, we generate a histogram with 45 equally spaced bins. The distribution is clearly normal. The question now becomes estimating the parameters of the normal in order to obtain a nice fit. We find that a normal with the above specified parameters provides a good fit, as shown on the graph. To strengthen this conclusion, we perform a χ^2 test. The hypotheses we are testing are

H_0 : residuals are normally distributed with parameter: $\mu = .05441$ and $\sigma = .85$

H_a : H_0 is not true.

We will divide the interval into K cells and calculate the number of observations that fall into each cell. We define this as Y_1, Y_2, \dots, Y_{10} . We define the range of our 10 cells as

$$\begin{aligned} &[-\infty, -0.35), [-0.35, -0.25), [-0.25, -0.15), [-0.15, -0.05), \\ &[-0.05, 0.05), [0.05, 0.1), [0.1, 0.2), [0.2, 0.25), [0.25, 0.35), [0.35, \infty). \end{aligned}$$

Let $K = 10$ and $n = 4231$.

We define

$$E_i = n \int_{t_{i-1}}^{t_i} f(x) dx,$$

where $f(x)$ is the density of the normal distribution.

Recall

$$\sum_{i=1}^K \frac{(Y_i - E_i)^2}{E_i} \sim \chi^2(K - 1).$$

We obtain that:

$$Y_1 = 1339$$

$$Y_6 = 94$$

$$Y_2 = 182$$

$$Y_7 = 190$$

$$Y_3 = 190$$

$$Y_4 = 205$$

$$Y_5 = 192$$

$$Y_8 = 94$$

$$Y_9 = 197$$

$$Y_{10} = 1546.$$

Using the definition of E_i gives:

$$E_1 = 1340.7$$

$$E_2 = 181.8$$

$$E_3 = 189.6$$

$$E_4 = 195.1$$

$$E_5 = 197.9$$

$$E_6 = 99.2$$

$$E_7 = 197.1$$

$$E_8 = 97.2$$

$$E_9 = 190.2$$

$$E_{10} = 1539.0$$

Using the definition of χ^2 we get:

$$\begin{aligned} & \frac{(1339 - 1340.7)^2}{1340.7} + \frac{(182 - 181.8)^2}{181.8} + \frac{(190 - 189.6)^2}{189.6} + \frac{(205 - 195.1)^2}{195.1} + \frac{(192 - 197.9)^2}{197.9} \\ & \frac{(94 - 99.2)^2}{99.2} + \frac{(190 - 197.1)^2}{197.1} + \frac{(94 - 97.2)^2}{97.2} + \frac{(197 - 190.2)^2}{190.2} + \frac{(1546 - 1539.0)^2}{1539.0} \sim \chi^2(9). \end{aligned}$$

More succinctly, we have just computed

$$\sum_{i=1}^K \frac{(Y_i - E_i)^2}{E_i} = 1.5884.$$

Referencing a standard χ^2 table shows that $\chi^2(9)$ can be as large as 1.73 with 99.5% probability. Hence, the null hypothesis that these values are normal cannot be rejected (at a significance level of 0.5%). Furthermore, the corresponding p-value is 0.9964363, which further strengthens our conclusion that the null hypothesis cannot be rejected at any reasonable levels of significance.

As a final means by which to accurately assess the strength of our conclusion, we simulate a GARCH(1,1) process using the specified parameters: $\omega = .005306, \alpha = 0.054962, \beta = 0.940181$. We then test to see the likelihood that the GARCH process we simulated came from a GARCH(1,1) process with the given parameters. We do this in order to demonstrate the strength of our results concerning the goodness-of-fit of our actual data to the GARCH(1,1) process. The computer output is below.

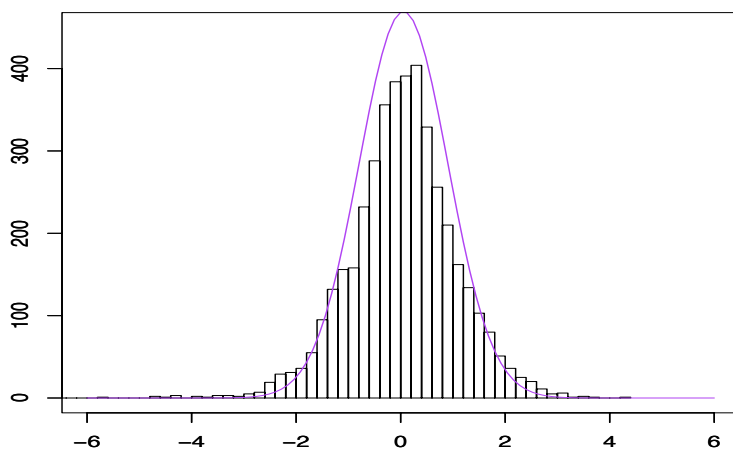


FIGURE 5.1 Histogram of returned residuals from the fitted GARCH model

Call:

```
garch(x = simgarch)
```

Model:

```
GARCH(1,1)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-3.32909	-0.69313	-0.03284	0.66887	3.27513

Coefficient(s):

	Estimate	Std. Error	t value	Pr(> t)	
a0	0.006154	0.001991	3.091	0.00199	**
a1	0.045466	0.006235	7.292	3.05e-13	***
b1	0.946528	0.007288	129.866	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 '.'

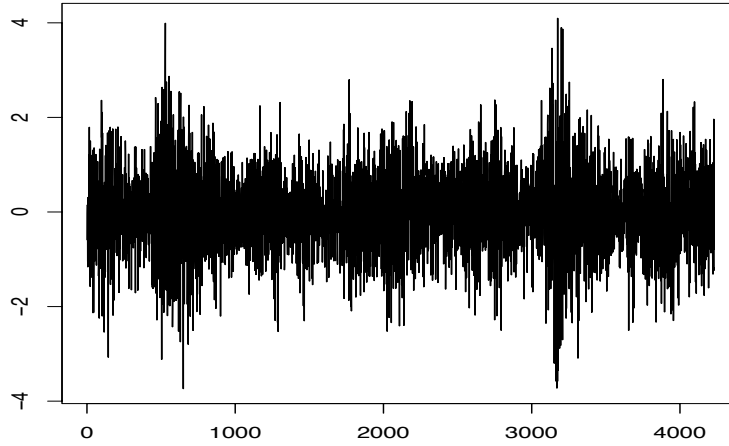


FIGURE 5.2 Simulations using $n = 4231$, $\omega = .0064298$, $\alpha = .0737144$, $\beta = .9208423$

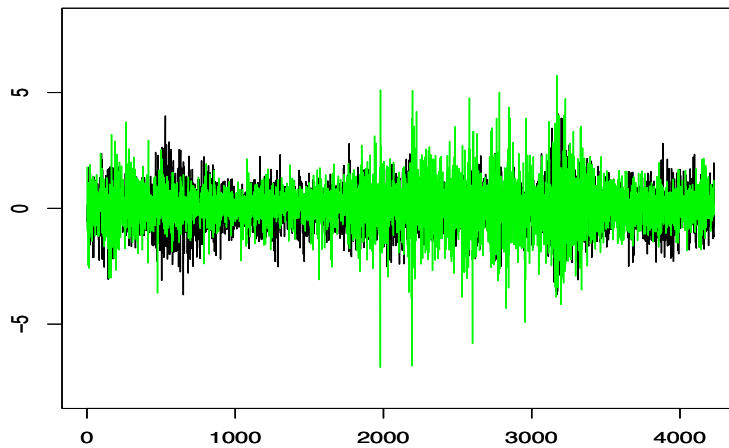


FIGURE 5.3 Simultaneous plot of simulated GARCH(1,1) process and original data

6 Conclusion

We closely examined the attributes of our chosen data set concerning financial returns. Through analysis of the visual properties, we were able to address stationarity and differencing in both the continuous and discrete cases. Particular attention was placed on the mathematical properties of various time series processes in order to determine which would provide the best fit for our data. Various combinations of autoregressive, moving average, and integrated models were considered. We carefully demonstrated the effects of slightly altering the parameters of the more common time series models through numerous simulations and discussed the autocorrelation and partial autocorrelation function and the importance thereof. We meticulously test our claim that the GARCH model with a specified set of parameters is the best fit. We do not merely accept a p-value as sole confirmation, but rather gain further confirmation through various statistical tests. These include χ^2 test for goodness-of-fit and Ljung-Box test for independence. As a further means by which to accurately assess the strength of our conclusions, we simulated a GARCH(1,1) process with the same specified parameters and equal number of observations and compare our findings. Furthermore, we superimpose the plots of our actual and simulated time series. The aforementioned rigorous analysis and statistical testing support our conclusions concerning the results.

7 Bibliography

- Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**:307-327, 1986.
- Brockwell, Peter J. and Richard A. Davis. *Time Series: Theory and Methods*. New York: Springer-Verlag, 1987.
- Chatfield, Christopher. *The Analysis of Time Series: An Introduction*. New York, NY: Chapman and Hall, 1984.
- Chung, Kai Lai. *A Course in Probability Theory*. New York: Academic Press, 1974.
- Engle, R.F. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**:987-1007, 1982.
- Fuller, Wayne A. *Introduction To Statistical Time Series*. New York: John Wiley & Sons Inc., 1976.
- Glosten, L., Jagannathan, R., and Runke, D. Relationship between expected value and the volatility of the nominal excess returns on stocks. *Journal of Finance*, **48**:1779-1801.
- Gunnip, Jon. *Analyzing Aggregated AR(1) Processes*. University of Utah: 2006.
- Hamilton, James D. *Time Series Analysis*. Princeton, NJ: Princeton University Press, 1994.
- Harris, Richard and Robert Sollis. *Applied Time Series Modelling and Forecasting*. Hoboken, NJ: John Wiley & Sons Inc., 2003.
- Horvath, Lajos and Remigijus Leipus. *Effect of Aggregation On Estimators in AR(1) Sequence*. Preprint, 2005.
- Lumsdaine, R.L. Consistency and asymptotic normality of the quasi-likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models. *Econometrica*, **64**:575-596.
- Nelson, D. B. Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory*, **6**:318-334, 1990.
- Williams, David. *Probability with Martingales*. New York, NY: Cambridge University Press, 1991.