Introduction

"Qianjiang Pharmaceutical" is the largest ophthalmic drug production base in China. With its amazing achievements during recent years, I am interested in investigating this stock. In this study, I describe how we collect the data; analyze the data and my conclusions, recommendations.

Methods and Materials

There are two data sets. Data1 is the daily closing price of Qianjiang stock from December 1 2003 to April 22 2004 during the trading days. "Closing price" generally refers to the last price at which a stock trades during a regular trading session. Data2 is daily Stock Market Index in Chinese Market as the same period as data1.

First, I differenced once of data1 according to the suggestion from ACF. The ACF and PACF of the differenced data tell us that the behavior of the stock price is a random walk.

Next, I am interested in the relationship between individual stock and stock market index. I then looked at the CCF of these two data sets.

Lastly, I fit a model of the index data in order to see the overall trend of the Stock Market.

All data analysis was conducted using ASTSA.

Results

Random Walk Model

Random walk theory is a basic theory of stock price behavior, which predicts that stock price behave as a random walk. I'm interested in the behavior of Qianjiang Pharmaceutical stock. The ACF of data1(Appendix1), exhibits a slow decay as lags increase. It is suggested we might try looking at a first difference and comparing the result to a white noise or completely independent process.

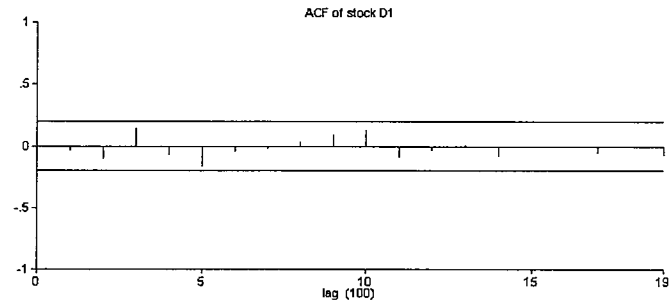
The first difference of data1 appears to be stationary (Appendix2), although an possible outlier is observed. The ACF and PACF of the differenced data1 (Graph1) reflect this predicted behavior, with no significant values for lags other than zero. Qianjiang Pharmaceutical Stock is a Markov Process, behaves like a random walk as described below.

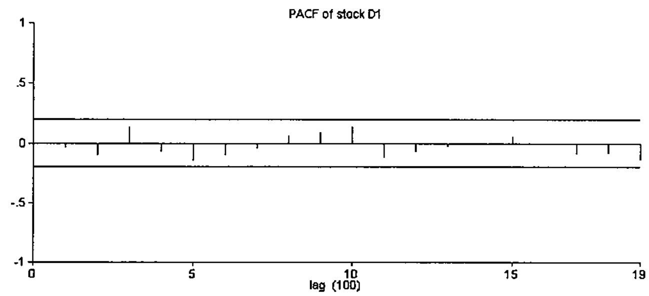
$$\mathbf{x_t=x_{t-1}+w_t}$$

where w_t is a white noise with mean zero and variance σ^2.

Consequently, we have the following equation $E(x_{t+k}|x_t)= x_t$, which makes the forecast no sense. So we will not talk about forecast for data1.

Graph1: ACF and PACF of data1 (stock price)

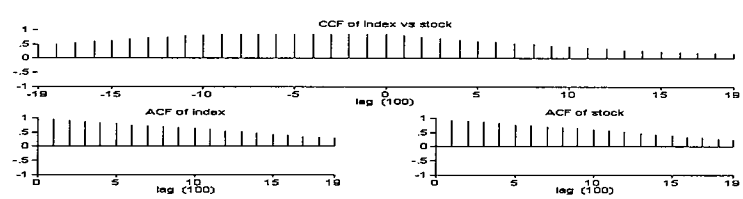




Relationship between Two Series

The CCF of the two series (data1 and data2) (Graph2) does tell us something. When the lags are negative, CCF is larger than those of positive lags. But they are always positively correlated. This sounds quite reasonable to me. Generally, with the increasing of index, individual stock will increase too. But I'd like to divide the stock into two groups. One is popular and huge market value stock. Another one is unpopular and small size stock. Since Qianjiang Pharmaceutical is not a popular stock and with a small size, its reaction to the market generally will lag behind the reaction of index. So when the lag is negative, accordingly CCF will be larger.

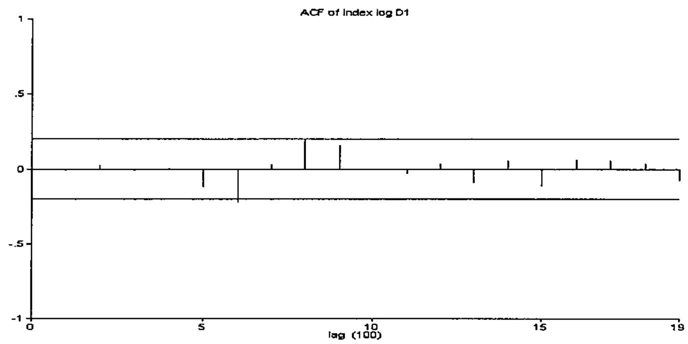
Graph 2: CCF of the two time series (Index vs.Stock)

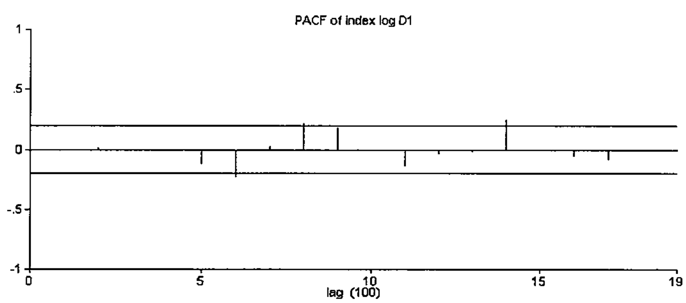


Model Building of Index Data

This section, I am interested in using data2 to predict the overall trend of the stock market. As a result, it's possible to help us to investigate the behavior of our stock. Model building is the first step. The process of index evolves as a fairly small and stable percent change. So it's suggested return (first use log transformation, then take first difference) be a relatively stable process. The graph of return shown in Appendix 3 looks stationary. Then we looked at the ACF and PACF of return (Graph 3),

Graph 3: ACF and PACF of return of data2





Slightly significant values are shown at lag 6 for both ACF and PACF. Also other slightly values occur at other lags of PACF. I doubt this is due to the similar trend of the investor's behavior during Friday and the previous Friday, Monday and previous Monday. Since if the index is increasing, investors expect a promising market, they wait and think it over and can't reject the allurement, then they will buy the stock at the last trading day of the week, Friday. After a weekend, another group of people will do the same investment next Monday. On the other hand, if the index is decreasing, the loosing party can't endure the panic and will sell their stock on Friday or next Monday. Generally, the behavior of index during the two continuous weeks is similar. So the value of lag 5 should be significant in practice. But our data hasn't shown this phenomenon. It may due to the fact that there are Spring Festival holiday and other holiday during the period I collected the data, so there may be only two, three or four trading days in some certain weeks. It is a shortcoming of my data. But I do believe significant value should occur, so model building is required.

ACF and PACF suggested the degree of AR may be between 6 to 14; the degree of MA may be 6. I use the ARIMA search for the best model. The result is ARIMA(6,1,6). The residuals look quite good. But since differencing may introduce dependence where none exist. We tried several models, such as ARIMA(6,0,6),ARIMA(5,0,8) and so on. I use residuals test, Ljung Box and forecast confidence interval to compare these models. Finally, I decided to use ARIMA(6,0,6) to fit index return model, equivalently, ARIMA(6,1,6) to fit log(index).
The model is described as the following

(1 -.12B1 +.09B2 -.05B3 -.15B4 -.13B5 -.25B6) (D1) x(t) =
(1 -.26B1 +.49B2 +.26B3 -.02B4 -.35B5 -.56B6) w(t)

All of the coefficients are significant.

The residual plot (Appendix4) shows no obvious pattern. No outliers are detected. ACF of the residuals (Appendix5) shows no significant values. In this case, all of the autocorrelations are within the white noise limits.

The Ljung-Box-Pierce statistic, for H=20, was Q=16.10, p-value shown in the result indicate no significance, which supports the white noise hypothesis.
We note that the empirical results for the residuals are very close to what would be expected for white noise. Q-Q plot of the residuals indicates a strong linear relationship, and the correlation between e_j and q_j is about .99479; this result strongly supports the null hypothesis of normality. Histogram of the residuals also supports the normality of the residuals. These three additional diagnostic graphs appear in Appendix 6.

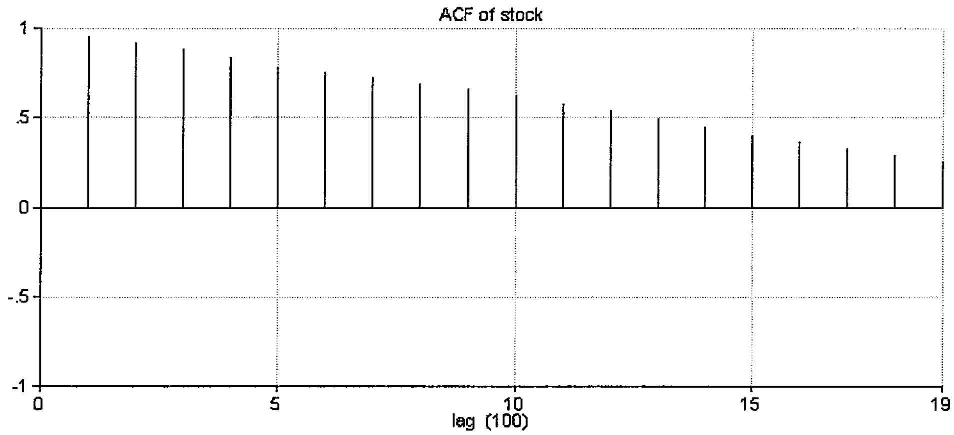
Forecast is always what we are interested. The result (Appendix 7) indicates that there is an increasing trend of index during the next week, then with very slightly changes. The confidence interval during the 20 days is about [1436.55,1900.74]. Since index data has a fairly small and stable percent change During 20 days, it's hard for me to see an overall trend.

Conclusions and Recommendations

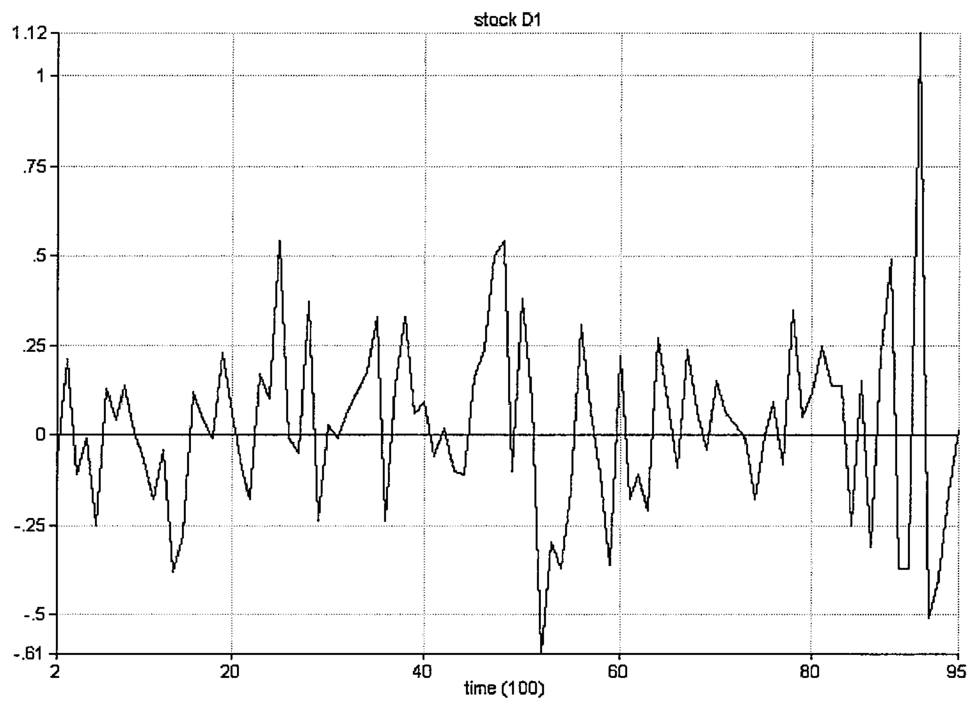
From my study, the behavior of Qianjiang Pharmaceutical stock is a random walk. Since it's not a popular stock and with a small size, its reaction to the market generally will lag behind the reaction of index. I also build a model for the log (index), but unfortunately, the forecast doesn't look very meaningful to me.

There still have many interesting things for us to discover. For example, if we choose a popular and huge size stock, will it behave like a random walk? If we choose the average weekly closing price to be the response, do we have the similar results? Future study may collect more data points for stock market index to predict the overall trend of the market. Due to the limited knowledge of financial time series, the model is still far away from practical operation. Also, there are many other factors that may affect the stock and index behavior, such as the government's policy, economic situation and so on. So we need to be very careful when we invest a stock.

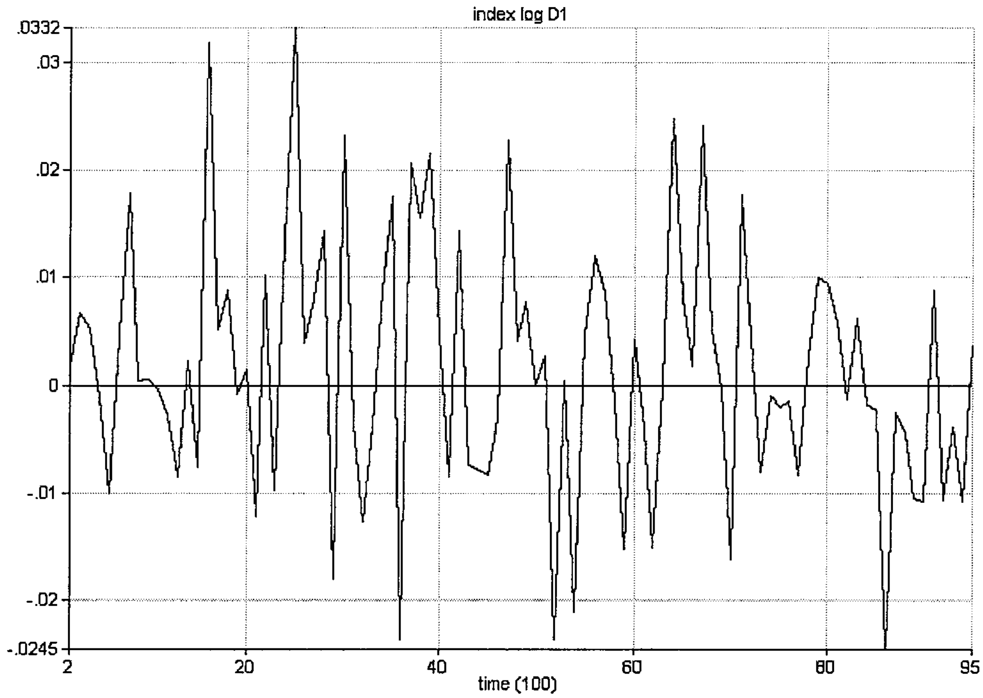
Appendix 1: ACF of stock



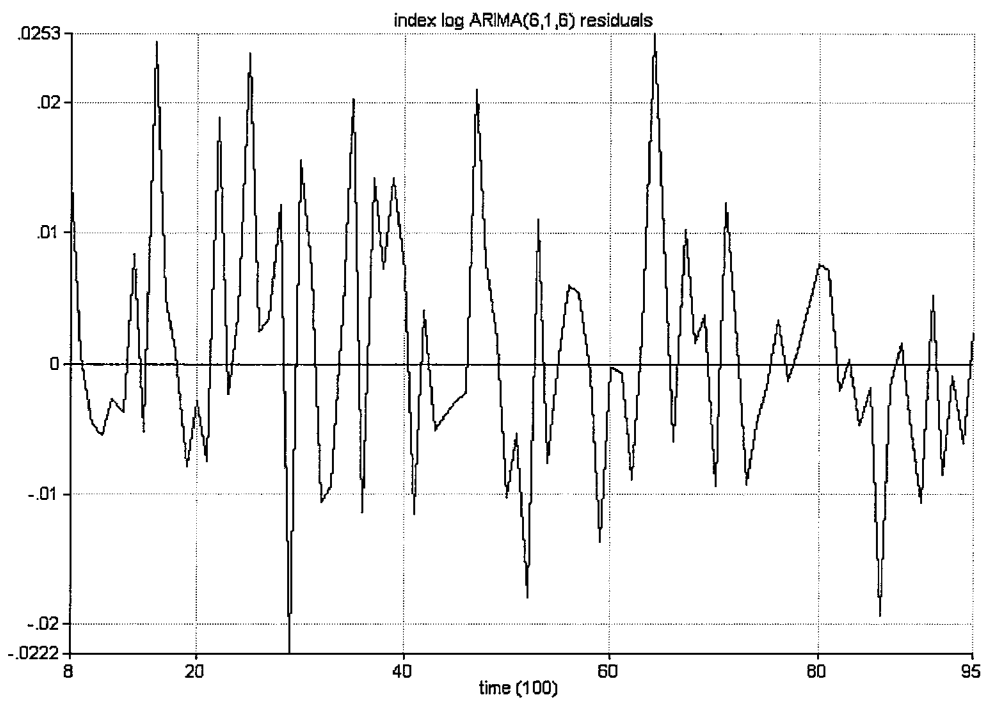
Appendix2: graph of the first difference of stock data



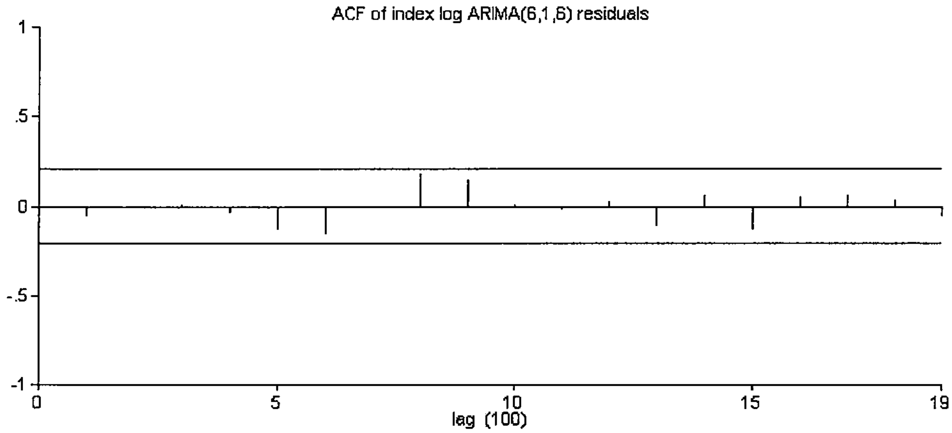
Appendix 3: Graph of return of index



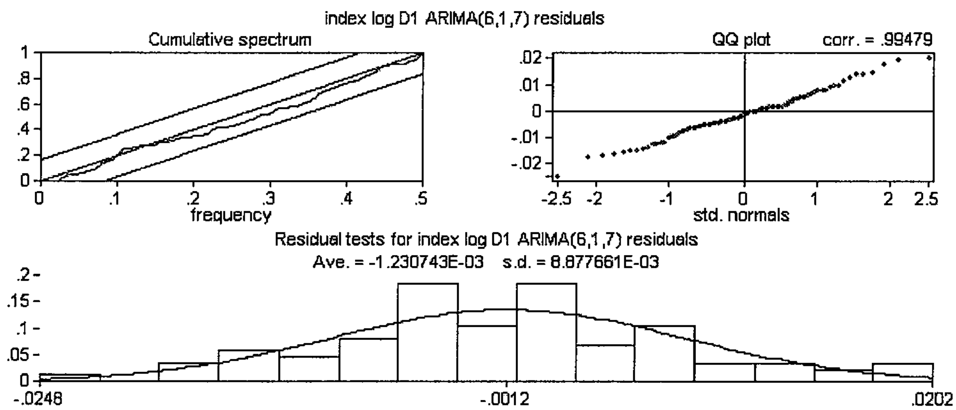
Appendix 4: residuals of ARIMA(6,1,6) model for log(index)



Appendix 5: ACF of index log ARIMA(6,1,6) residuals



Appendix 6: Additional residuals diagnostic graphs



Appendix 7: Forecast of index during the next 20 days

Forecasts for ARIMA(6,1,6) from index log

		95% Conf. bound		
		95% Conf. bound		
t	forecast	lower	upper	S.E.
96	7.4118	7.3916	7.4319	.0103
97	7.4106	7.3761	7.4451	.0176
98	7.4117	7.3585	7.4650	.0272
99	7.4126	7.3468	7.4785	.0336
100	7.4127	7.3373	7.4880	.0384
101	7.4118	7.3240	7.4997	.0448
102	7.4114	7.3189	7.5039	.0472
103	7.4114	7.3150	7.5079	.0492
104	7.4118	7.3126	7.5111	.0506
105	7.4120	7.3106	7.5133	.0517
106	7.4118	7.3077	7.5159	.0531
107	7.4115	7.3055	7.5175	.0541
108	7.4115	7.3028	7.5201	.0555
109	7.4115	7.3001	7.5230	.0568
110	7.4116	7.2975	7.5258	.0583
111	7.4116	7.2943	7.5289	.0599
112	7.4115	7.2915	7.5315	.0612
113	7.4114	7.2885	7.5344	.0627
114	7.4114	7.2858	7.5371	.0641
115	7.4115	7.2833	7.5397	.0654