Midterm Exam - Fall 2020

FE-570

October 17, 2020

Problem 1 (5pt)
Which of the following are true about a limit order? (multiple answers are allowed)
i) It is executed on arrival
ii) It is typically placed by patient traders
iii) It creates liquidity.

Problem 2. (5pt)
Write one short comment (1-2 sentences) for each of these statements about the bid-ask spread, explaining why it is (or not) true:
i) It is related to trading costs
ii) It includes dealer's inventory costs
iii) It can be used as a liquidity measure.

Problem 3. (5pt)
Explain briefly the liquidity measures *effective spread* and *realized spread.*

Problem 4. (5pt)
One of the stylized facts of the financial markets states that the log-returns of the stock price on consecutive days $r_i = \log(S_i/S_{i-1})$ are (pick one):
i) Positively correlated
ii) Uncorrelated
iii) Negatively correlated

Does this stylized fact remain valid for microstructure data? If not, what is different at the microscale?

Problem 5. (5pt)
Rank the following distributions in increasing order of heavy tails. This does not require any calculations, just examine the analytical formula for the distribution in the tails.

i) Log-normal distributions
ii) Gaussian distribution
iii) Student-t distribution with 3 degrees of freedom $f_3(t) = \frac{1}{(1+\frac{1}{3}t^2)^2}$.

Problem 6. (5pt)
State the stylized property of the financial markets known as *Aggregational Gaussianity*, and explain briefly why do we expect it to hold.

Problem 7. (20pt)
Roll model with stale prices. (Exercise 8.1 in Hausbruck.) The beliefs of market participants at time t are summarized in m_t (the efficient price), which follows a random walk $m_t = m_{t-1} + w_t$. But due to operational delays, trades actually occur relative to a lagged value: $p_t = m_{t-1} + cd_t$, where $d_t = \{\pm 1\}$ is the trade indicator ($d_t = +1$ is a buy, $d_t = -1$ is a sell).

Compute the autocovariances of the trade price changes $\Delta p_t = p_t - p_{t-1}$.

Problem 8. (50pt) [Volatility estimation at the microscale]
For this problem we will use the R code provided to determine the volatility of the stock price from microstructure data. We will use the *sample_tdata* and *sample_qdata* files in the *highfrequency* package containing trade price and quotes data of a fictitious stock covering an entire trading day.

Recall that the realized variance $RV(\tau)$ of the price process p_t over the time period $[0, T]$ sampled with time step τ is defined as

$$RV(\tau) := \sum_{i=1}^{N} |p_{\tau(i+1)} - p_{\tau i}|^2 \tag{1}$$

with $N = T/\tau$ the number of time sampling steps.

We work in trading time (as opposed to calendar time), measured in time ticks 1,2,3, $\cdots, n_{\text{trades}}$. Trading time i increments by one whenever a trade occurs.

The code uses two functions: i) *realizedVar(q)* computes the realized variance $RV(q)$ of the trade price p_t over a lag q in trading time (this means that it uses only every q-th trade price)

$$\text{realizedVar(q)} = RV(q) = \frac{1}{q} \sum_{t=1}^{[n_{trades}/q]} (p_{q(t+1)} - p_{qt})^2 \tag{2}$$

ii) $realizedVarLog(q)$ is defined in a similar way but computes the realized variance of the log-price $\log p_t$.

i) Using the functions $realizedVar(q)$ and $realizedVarLog(q)$ plot the signature plot for the realized variance $RV(q)$ at lag $q = 1 : 200$ for both arithmetic price changes $\Delta p_t = p_t - p_{t-1}$ and log-normal price changes $\Delta \log p_t = \log p_t - \log p_{t-1}$. Discuss the shape of the plot.

ii) Assuming that trading volume is constant throughout the trading day, compute the realized variance $rv5 = RV(q_{5\text{min}})$ corresponding to sampling every 5 minutes. This assumption allows us to convert results to calendar time.

How does $rv5$ compare with the realized variance $RV(1)$ with lag 1? Which one do you think is a more precise estimate of the realized variance of the underlying stock price and why?

iii) Fill out the numerical values in the table below, showing the daily estimated volatility σ_{day} from arithmetic trade price changes (first three columns) and log-normal price changes (last two columns), for three lag values $q = 1, 2, q_{5\text{mins}}$.

iv) Estimate the daily volatility from the Roll model. Recall that in the Roll model the trade price $p_t = m_t + cd_t$ contains the efficient price m_t and the bid-ask bounce term cd_t proportional to the trade indicator $d_t = \{\pm 1\}$.

The code computes the variance of the price changes $\gamma_0 = var(\Delta p_t)$ and their covariance at lag-1 $\gamma_1 = \text{cov}(\Delta p_t \Delta p_{t-1})$.

The efficient price m_t follows a random walk with increments of variance $\text{var}(m_t - m_{t-1}) = \sigma_u^2$. Estimate σ_u using $\sigma_u^2 = \gamma_0 + 2\gamma_1$. The daily realized variance is estimated as $rv_{\text{Roll}} = \sigma_u^2 n_{\text{trades}}$ where n_{trades} is the total number of trades during the day.

Table 1: Daily realized variance $RV(q)$ from the functions $realizedVar(q)$ and $realizedVarLog(q)$ at different sampling frequencies q. The lag $q_{5\text{min}}$ corresponds to sampling every 5 minutes. For each lag q we obtain an estimate of the daily volatility $\sigma_{\text{day,q}} = \sqrt{rv(q)}$. $\sigma^{LN}_{\text{day,q}}$ is defined in the same way using $\log p$. The Roll model estimates the daily realized variance as $rv_{\text{Roll}} = (\sigma^{\text{Roll}}_{\text{day}})^2 = \sigma_u^2 n_{\text{trades}}$. p_1 is the price of the first trade.

q	$rv(q)$	$\sigma_{\text{day,q}}$	$\frac{1}{p_1}\sigma_{\text{day,q}}$	$rv_{\log}(q)$	$\sigma^{LN}_{\text{day,q}}$
1					
2					
q_{5min}					
	σ_u	$\sigma^{Roll}_{\text{day}}$			$\frac{1}{p_1}\sigma^{Roll}_{\text{day}}$
Roll model					

How does the Roll model estimate of the daily volatility compare with the direct evaluation from realized variance? Which one do you think is more precise?

Problem 9. (50pt)
Consider the GARCH(1,1) model for conditional variance for one time period σ_i^2

$$\sigma_{i+1}^2 = \omega + \alpha\varepsilon_i^2 + \beta\sigma_i^2 \tag{3}$$

where the log-return $\varepsilon_i = \sigma_i Z_i$ with $Z_i = N(0,1)$ is a normally distributed random variable with mean zero and variance 1.

i) Assuming that the model is weakly-stationary, compute the average $v_\infty = \mathbb{E}[\sigma_i^2]$ and the kurtosis of the conditional variance $\mathbb{E}[\sigma_i^4]$ in this model.

ii) Simulate one year of data (252 business days), starting with $\sigma_i^2 = v_\infty$ equal to the stationary value. Estimate the mean and variance of the resulting sample, and compare with the theoretical results.

For the numerical simulation assume the model parameters determined by V-Lab for SP500

$$\alpha = 0.0923\,, \quad \beta = 0.8952\,, \quad \omega = 0.0149\,. \tag{4}$$