Stevens Institute of Technology **Homework 3**
QF 202 – Introduction to Financial Time Series **Due Date: March 21, 2022**

Problem 1. (70 points) Please download the attached data (T-notes.csv). This data set contains 10-year treasury notes daily prices from 2021 March to 2022 Feb. Please import the CSV file into R and make sure the time series is in the correct time order.

(a) Calculate the daily simple return of T-notes for the given period. Split the result into two separate subsets. The first one (DATA 1) contains the first 200 daily returns. The second subset (DATA 2) contains the remaining daily returns.

(b) Create an ACF plot usinf the T-notes return in DATA 1. Report the values of the first 20 lags of autocorrelation.

(c) Use the PACF function to plot the partial autocorrelations using DATA 1. Based on the plots what is the recommended order p if we want to fit an AR(p) model on DATA 1.

(d) Use the ARIMA function and the order p you determined in previous problem fit an AR(p) model on DATA 1. Name this fitted model as MODEL 1.

(e) Use the builtin function ar() and your preferred criteria to fit another AR model on DATA 1. Name this fitted model as MODEL 2.

(f) For each model, predict the daily returns for the next 20 days. Plot the predicted values following the original values in DATA 1.

(g) Using DATA 2, compute the mean squared error (MSE) for each AR model you constructed. The equation for the MSE is given below. In that equation $\hat{Y}_i$ is the predicted value and Y_i is the real value at time i. The real return data is calculated using DATA 2. Compare the MSE values for the models. Which model is better?

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2$$

Problem 2. (30 points) Let $w_t \overset{iid}{\sim} N\left(0, \sigma_w^2\right)$. The 1^{st} order moving average model, denoted by MA(1) is:

$$x_t = \mu + w_t + \theta_1 w_{t-1}$$

The 2^{nd} order moving average model, denoted by MA(2) is:

$$x_t = \mu + w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2},$$

where θ_1 and θ_2 are parameters.
For these two models assume the following formulas. These can be easily proven.

- For MA(1):
  - Mean is $E\left(x_t\right) = \mu$

- Variance is $\text{Var}(x_t) = \sigma_w^2 \left(1 + \theta_1^2\right)$
- Autocorrelation function (ACF):

$$\rho_1 = \frac{\theta_1}{1 + \theta_1^2}$$

- For MA(2):
  - Mean is $E(x_t) = \mu$
  - Variance is $\text{Var}(x_t) = \sigma_w^2 \left(1 + \theta_1^2 + \theta_2^2\right)$
  - Autocorrelation function (ACF):

$$\rho_1 = \frac{\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2}, \rho_2 = \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2}$$

Please answer the following questions.

(a) Suppose that an MA(1) model is $x_t = 10 + w_t + .7w_{t-1}$, and the MA(2) model $x_t = 10 + w_t + .5w_{t-1} + .3w_{t-2}$. Calculate the theoretical autocorrelation values for these models.

(b) Using arima.sim() function to generate 5000 data points of corresponding MA(1) model and calculate the ACF value for this data set.

(c) Using arima.sim() function to generate 5000 data points of corresponding MA(2) model and calculate the ACF value for this data set.

(d) Present a table with the theoretical values and estimated values. Comment on what you observe.

BONUS The values above are for $n = 5000$. Repeat the procedure for a series of n values and study the differences between the real and estimated values as a function of n. In your opinion what is the order of this function. That is, how does this difference goes to 0 as n goes to ∞?