HOMEWORK 1

Please submit your homework to xm@bu.edu. Don't forget to attach your figures and code. Feel free to ask me if you have any question. GLHF! -Sean.

Problem 1: simple statistics, PDF, and time series

For this problem you will need the daily closing prices of S&P 500 index from 01/03/2006 to 12/31/2009 (1007 data points).

1. Make a list plot of S&P 500 index prices. Can you tell from the graph when the latest global financial crisis happened? You might have become a millionaire if you had sold short in the stock market at the right time!

2. Make a list plot of the log returns (differences of daily log prices) of S&P 500 index. Can you tell from the graph when the volatility of U.S. stock market reached local minimum? Local maximum?

3. Draw a histogram of the log returns. Does it look like a Gaussian distribution? Calculate the expectation and variance of the log returns and use them to fit your data with a Gaussian PDF (probability density function). Make a plot to compare your fit to the real PDF (Be aware that a PDF must be normalized). What do you find? Now make a log plot to confirm your findings.

4. Create a list of 1006 random data points sampled from the Gaussian PDF. Make a list plot of the data points and compare it to what you just plotted in Question 2. Can you tell the difference? Which one is heteroskedastic (exhibits time-varying volatility)?

5. A collection of random variables can usually be defined as a stochastic time series if indexed by time. The set of random variables $\{X_t|t = 1, 2, 3\ldots\}$ you created in Question 4 is an important stochastic process called **white noise**. Now, we introduce a new set of random variables $\{Y_t = \sum_{s=1}^{t} X_s|t = 1, 2, 3\ldots\}$. Plot Y_t over time. Do you know what name of the new time series is ? What is its difference from X_t?

Problem 2: scaling behavior and power law

For this problem you need to find a company by yourself which must have been publicly traded in NYSE for at least 30 years. Its P/E ratio by year in 2016 must be smaller than the current S&P 500 index P/E ratio (≈ 25.54 by year).

1. Find such a company. You may consider an investment!

2. Collect daily closing prices of that company in the last 30 years. We know that the log price return is defined as the difference between two consecutive log prices with time lag Δt. In Problem 1 we have implicitly chosen $\Delta t = 1$ (day). In this question, you need to choose $\Delta t = 1, 2, 3, \ldots, 10$ (days) in order to generate different datasets of price returns and calculate their expectation and standard deviation correspondingly. Plot expectation $E[\ln(P_{t+\Delta t}/P_t)]$ and standard deviation $\sigma[\ln(P_{t+\Delta t}/P_t)]$ of log returns with respect to Δt. They should both increase if increasing the time lag Δt. What are the scaling exponents for the expectation and standard deviation? i.e., find α and β so that $E[\ln(P_{t+\Delta t}/P_t)] \sim (\Delta t)^{\alpha}$ and $\sigma[\ln(P_{t+\Delta t}/P_t)] \sim (\Delta t)^{\beta}$. You may need log plots to do linear fits so as to find the power-law relations.

3. If choosing a different Δt, not only the expectation and standard deviation but the shape of PDF itself will change as well. Draw price return PDFs (histograms) with respect to $\Delta t = 1, 2, 3, \ldots, 10$ (Be aware that a PDF must be normalized). Find the maximum of each PDF, $\max\{\mathcal{P}_{\Delta_t}\}$. It should decrease if you increase the time lag. Find γ, the negative scaling exponent which tells you the power law: $\max\{\mathcal{P}_{\Delta_t}\} \sim (\Delta t)^{\gamma}$.