1. Download daily returns data from Yahoo Finance for the S&P500 index from 1990 to present. Report the threshold GARCH(1,1) model.
   a. Is the asymmetric term significant?
   b. Plot the histogram of the standardized residuals \( z_t \). Does it look Normal? What are the skewness and kurtosis?
   c. There is no closed form solution for the distribution of a multiple-step ahead forecast obtained from a GARCH model. We must figure the distribution out by simulating. We will simulate multiple steps ahead for a GARCH model with Normal error and with the bootstrapped errors. First, with normal errors. From the end of the sample, build a forecast of the two week return (i.e. 30 day return) using a GARCH of problem 1. You will condition on the last value of \( h \) and the last value of \( r \). Then simulate the sequence of returns 30 days ahead. Add them up to get the total cumulative return over the 30 day period. Let the last observation in your sample be at time \( T \). The first draw \( (r_{T+1}) \) will be obtained by randomly drawing a value from a Normal(0,1) distribution and multiplying it by \( \sqrt{h_{T+1}} \). You now have a simulated draw of \( r_{T+1} \) that can be used to update and get \( h_{T+2} \). Again, take another random draw from a Normal(0,1) and multiply it by \( \sqrt{h_{T+2}} \) to get a simulated value for \( r_{T+2} \). Continue this out through \( r_{T+30} \) and sum them up to get the cumulative return over the 30-day period. This is one realization (possible outcome) of the 30-day return obtained by simulating the model. Repeat this 2000 times and plot the histogram.
   d. Now we will repeat part a, but instead of drawing from a Normal, we will draw from the empirical distribution of the \( z \)'s obtained in question 1. The first draw \( (r_{T+1}) \) will be obtained by randomly selecting a single value of the \( z \)'s obtained from your fitted model in question 1 and multiplying it by \( h_{T+1} \). You now have a simulated draw of \( r_{T+1} \) that can be used to update and get \( h_{T+2} \). Again, take another random draw from the \( z \)'s and multiply it by \( h_{T+2} \) to get a simulated value for \( r_{T+2} \). Continue this out through \( r_{T+30} \) and sum them up to get the cumulative
return over the 30 day period. This is one realization (possible outcome) of the 30 day return obtained by simulating the model. Repeat this 2000 times and plot the histogram.

e. Find the kurtosis and skewness of the 30-day ahead returns for both sets of simulations in parts a. and b.

f. Find the 1% VaR for the 30 day returns for both sets of simulations in parts a. and b.

g. Find the probability that an option that gives you the right to sell at 15% under the current price will be in the money (i.e. the probability that the index falls by more than 15%) at the end of the 30 day period? Do this again for both sets of simulations.

h. Find the probability that an option that gives you the right to buy at 15% above the current price will be in the money (i.e. the probability that the index rises by more than 15%) at the end of the 30 day period. Do this again for both sets of simulations.

i. Given your results in parts e. and f., which option should be more valuable (i.e. which one has the higher expected payoff)? Discuss.

2. Use the dataset NASDAQ.xls attached with this email for this problem. The dataset contains 5 minute continuously compounded returns for NASDAQ index from 9:30 in the morning to 2:45 in the afternoon. The returns are constructed using the difference of the logarithmic midpoint of the bid and ask prices. It also contains a variable indicating the day (1 through 32). The timestamp is in milliseconds past midnight. If you divide the timestamp by 1000 you recover seconds past midnight. Intraday returns have a deterministic component in the volatility. Volatility is high near the open and the close and relatively low in the middle of the day, a U shape or hockey stick shape over the day. This pattern repeats itself every day. You will estimate a component GARCH model for this data given by: \( r_t = \sqrt{h_t} \sqrt{\phi_t z_t} \). \( \phi(t) = E(r^2 | t) \) is simply the deterministic component that is purely a function of time of day. Notice that the return divided by the deterministic component results in a process that is free of any deterministic elements: \( \tilde{r}_t = \frac{r_t}{\sqrt{\phi(t)}} = \sqrt{h_t} z_t \). Also the model says that \( \tilde{r}_t \) follows a GARCH process.

a. Since the deterministic component is defined in terms of time of day, create a time of day variable.

b. Choose a functional form for this time of day variable that can account for the U shape. Suggestions include a quadratic spline, a polynomial, or perhaps a piecewise linear function (I find the quadratic or cubic splines work well). You could even use a step function where volatility is assumed constant over regions of the day like every half hour. Let \( \phi(t) \) denote your chosen functional form.
Since $\phi(t) = E(r_t^2 \mid t)$ you can estimate parameters in $\phi(t)$ by running a regression $r_t^2 = \phi(t) + \varepsilon_t$.

c. For a single day, plot the estimated function $\phi(t)$

d. Next, create the series $\tilde{r}_t = \frac{r_t}{\sqrt{\phi(t)}} = \sqrt{h_t} z_t$. Fit a GARCH model for $\tilde{r}_t$. Report your model. This GARCH model should have a mean of about 1 and represents the fraction above or below the normal variance for a given time of day.

e. $E(r_t^2 \mid F_{t-1}) = h_t \phi_t$. Pick one day and plot $\sqrt{\phi(t)}$ and $\sqrt{h_t} \sqrt{\phi_t}$ on the same picture. Put the time of day on the horizontal axis.