Project 3: Jump Returns, Truncated Variance and Inference for the Integrated Variance

1 Instructions

Project 3 is due on September 20th by 10:00 pm. This is a hard deadline, so no exceptions. You must push your local repository back to GitHub before the deadline. Your repository must contain:

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

You can obtain the repository for this project by clicking **on this link**.

Questions

The purpose of this project is to implement the truncated variance estimator and conduct inference for the integrated variance. You will learn about the intraday pattern of returns, how it is used in the separation of diffusive returns from jump returns. You will also learn how to use the delta-method to construct valid confidence intervals for the annualized integrated variance. Lastly, you will learn how to run the bootstrap method to obtain confidence intervals without necessarily deriving the asymptotic results for an estimator.

Exercise 1

The objective of this exercise is to learn how to separate diffusive returns from jump returns, and to understand the distribution of the jump returns.

A.

Compute the time-of-day factor (τ) and plot it. Interpret the results.

B. (Optional, PhD Required)

Why does the time-of-day factor change over the day? Why is it not a constant? Cite the literature to support your arguments. Hint: find motives for why the volatility is U-shaped.

C.

Let $\alpha = 4$ and compute the jump separation threshold, $cutoff_{t,i}$, for your entire sample. Then compute the diffusive and jump returns:

$$r^c_{t,i} \equiv r_{t,i} 1_{\{|r_{t,i}| \leq \text{cutoff}_{t,i}\}}$$
$$r^d_{t,i} \equiv r_{t,i} 1_{\{|r_{t,i}| > \text{cutoff}_{t,i}\}}$$

Verify that $r_{t,i} = r^c_{t,i} + r^d_{t,i}$. Plot the diffusive and jump returns. Interpret the results.

D.

How many jumps occur in each year? Are there more jumps during periods of crisis? Are jumps approximately evenly distributed throughout the years?

E.

What is a density function? If you could estimate the density of the diffusive returns what would you expect it to look like? Explain. If you could estimate the density of the jump returns what would you expect it to look like? Explain. If you ignored the zeros in the jump returns and estimated the density for the non-zero jump returns, then what do you expect it to look like? Explain.

G.

Use the ksdensity function from Matlab to estimate the density of the diffusive and jump returns. The function takes a few optional arguments which we will make use of. Specify the 'Kernel' to be 'epanechnikov', and adjust the 'Bandwidth' argument so that the density is reasonably smooth, but not over-smoothed. Plot the densities and interpret.

H.

Many people believe that stock crashes (negative jump returns) are larger and occur much more often than rallies (positive jump returns). What do your density plots suggest about this belief?

Exercise 2

The objective of this exercise is to implement the truncated variance estimator and conduct valid inference on the integrated variance.

A.

Compute the truncated variance for all days in the sample and make plots. Remember to annualize the values for plotting. Interpret the results.

B.

Plot the truncated variance and the realized variance and compare. Remember to annualized the values for plotting.

C.

Estimate the quartic integrated variance and plot it. Interpret the results.

D.

Compute the 95% confidence intervals for the integrated variance based on the asymptotic distribution of the truncated variance. There is no need to annualize the estimate in this exercise. Plot the truncated variance alongside the confidence interval.

E.

The plot in the previous exercise is quite condensed. Zoom in on a month of interest (say October 2008) and plot TV and the confidence intervals. Discuss.

F.

It turns out that annualizing the confidence interval does not lead to the correct confidence interval for the annualized TV. Remember, the asymptotic distribution for the TV is:

$$\Delta_n^{-\frac{1}{2}}(\mathrm{TV}_t - \mathrm{IV}_t) \xrightarrow{d} \mathcal{N}(0, 2\mathrm{QIV}_t)$$

But we want to find the asymptotic distribution for the annualized TV:

$$\Delta_n^{-\frac{1}{2}}\left(100\sqrt{252\mathrm{TV}_t} - 100\sqrt{252\mathrm{IV}_t}\right) \xrightarrow{d} \mathcal{N}(0, 1)$$

To do so, we will use the Delta Method theorem.

- Read the univariate case of the delta method theorem.
- What does the Delta Method theorem allow us to do?
- Why do we need it?
- What should be g defined as in our case?
- What is g'?
- What is the asymptotic distribution of $\Delta_n^{-\frac{1}{2}}\left(100\sqrt{252TV_t} - 100\sqrt{252IV_t}\right)$?

G.

Given the asymptotic distribution of: $\Delta_n^{-\frac{1}{2}}\left(100\sqrt{252\text{TV}_t} - 100\sqrt{252\text{IV}_t}\right)$. What is the confidence intervals for the annualized integrated variance:

$$CI(100\sqrt{252 \times IV_t}, 1-\alpha) = \ ?$$

H.

Plot the annualized TV (zoom-in on some month) and the correct 95% confidence intervals. Interpret the results.

Exercise 3

In this exercise you will do a Monte Carlo analysis to evaluate the accuracy of the IV estimators (and the asymptotic theory) we have developed so far. The idea is to simulate a jump-diffusion model, and use the simulated values of the c_t process to compute the actual integrated variance. Then, we will put ourselves in the place of the practitioner, who only has access to the stock returns, and uses these returns to estimate IV via the estimators we have studied.

Let X denote the process for the log-price of an asset. We want to simulate:

$$\begin{aligned} dX_t &= \sqrt{c_t}dW_t \\ dc_t &= \rho(\mu_c - c_t)dt + \sigma_c\sqrt{c_t}dW_t^c \end{aligned}$$

where W and W^c are independent.

Assume that:

$$\begin{aligned} n &= 80 \\ T &= 1.25 \times 252 \\ n_E &= 20 \times n \\ \rho &= 0.03 \\ \mu_c &= 0.011^2 \\ \sigma_c &= 0.001 \\ c_0 &= \mu_c \\ X_0 &= \log 292.58 \end{aligned}$$

Use the following additional rule. If the simulated c becomes smaller than $\mu_c/2$ change the value to $\mu_c/2$ instead.

A.

Simulate the stochastic variance process c_t. An Euler discretization scheme is now necessary to reduce the discretization error. Plot the simulated variance process.

B.

Because we "observe" the values of the variance process c we can compute the actual IV:

$$IV_t \equiv \int_{t-1}^{t} c_s ds$$

- What does the numerical value of IV_t represent?
- We do not have a continuum of values for c_t, but only as many as we simulated. How can we use those to compute the actual IV_t?

C.

Let $c_{t,i}$ denote the simulated value of the variance on day t at interval i. Compute the actual IV for each of the simulated days:

$$IV_t = \frac{1}{N_{euler}} \sum_{i=1}^{N_{euler}} c_{t,i}$$

Plot the integrated variance. Interpret the results.

D.

We now take the position of the researcher that does not observe the underlying process, but has access only to samples of the price at discrete times.

Use the prices from the simulation to compute the intraday returns. Remember, we use the Euler scheme to reduce the discretization error, but we are simulating 5-minute prices, not 15-second prices. Use the intraday returns to compute RV_t for all simulated days.

Plot the realized variance alongside the actual integrated variance. Discuss.

E.

Remember that for RV the asymptotic distribution is (we are ignoring jumps for now):

$$\Delta_n^{-\frac{1}{2}}(\mathrm{RV}_t - \mathrm{IV}_t) \xrightarrow{d} \mathcal{N}(0, 2\mathrm{QIV}_t)$$

Compute the 95% confidence interval for IV using the simulated returns. Plot the IV and the confidence intervals.

F.

The theory says that, for n big enough, the probability of the actual integrated variance be within the confidence interval should be 95%.

- For how many days is IV_t within the confidence interval?
- What is the average coverage rate?
- Is it close to the expected 95%?

G.

We know that when the stock price can jump the RV_t estimator becomes biased. Let's see this effect in practice.

First, let's add jumps (compound Poisson) to the model

$$\begin{aligned} dX_t &= \sqrt{c_t}dW_t + dJ_t \\ dc_t &= \rho(\mu_c - c_t)dt + \sigma_c\sqrt{c_t}dW_t^c \end{aligned}$$

Use the following parameters to simulate the jumps:

$$\lambda = 20/252$$
$$\sigma_{jump} = 30 * \sqrt{\mu_c/N}$$

Add the jumps to the simulated prices. Plot the simulated prices with jumps.

H.

Use the simulated prices from the model with jumps to compute the intraday returns. Then compute RV for each day of the sample, and compute the confidence intervals for the annualized IV based on the asymptotic theory.

- What is the coverage rate of the confidence intervals?
- How does it compare to when there are no jumps?
- Is the value of the coverage rate what you expected? Explain.
- How did the failure to account for jump returns affect the inference?

I.

Change the jump intensity to a higher value, say $\lambda = 1$, and redo the previous exercise.