Midterm Exam

Instructions

The purpose of this exam is to test your knowledge of the contents discussed in class and in the projects, and to test your ability to work with new concepts.

Exam rules:

- The exam is open book: You can use all texts, lecture notes and previous Matlab codes you developed. Note that if your previous Matlab codes were not completely correct and you use them when solving this exam, then you may find incorrect solutions, thus resulting in not getting full credit for your solution.
- Students must uphold the Duke Community Standard: Exams with excessive overlap with other student's answers will receive a zero grade.
- Exam is strictly individual: You must not talk or communicate with each other or anyone else about the exam or course material. This restriction includes e-mail, voice, phone, text, or any other means of communication. The answers you turn in must be your own. Failure to do so will lead to a zero grade.
- All results must be interpreted: Half of the work in the exam is doing the computations. The other half of the work is interpreting the results. You must interpret results regardless of whether the exercise explicitly asked for it or not.
- All plots in the report must be self-contained. Self-contained means that a reader who only sees your figure (image and caption, but not the surrounding text) can understand what you are plotting. This translates to all plots having axis titles, correct units on the axis, and a caption that summarizes what is plotted.
- This exam makes use of stock data. Refer to the Data page for instructions on how to download the data and which files to download (requires Duke login). You must complete all exercises for **both** of your stocks using the data at the 5-minutes sampling frequency, unless stated otherwise.
- If you need to separate diffusive from jump returns, use the value $\alpha = 5$.
- Whenever asked to fill a "Summary Statistics" table, there is no need to convert the units (use the raw value of the statistic).
- Question 0 must be completed for the exam to be graded.

The exam is due on October 3rd by **10:00 am**. You must push your local repository back to GitHub before the deadline. Your repository must contain:

- The Matlab code you used to complete the exam;
- A script file named `main.m` that generates all required plots;
- A `report.pdf` file with your answers to the exam questions. The report must also contain an Appendix with the code used to solve the exam;

The repository for this exam can be obtained by clicking **on this link**.

Question 0

During the exam you will use data on your two stocks and also on the market index (SPY). If one of your stocks already is SPY, pick a different one from the available stocks and clearly state what stock you are using. Fill the table below with the stocks you will use for solving this exam:

Stock	Ticker
1	
2	
3	SPY

Question 1 - Value at Risk [40 points]

Suppose it is 9:00 the morning of September 16, 2008, in the middle of the financial crisis. You work for a mutual fund manager who has a \$200 million position in a particular stock. The fund manager is worried about a very large loss over the upcoming trading day starting in the morning and ending at 16:00. Specifically, the manager asks you for the probability that the value of the position (ignoring rare jumps) will decline by 2 percent or worse over the day and for the probability of decline by 4 percent or worse over the day. The position size of \$200 million is very realistic and the potential losses of 2 or 4 percent translate to \$4 and \$8 million, respectively. The manager is asking you for:

$$\mathbb{P}(r_t^c \leq -q) \text{ for } q = 2\% \text{ or } q = 4\%$$

where $r_t^c = \sum_{i=1}^{n} r_{t,i}^c$ is the return over the day and $r_t^c \overset{d}{\sim} \mathcal{N}\left(0, \mathrm{IV}_t\right)$.

A. [5 points]

Identify the correct value of t in your data set for the date: September 16, 2008. If it is the morning of September 16, 2008, is it possible to estimate IV for today? Why or why not?

B. [5 points]

We cannot estimate IV_t, but the best we can do is forecast its value based on the historical data. A simple forecast for IV_t (that works reasonably well in practice) is to use yesterday's estimate for the integrated variance. Report the estimate for yesterday's integrated variance. Interpret its value, and justify its use as a forecast for the variance over the next day.

C. [12 points]

For both of your stocks, estimate the probabilities described in the exercise. Compute 95% confidence intervals for yesterday's integrated variance, and use those to construct 95% confidence intervals for the probabilities. Fill in the table below: Comment on the

		Estimate of $\mathbb{P}(r_t^c \leq -q)$		
Stock	q	Estimate	Lower bound	Upper bound
1	2			
1	4			
2	2			
2	4			

estimation accuracy. If the width of the confidence interval is in the range 0.01–0.06, then that level of accuracy would likely suffice for most applications. However, if the width is something like 0.20–0.30, then the estimates would be too imprecise for any practical application.

D. [18 points]

Above, we fixed the loss (2% or 4%) and determined the probability of the loss. In risk management, it is common to turn the problem around by fixing the probability at some $p \in [0, 1]$ and determine the associated loss in dollars that could occur with probability fixed at p. This value is known as the Value at Risk (VaR).

If r_t^c is expressed as a percent, that is, $r_t^c \pm 1$ percent or $r_t^c \pm 2$ percent, the actual dollar gain or loss on an investment of $V = \$200$ million is $\pm\frac{r_t^c}{100}V$. Therefore, the VaR is the number Q such that:

$$\mathbb{P}\left(\frac{r_t^c}{100} \times V \leq Q\right) = p$$

That is, the probability of losing more than Q is given by p.

Complete the table below for each of your stocks: Comment briefly on the accuracy

VaR Estimate and 95% Confidence Interval				
Stock	p	VaR	Lower Bound	Upper Bound
1	1%			
1	5%			
2	1%			
2	5%			

of the VaR estimates. Would your boss find them helpful or uselessly imprecise?

Hint: To compute the Value at Risk, write the probability in terms of a standard normal (what is the distribution of r_t^c?), then use the fact that the cumulative density function of a normal can be inverted and solve for Q.

Question 2 - Market Neutral Strategy [30 points]

A long-short hedge fund follows a trading strategy defined by a long (buy) position in one stock and a short (sell) position in the other. The idea is that the fund has identified positive future prospects for the stock held long and negative future prospects for the stock that is shorted.

The rules of short trading are set up so that the return on the short stock is just the negative of the return if it were held long. Let $r_{t,i}^{A}$ denote the return on stock A (long) at day t over the i-th interval, and $r_{t,i}^{B}$ denote the return on stock B (short). Then, the return on the long-short portfolio is just:

$$r_{t,i}^{\text{Long-Short}} = r_{t,i}^{A} - r_{t,i}^{B}$$

Usually, hedge funds charge investors rather large fees, and investors thereby demand that the fund demonstrate that its strategy is market neutral. A market neutral strategy is a strategy that provides returns that have zero beta with the market. The reason investors demand a market neutral strategy is that investors can easily (and cheaply) earn the average market return by using a mutual fund or ETF that tracks the market index. Therefore, there is no point in paying the hedge fund the fees for the market return.

For this exercise you will also use data on the market index (SPY). If one of your stocks already is SPY, pick a different one from the available stocks and clearly state what stock you are using.

A. [2 points]

Compute the long-short 5-min returns (long on stock 1, short on stock 2) for your two stocks across all days. Fill out the summary statistics table below for the diffusive returns of your long-short portfolio:

Summary Statistics of Long-Short Continuous Returns				
Average	Min	5% Percentile	95% Percentile	Max

B. [4 points]

Estimate the realized beta between your long-short portfolio and the market index (SPY). Fill the table below:

Summary Statistics of the Realized Beta				
Average	Min	5% Percentile	95% Percentile	Max

C. [8 points]

Compute 95% confidence intervals for the realized beta. Use $k_n = 11$ and 1000 repetitions to keep the computations doable. Plot the entire trajectory of the realized beta and intervals in a suitable way to show an investor. Comment on whether at least visually the long-short position appears market neutral.

D. [8 points]

Plot the realized beta and 95% confidence intervals for October, 2008 (be careful with the x-axis, only plot the business days in October). Interpret the results.

E. [8 points]

Most investors demand more than just a plot and insist on formal statistical tests. We can test the null hypothesis of zero beta at the 5 percent level by checking whether a 95 percent confidence includes zero or not. That is, we reject the null hypothesis if zero does not lie in the 95 percent confidence interval. Run the test day-by-day for your data set and compute the average number of rejections over all the days in your data set. Report the total number of rejections. A common rule is to declare a hedge fund **not** market neutral if there are 10 percent or more rejections. (Five percent would be expected by simple statistical fluctuations so the 10 percent rule adds in a small additional allowance.) Would the hedge fund holding long stock 1 and short stock 2 be considered market neutral?

Question 3 - Common Variation [30 points]

For this exercise you will also use data on the market index (SPY). If one of your stocks already is SPY, pick a different one from the available stocks and clearly state what stock you are using.

A. [2 points]

Consider:

$$\begin{aligned} r^c_{1,t,i} &= \text{continuous returns of stock 1} \\ r^c_{2,t,i} &= \text{continuous returns of stock 2} \\ r^c_{m,t,i} &= \text{continuous returns of SPY} \end{aligned}$$

Obtain the continuous returns above. Complete the table below:

Summary Statistics for Continuous Returns					
Stock	Average	Min	5% Percentile	95% Percentile	Max
1					
2					
SPY					

B. [2 points]

How many jumps did you detect in the market?

C. [2 points]

Estimate the realized betas for both stocks, and fill the table below:

Summary Statistics of the Realized Beta					
Stock	Average	Min	5% Percentile	95% Percentile	Max
1					
2					

D. [2 points]

Plot the realized betas of both stocks. From viewing the plot, would you say that your stocks generally more risky than the market ($\beta > 1$), less risky than the market ($\beta < 1$), or as risky as the market ($\beta = 1$)?

E. [4 points]

The residual (non-market) part of the returns on the stocks are:

$$\begin{aligned} e_{1,t,i} &= r^c_{1,t,i} - R\beta_{1,t} r^c_{m,t,i} \\ e_{2,t,i} &= r^c_{2,t,i} - R\beta_{2,t} r^c_{m,t,i} \end{aligned}$$

The residual stock returns can be correlated because of common factors other than the market. The realized correlation between the residuals is denoted by:

$$\rho_{e,t} = \text{Correlation between } e_1 \text{ and } e_2 \text{ on day } t$$

The correlation $\rho_{e,t}$ between the residuals is the simple correlation between the residuals over the day (Matlab function `corr`). Compute the correlation day by day and fill the below:

Summary Statistics of the Residual Correlation				
Average	Min	5% Percentile	95% Percentile	Max

F. [8 points]

Compute 95% confidence intervals for the realized correlation between the residuals using the bootstrap method (with $k_n = 11$ and 1000 repetitions). Plot the correlations and confidence intervals. Interpret the results.

G. [5 points]

Plot the correlation and 95% confidence intervals for October, 2008 (be careful with the x-axis, only plot the business days in October). Interpret the results.

H. [5 points]

Consider the null hypothesis $H_0 : \rho_{e,t} = 0$ versus the alternative $H_1 : \rho_{e,t} \neq 0$. Use your confidence intervals to compute the following proportions:

$$p_0 = \frac{\text{number of days do not reject H}_0\text{: } \rho_{et} = 0}{252}$$
$$p_+ = \frac{\text{number of days reject H}_0\text{: } \rho_{et} = 0 \text{ in favor of } \rho_{et} > 0}{252}$$
$$p_- = \frac{\text{number of days reject H}_0\text{: } \rho_{et} = 0 \text{ in favor of } \rho_{et} < 0}{252}$$

Report the numbers and interpret the results.