

FINAL EXAM

NAME

SIGNATURE

1. There are total 100 points.
2. You have till 6pm, May 7th to submit the exam.
3. **You are not allowed to collaborate with other students.** I will be very strict on this and check for similarity. Your answers must be concise and well organized. For problem A and B, you can scan or take clear pictures and upload them. For problem C you should summarize the results into a report using outputs, figures, tables and detailed explanations. The report should be in either **word document or pdf format**. Please upload both the report and your R code (the R code should reproduce the results in your report, and should have appropriate comments). It's your responsibility to make sure the code run smoothly and can produce the results. Late exam will not be accepted.
4. "GOOD LUCK!"

• **Problem A** (22 points total)

Consider the nonsynchronous trading model discussed. Let r_t be the true return of a stock at time t . It is assumed to be independent and identically distributed with $E(r_t) = \mu$ and $Var(r_t) = \sigma^2$. Let r_t^o be the observed return. When there is no trade at time t , we have $r_t^o = 0$. When there is a trade, then $r_t^o = r_t + r_{t-1} + \cdots + r_{t-k_t}$ where k_t is the largest nonnegative integer such that no trade occurred in the periods $t - k_t, t - k_t + 1, \dots, t - 1$. For each time period, the probability that the security is not traded is $\pi = 0.3$, which is independent from r_t . Furthermore, whether there is a trade or not is independent from period to period.

(a) What is the probability that there is a trade at time t and $t - 4$, but no trade in between? Why? (4 pts)

(b) What is the observed r_{t-3}^o in terms of the true log returns $r_t, r_{t-1}, r_{t-2}, r_{t-3}, \dots$ when there is a trade at time $t - 3$ and $t - 5$, but no trade in between? (4 pts)

(c) What is $P(r_t^o = r_t | r_{t-1}^o = r_{t-1})$? (6 pts)

(d) Show in details that $COV(r_t^o, r_{t-2}^o) = -\pi^2 \mu^2$. (8 pts)

• **Problem B** (16 points total)

Let $\{X_t\}$ denote the time series of interest. However, because measurement process is not perfect, we actually observe $Y_t = X_t + W_t$. We assume that $\{X_t\}$ and $\{W_t\}$ are independent processes, and that $\{W_t\}$ is a mean 0 white noise process with variance 1, and that $\{X_t\}$ is weakly stationary. We call $\{X_t\}$ the **signal** and $\{W_t\}$ the **measurement noise**. We call σ_X^2/σ_W^2 the signal to noise ratio (SNR).

(1) Prove that the process Y_t is weakly stationary. (7 pts)

(2) Prove that the larger the SNR, the closer the autocorrelation function of the observed process $\{Y_t\}$ is to the autocorrelation function of the signal $\{X_t\}$. (9 pts)

(hint: express the autocorrelation function of the process $\{Y_t\}$ as a function of σ_X^2 , σ_W^2 and the autocorrelation function of the process $\{X_t\}$).

In problem C, you should carry out the analysis in R, and summarize the results using figures, tables and detailed explanations. Do not simply turn in your R codes and R outputs as your report.

- **Problem C** (62 points total)

Pick a stock whose FIRST letter is the same as the FIRST letter of your FIRST name. The stock should have market capital greater than 1 billion. Obtain adjusted daily closing pricing of the stock. Use dates from='2015-01-01',to='2019-12-31'. Compute daily log returns of the stock. Note that there should be 1341 returns (the first return is on 2015-01-05). Use the first 1200 as training(the return up to Oct 9th, 2019) and the last 141 as the holdout set. If there doesn't exist any autocorrelation in the data, choose another stock that satisfies the criteria.

1. Write a procedure using for loops to identify the best arima orders p and q of the training data(returns from 2015-01-05 to 2019-10-09) based on the AIC value. When running the for loops,set the maximum number of p and q to be both 5. Perform model fitting and model diagnosis on the model chosen. (10 pts)
2. Now fix the order of the arima model, and at each step, only parameter estimation changes. Use this model to predict one day ahead. Compare the predicted direction (namely if the prediction return is positive or negative) with the direction of the actual return. Were you right? For the next day re-estimate the parameters using returns from 2015-01-06 to 2019-10-10 and predict the next day's direction. Repeat until the test data is exhausted. Report the probability of being correct (calculated as the percentage of times the model predicts the directions correctly). (10 pts)
3. Check if there exist significant ARCH effect in the time series. If so, identity the best ARMA+GARCH model. Pay attention to the error distribution, choose between normal, t , and skewed t distributions. Perform model fitting and model diagnosis on the fitted model. (8 pts)
4. Suppose you are given a \$100, 000 account. Starting from Oct. 9th, at the end of the trading day, you decide whether to long the stock based on the arima model prediction for Oct. 10th. If the prediction is positive, then you long(buy) using all your money at the adjusted closing price of Oct. 9th and close(meaning sell) at the end of Oct 10th using the corresponding adjusted closing price. Otherwise you do nothing. We neglect transaction costs here. For the next day, re-estimate the model and predict, and carry out the trade. What is the value of your portfolio at the end of the process? (8 pts)
5. Now, do the same thing as in the previous step, except that you long only portions of account according to your prediction. This time we use the ARMA+GARCH model identified in (3). When the predicted return is in between 0 and $0.1 \times \sigma_t$, you long using 50% of your money, when the predicted return is in between $0.1 \times \sigma_t$ and $0.2 \times \sigma_t$, you long using 80% of your money, and when the predicted return is greater than $0.2 \times \sigma_t$, you long using 100% of your money. σ_t is the predicted volatility. You do this in a rolling fashion as previously. What is the value of your portfolio at the end of the process? (8 points)
6. Define and carry out one more trading strategy the best you can. The strategy needs to use the predicted returns and predicted volatilities from the fitted models . You can neglect the transaction cost. This part will be graded on the substantiality and complexity of the strategy. (18 pts)