Project

Zhikang Dong zd2241

5/9/2020

Introduction

In this project, we would like to analyze JPMorgan Chase&Co stock price stationarity and volatility. Then based on these results, we predict price log returns and perform risk management related to Nasdaq 100 index.

The two data sets we used here are:

Nasdaq 100 index: Weekly open price from 2010.01.01 to 2020.01.01

JPMorgan Chase&Co: Weekly open price: from 2010.01.01 to 2020.01.01

Stationarity

First, we'd like to check the stationarity of JPMorgan Chase&Co stock price. We plot original open price here, and find there is an upward trend in the plot, which means the stock price may not be stationary.

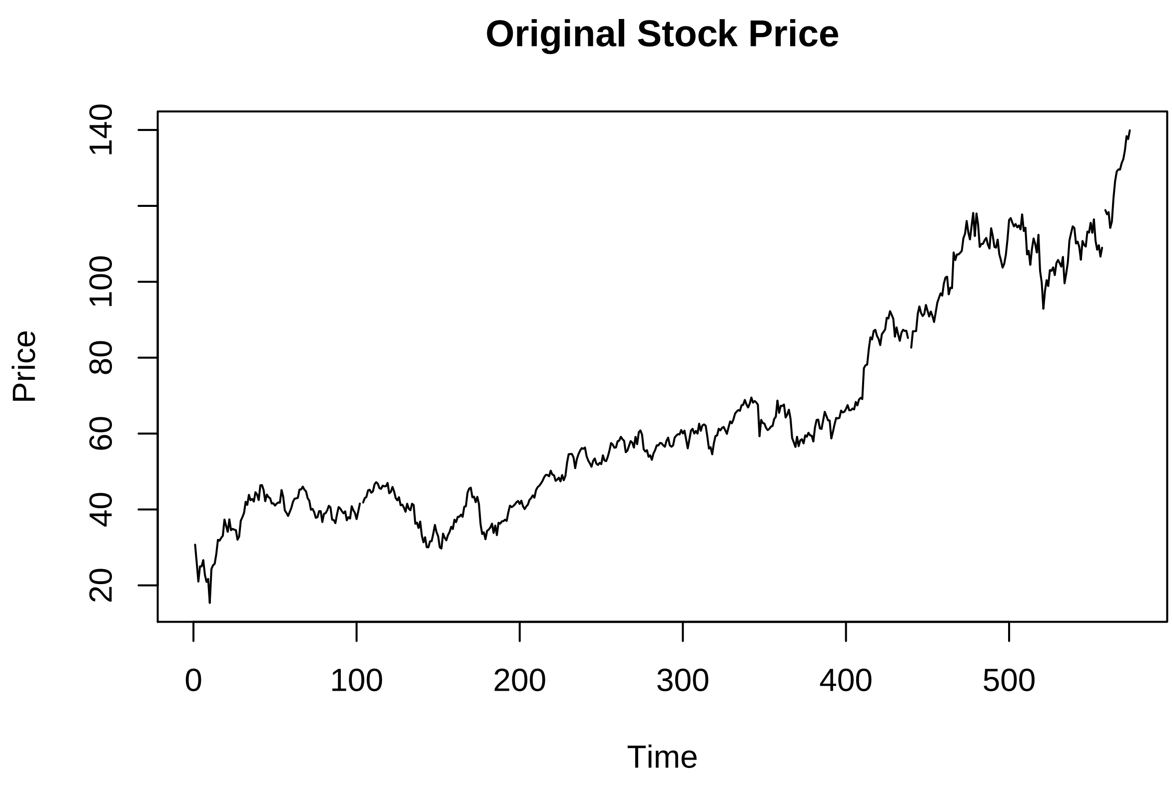

Figure 1: Original Stock Price

Then we plot autocovariance funtion and autocorrelation function of this time series.

From ACF plots, we see that the acf decreases when lag grows, we know that the time series is non-stationary[1]. Also, we don't see seasonality from the plot.

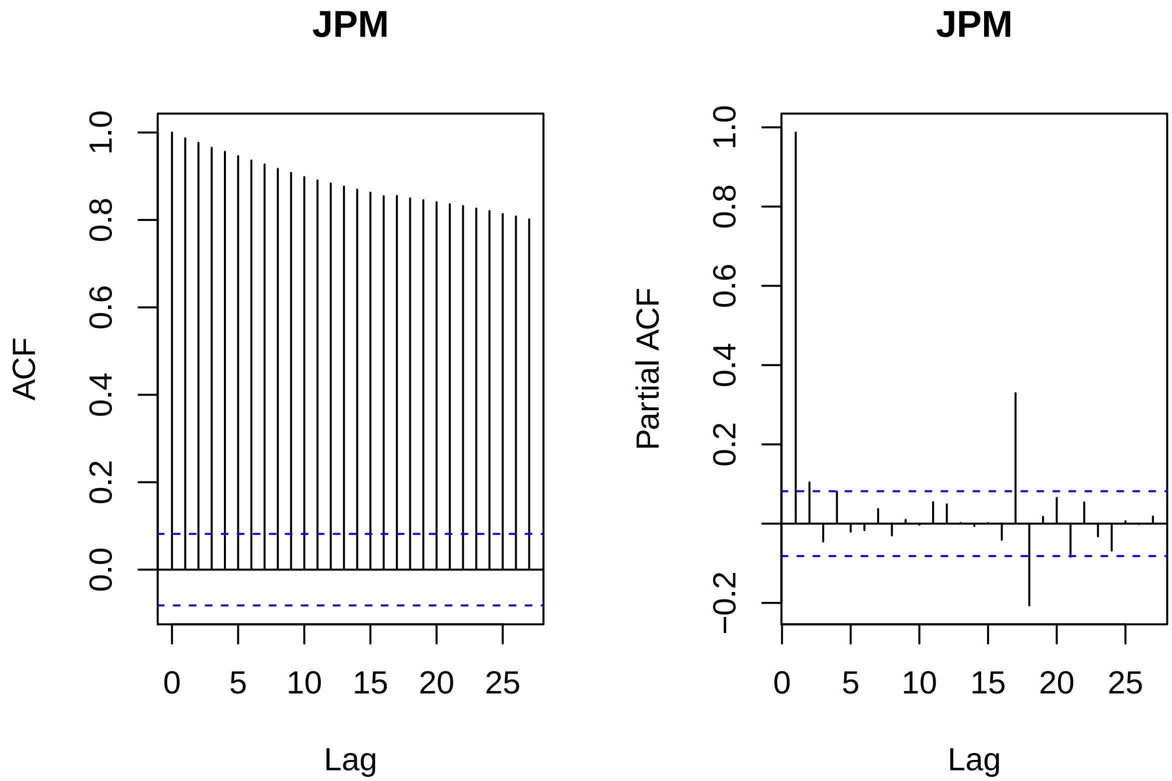

Figure 2: ACF and PACF of original stock price

In order to obtaining a stationary data, we make a difference with d=1 and plot ACF again. Roughly speaking, the differenced data is stationary without seasonality.

To make sure our assumption, we fit the differenced data with SARIMA model[1].

```
## Series: jpm_diff
## ARIMA(2,0,1) with non-zero mean
##
## Coefficients:
##           ar1      ar2     ma1    mean
##       -0.8494  -0.0386  0.7509  0.1786
## s.e.   0.1895   0.0534  0.1838  0.0848
##
## sigma^2 estimated as 4.77:  log likelihood=-1245.48
## AIC=2500.97   AICc=2501.07   BIC=2522.72
```

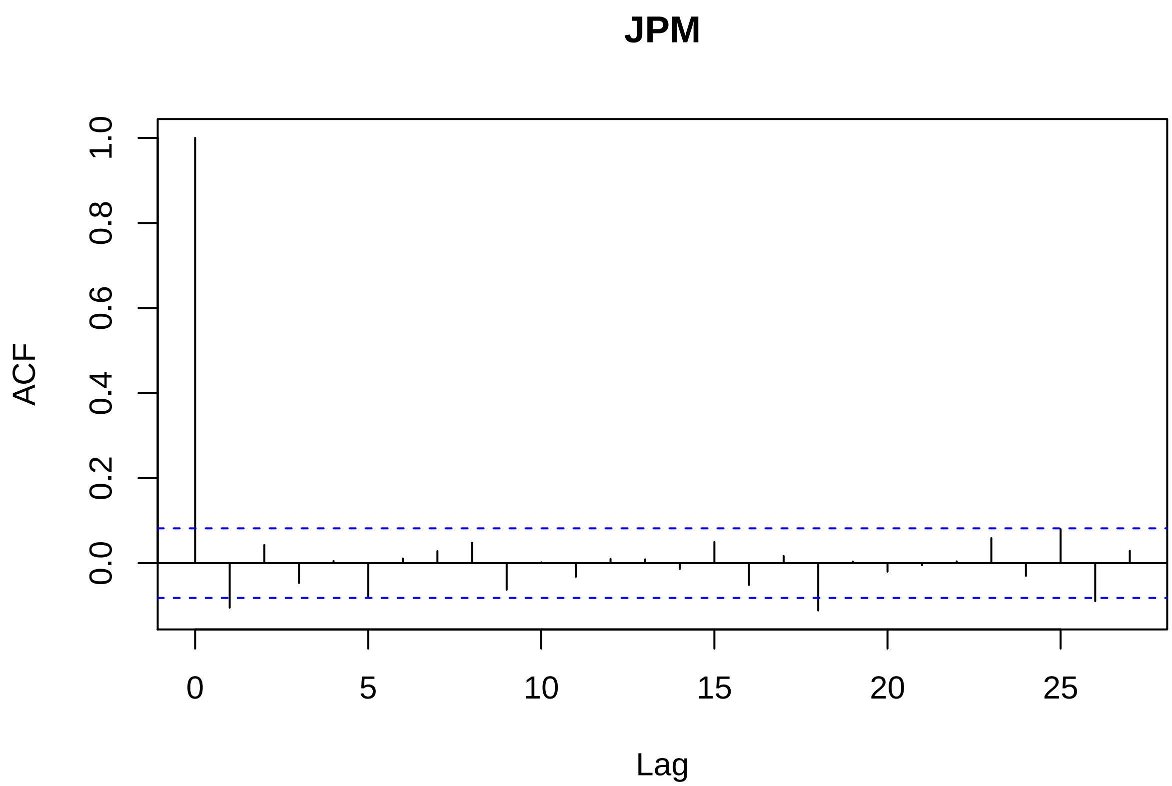

Figure 3: ACF plot of differened price

From the output, we can make sure that the differenced data is stationary, non-seasonal time series.

Next, we check the goodness of fitting differenced data into ARMA(2,1) model.

Intuitively, we can plot diagnostic figures (Figure 4) for residuals. We may assume that residuals is a white-noise process. To support our assumption, we preform Ljung-Box test, and find that p-value is .4766 which means the Ljung-Box test fails to reject the adequacy of the model. We may think that the fitted model is approriate[1].

```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(2,0,1) with zero mean
## Q* = 6.5556, df = 7, p-value = 0.4766
##
## Model df: 3.   Total lags used: 10
```

Volatility

We would like to estimate volatility of log-returns of JPM. And log-returns can be computed as:

$$r_t = log(\frac{P_t}{P_{t-1}})$$

Then we just pass missing data, and get log-returns data:

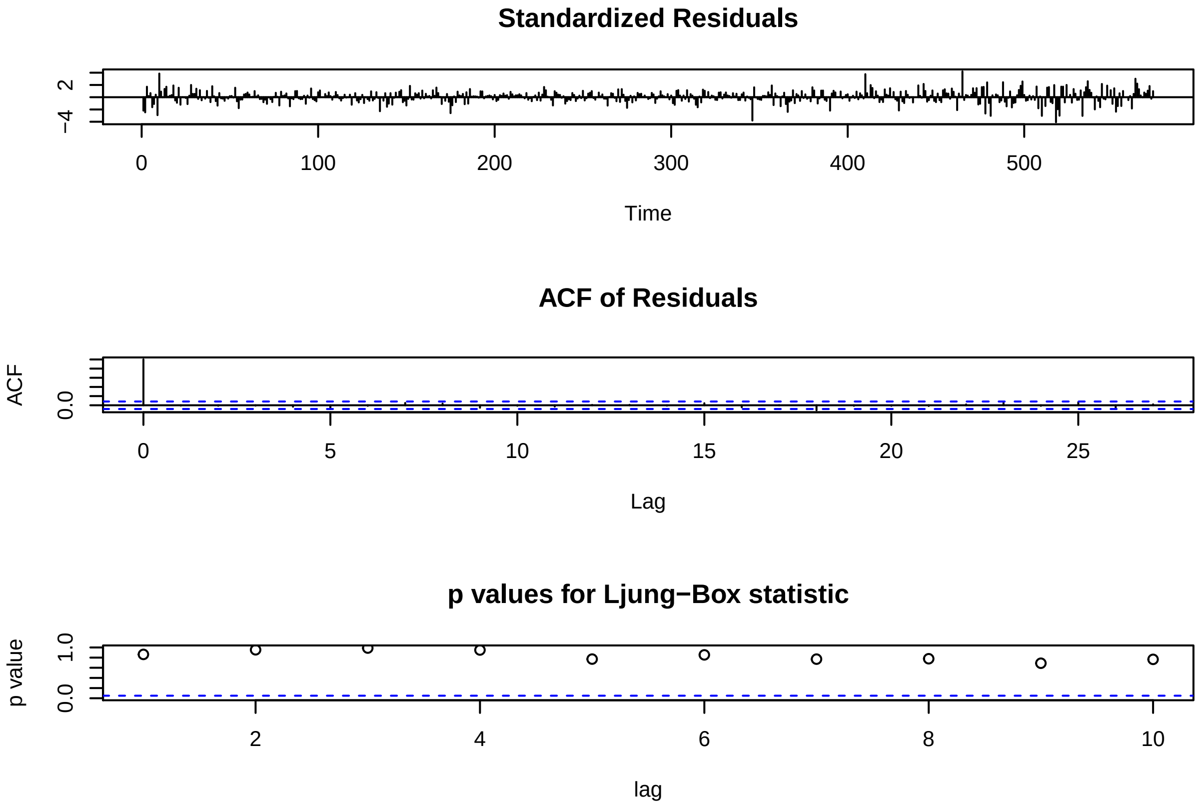

Figure 4: Goodness of fit

Here we choose GARCH(1, 1) model to fit the log-returns data. Let's check if the residuals follows normal distribution (Figure 5)

From the plot we see heavy tails exists, for better fitting, we choose skewed studentized-t distribution as conditional distribution and fit the data again:

It seems like this model fits log-returns better, and we derive coefficients:

```
##           mu        omega       alpha1        beta1         skew        shape
## 3.018055e-03 8.101332e-05 1.043777e-01 8.346767e-01 9.095231e-01 5.144113e+00
```

More precisely, performing several test on residuals and residuals square, we have results:

			Statistic	p -Value
Jarque-Bera Test	R	Chi $\wedge$ 2	227.8313	0
Shapiro-Wilk Test	R	W	0.9642111	1.596761×10^{-10}
Ljung-Box Test	R	Q(10)	11.96672	0.28729
Ljung-Box Test	R	Q(15)	13.11482	0.5934297
Ljung-Box Test	R	Q(20)	19.0799	0.5166382
Ljung-Box Test	R^2	Q(10)	11.83752	0.2960776
Ljung-Box Test	R^2	Q(15)	14.2206	0.5088688
Ljung-Box Test	R^2	Q(20)	18.40313	0.5608705
LM Arch Test	R	TR $\wedge$ 2	5.538428	0.9375422

We can know that p-value is relatively large in Ljung-Box Test, which means we fail to reject the adequancy of model. We have a good fitting here.

After deciding coefficients of the model, we can plot volatility of log-returns(Figure 6):

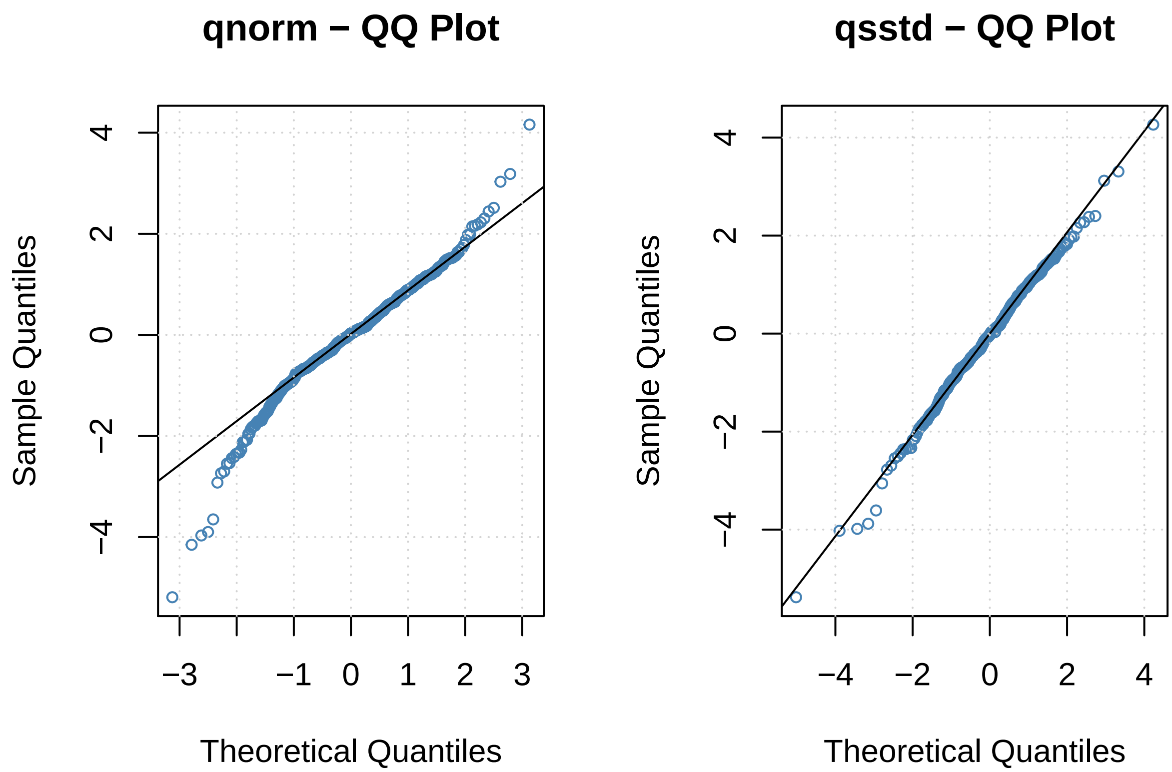

Figure 5: Normal Distribution and T distribution QQ-Plot of Standardized Residuals

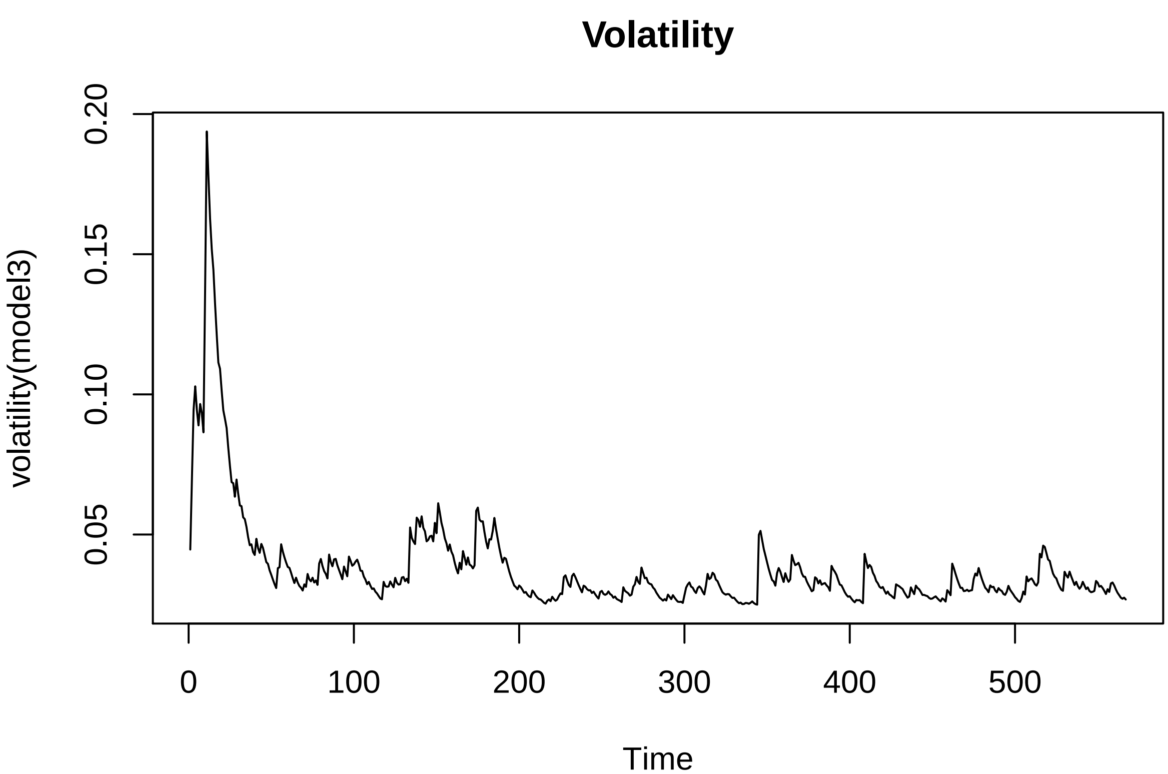

Figure 6: Volatility of log-returns

Prediction

We can also combine ARMA model and GARCH model to predict future log returns, which ARMA can predict expected mean and GARCH can predict expected variance[1].

First, we have to decide coefficients of ARMA model by using auto.arima function in R. Then we have an AR(3) model with coefficients:

```
##          ar1          ar2          ar3    intercept
## -0.180271237 -0.008606937  0.010249079  0.002607148
```

Then we combine AR(3) model and GARCH(1, 1) model. Again, we use skewed t-distribution as conditional distribution.

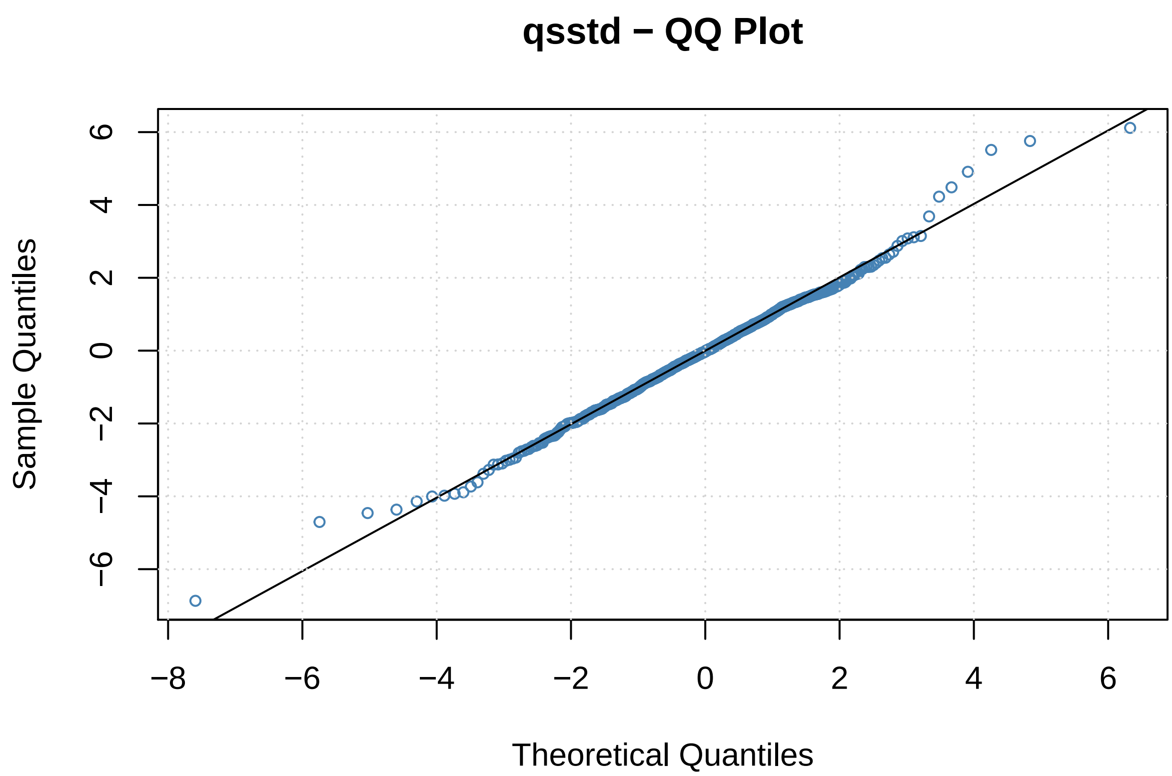

Figure 7: T-distribution QQ Plot

Based on this model, we can predict two-steps ahead log-returns.

meanForecast	meanError	standardDeviation
0.0134124	0.3671459	0.3671459
-0.0244506	0.3719956	0.3718716

Time-varying Beta

In this section, we would like to analyze risk of JPMorgan Chase&Co related to Nasdaq 100 index in last 10 years.

CAPM (capital asset pricing model) is one of the most common financial model. People use this model to establish the portfolio and estimate returns and market sensitivity. Here, we can use GARCH model to find betas (stock sensitivity) of the stock in different time.

We have the CAPM model like this[2]:

$$r_t = \alpha + \beta r_{m,t} + e_t, \quad t = 1, \ldots, T$$

where α (Jensen index) means the mispricing of the stock compared with the market.

Generally, if β is significantly greater than 0, which means that the stock responds aggresively to the market. On the other hand, if β is relatively close to 0, then the market doesn't have much impact on it. Thus $\beta < 1$ is regarded as less risky than the market, and $\beta > 1$ indicates a high risk investment.

In practice, we would like to see an asset outperform the market with less risk. Mathematically, $\alpha > 0$ and β is small.

For the CAPM model above, we have[3]

$$\hat{\beta} = \frac{\operatorname{Cov}(r_t, r_{m,t})}{\operatorname{Var}(r_{m,t})}$$

where r_t and $r_{m,t}$ are the log-return of the stock and the index we choose at time t.

By fitting a good GARCH(1, 1) model, we can easily get volatility of the stock and the market[2].

Here we can use

$$\operatorname{Cov}(r_t, r_{m,t}) = \frac{\operatorname{Var}(r_t + r_{m,t}) - \operatorname{Var}(r_t - r_{m,t})}{4}$$

to obtain the result.

Firstly, we consider β in traditional CAPM model[4].

r_t	beta
0.0011153	0.4368749

Then we calculate time-varying betas and plot them:

From the plot, we know that time-varying plot can reflect risk more accurately.

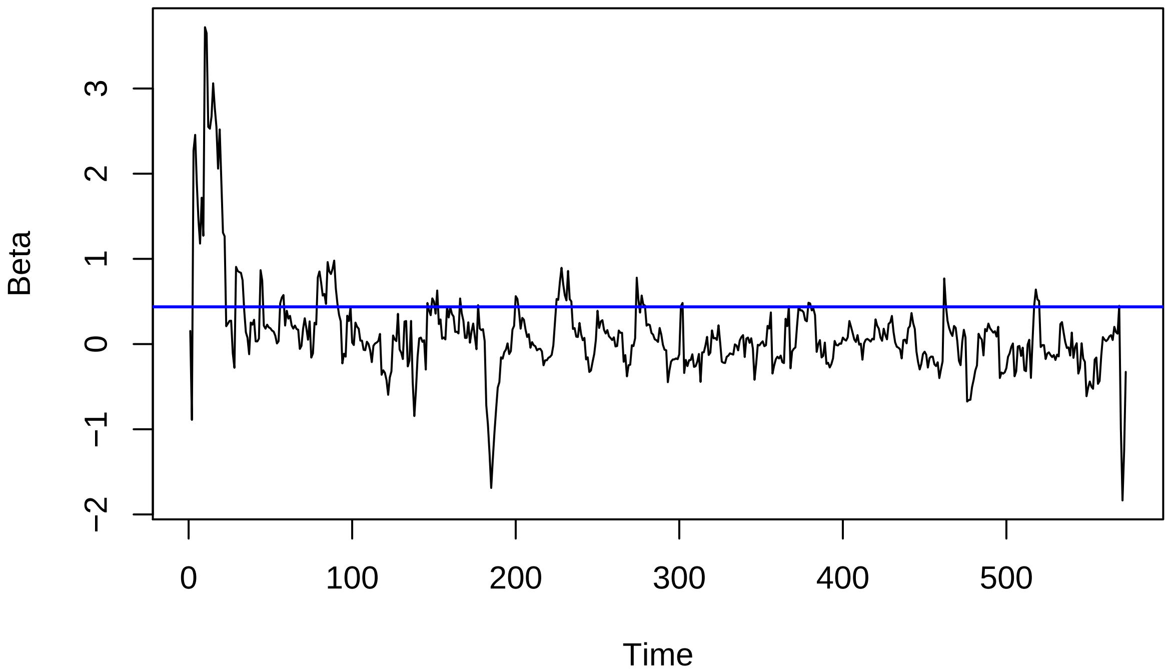

Figure 8: Time varying betas and blue line represents traditional beta

Conclusion

In this report, we find that JPMorgan Chase&Co. stock is an non-stationary and non-seasonal time series. After making difference, we can obtain a stationary process. We also show the volatility of the differenced data, which help us to dig out more information. Then we fit data to ARMA model and GARCH model and perform several important hypothesis testing to support our assumptions. We find GARCH(1,1) model explains the data well. We also use ARMA-GARCH model to find future price of the stock. At last, we calculate time varying betas of JPMorgan Chase&Co., we find although the stock outperforms the market a little, it has a very low risk, the price won't be fluctuate a lot.

Future works

In the prediction section, the model we used here is still naive, and the prediction is not perfect. In future works, we may optimize our model and introduce more coefficients to help us make a good prediction.

References

[1] Robert H Shumway, D. S. S. (2016). *Time series analysis and its applications with r examples.* Springer.

[2] Tsay, R. S. (2012). *An introduction to analysis of financial with r.* WILEY.

[3] Morris H DeGroot, M. J. S. (2011). *Probability and statistics.* Pearson.

[4] Trevor Hastie, J. F., Robert Tibshirani. (2009). *The elements of statistical learning.* Springer.