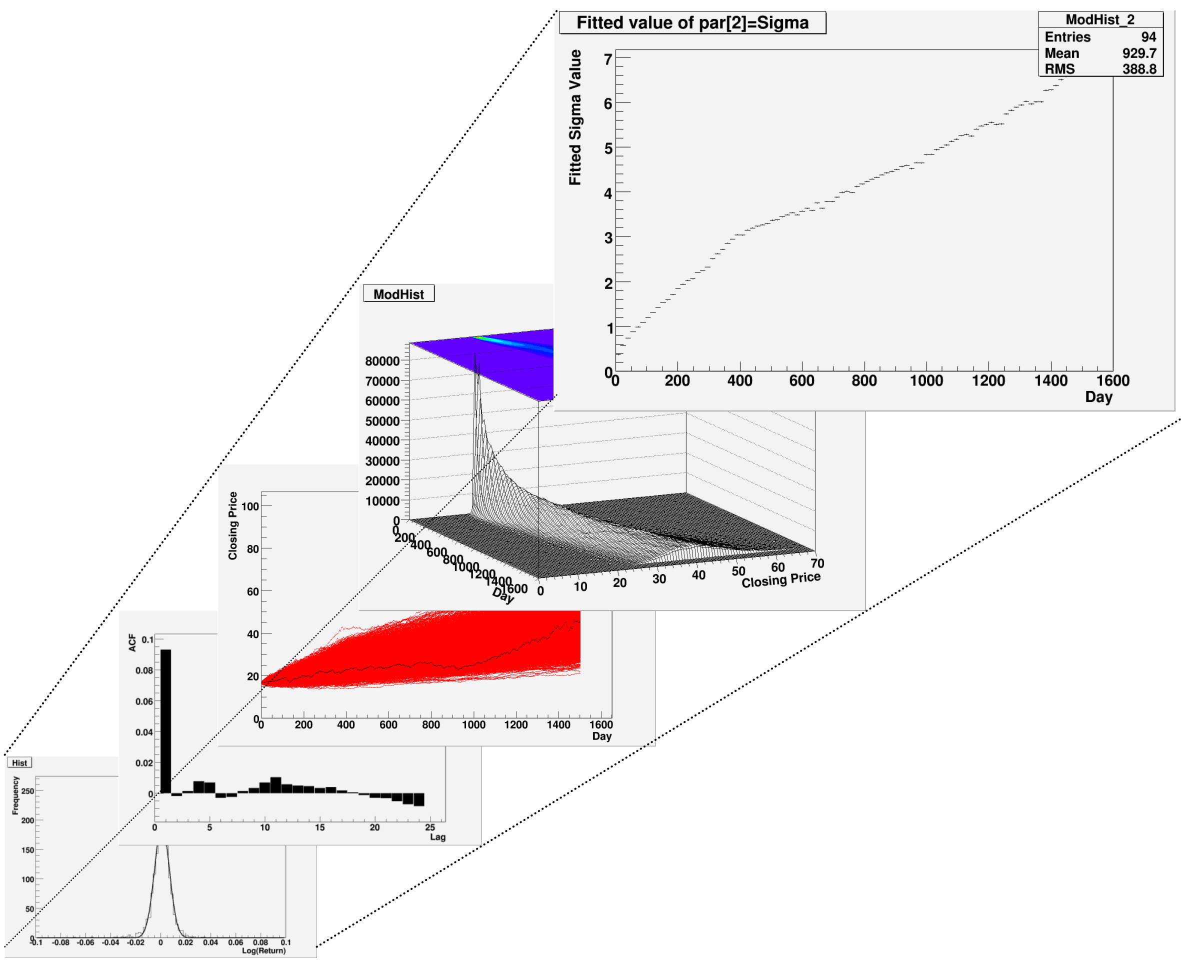
Analysis of the S&P 500 Closing Price as a Time Series

Eric D. Johnson, PhD

NEAS Times Series Analysis Course

Summer 2012



Introduction

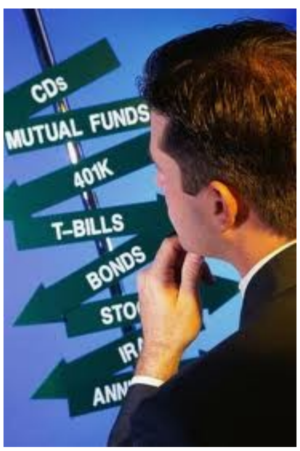
When people at cocktail parties hear that I work in the financial services industry I often end up in a conversation that goes something like this:

People: What should I invest in?

Me: What do you mean by invest?. In the simplest sense investing means to expend something, be it money, time or knowledge, with the expectation of replacing that expense with something of greater value in the future. Everyday investments are made in homes, businesses, education, and perhaps most frequently in financial products.

People: Oh, I want to invest in financial products.

Me: Great. Do you require income, capital retention, growth, maybe several or all of the above?



People: Huh?

Me: If you need to keep your money safe you can buy CDs, Treasury Bonds, some Corporate Bonds. For growth stocks and mutual funds are best bet.

People: Right, I want to invest in the stock market.

Me: Great. What is your investment philosophy?

People: I want to make money.

Me: But how? Using a companies financials for information, relying on an efficient market, or maybe using high frequency trading.

At this point I usually get interrupted because the person has something very important that requires there immediate attention. Fortunately, this is exactly the result I'm looking for because there are way too many variables involved in making an investment decision for me to give good advice at a cocktail party.

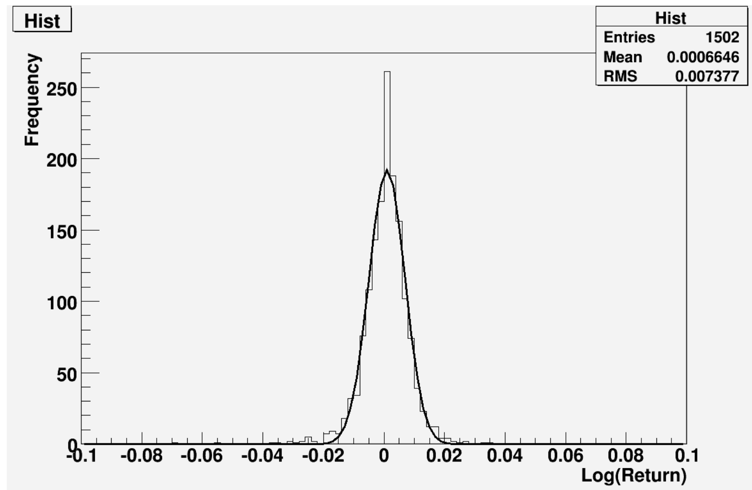
However, some people become very intrigued by the different types of investment philosophies and no longer care about advice but want to learn more. I am perfectly willing to indulge those people since it is a topic that I love to discuss.

This brings me to the topic and purpose of this report. Using the fundamentals of time series analysis I plan to show that the efficient market theory works but not very well. This argument is based on the fact that if all of the information needed to know tomorrow's stock price is accounted for in the current market price, then the current and past market prices can be used to accurately predict the future prices.

The key to whether or not the previous statement is true is in how one defines accurately.

Analysis

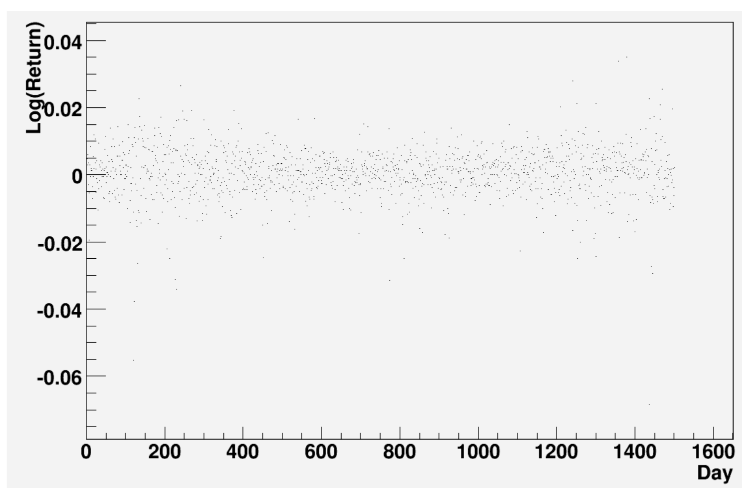
Historical price data for the S&P 500 index was downloaded from Yahoo Finance (http://finance.yahoo.com/q/hp?s=%5EGSPC+Historical+Prices). Price data is available from 1950 until the present. A time series analysis was performed on a subset of the data that started in 1950 and ended in 1979.



Analysis of the data was performed using the statistical analysis software ROOT (root.cern.ch). The first stage to the analysis process was creating a histogram of the log of the daily returns of the closing price. That histogram was then fit to a Gaussian function using a χ^2 minimization (see right).

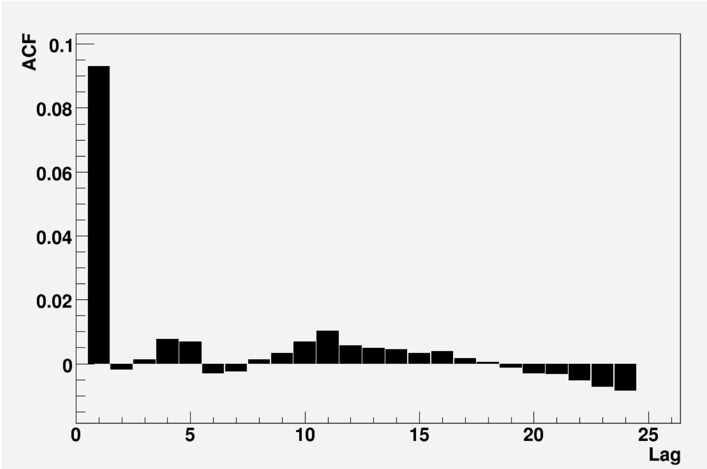
The parameters of the Gaussian function were used in determining randomized parameters within the time series. The fit parameters are given in the table below. The lower chart shows the log of the daily return as a function of time. From the charts on this page it is clear that the return is centered very near 0%, with the majority of data points falling between ±1%.

	Value	Error
Amplitude Constant	1.92E+02	6.97E+00
Mean	9.89E-04	1.58E-04
Standard Deviation	5.94E-03	1.48E-04



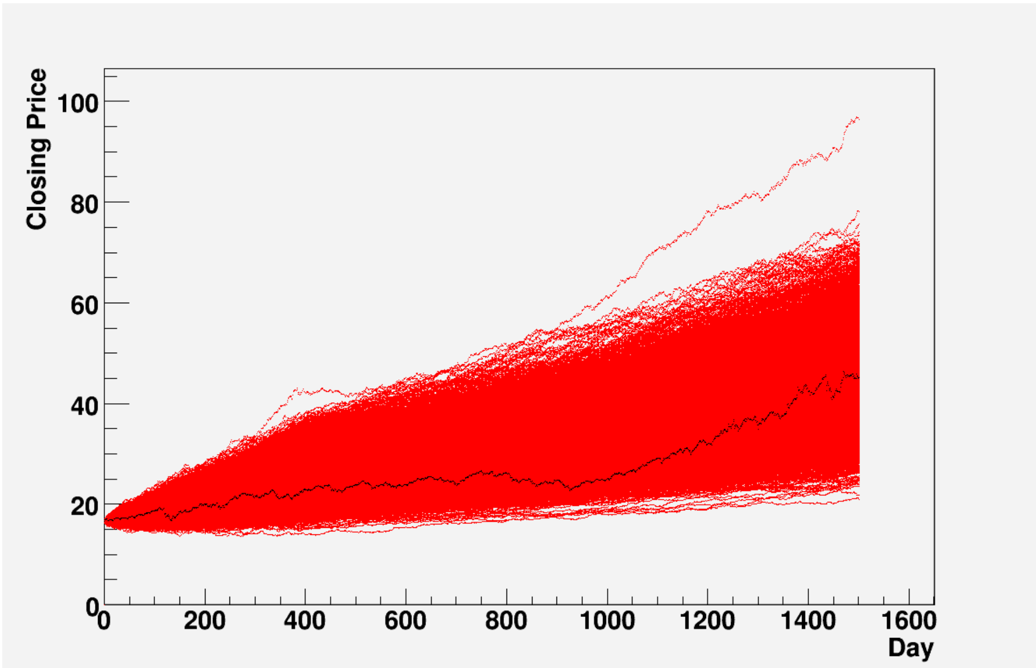
Analysis

The next step was to plot the auto-correlation function (ACF) for the closing price at various lags in order to determine what type of function to use to replicate the S&P 500; and also how many lags would be required. The figure to the right shows the ACF for lags 1 through 25. The pattern shown indicates that the times series should be modeled using an AR(1) function.

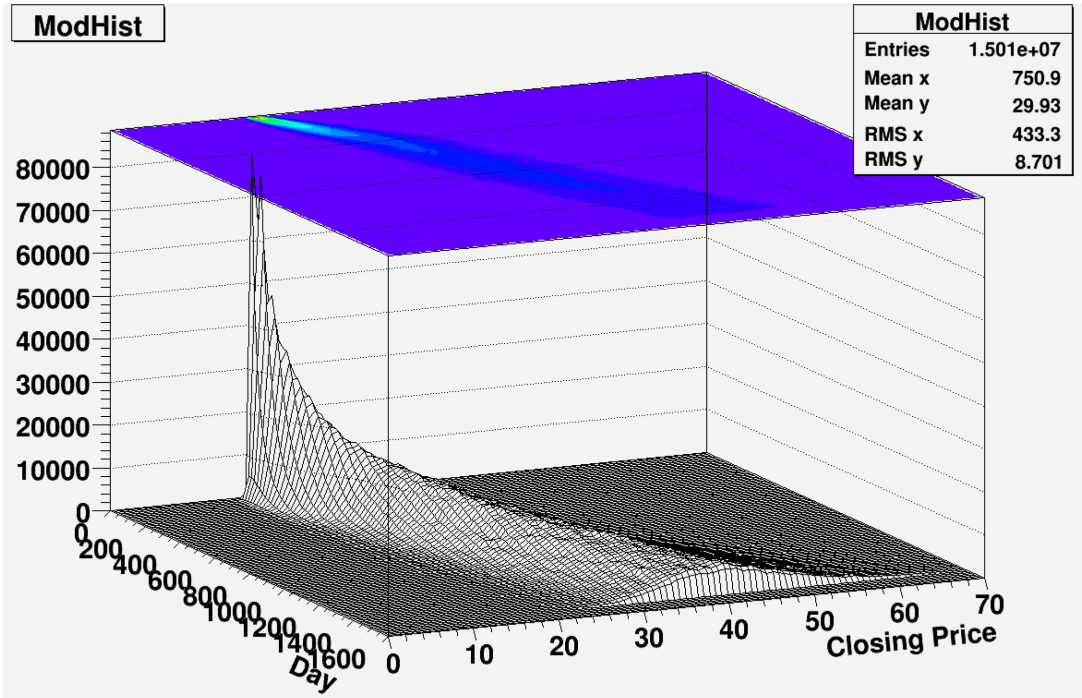


The return of the index was modeled using an AR(1) function with the stochastic term modeled using a random Gaussian distribution. The mean of the distribution was set using the mean parameter from the fit of the log-return histogram. The standard deviation parameter was also used; however, it was randomized using a uniform distribution with values between 1.0 and 1.5. The return value was also multiplied by a time-dependent scaling factor of 0.5 for T>400.

The AR(1) time series was produced 10,000 times in order to determine the range of possible values which the model would give. The chart below shows all the series (in red) along with the actual values of the S&P 500 stock index. The chart at the top of the next page shows a histogram of the modeled data points with frequency along the z-axis. From these 2 charts it is clear that the further one attempts to project into the future the less certain the projection becomes.



Analysis and Conclusion



The final step in the analysis process was to create daily slices of the histogram above and fit those slices using a Gaussian function in order to determine the standard deviation of the projection as a function of time. The chart below shows the standard deviation of the 10,000 projected time series for each day in the projection. As noted earlier the standard deviation and therefore the uncertainty in the projection increases with time.

The chart also shows that after projecting less than one year the uncertainty in the projected closing price has become greater nearly 10% of the actual value. This large of a fluctuation in the closing price of the index indicates that other forces may be involved in driving the market value.