# TIØ4317
# Empirical and Quantitative Methods in Finance
# Exercise 3

## Instructions

Solutions to the problems will be posted on BlackBoard after the deadline. You can use either Excel or a high-level programming language, e.g., R or Python, to solve the programming exercises. We suggest that you write your solutions using MS Word or LaTeX. Also, hand in all code and/or Excel files.

**Deadline:** Monday February 17th, 2025, 23:59. **Grading:** Passed/Failed.

## Tasks

1. What are the differences between autoregressive and moving average models?

2. Consider the following three models that a researcher suggests might be a reasonable model for stock market prices:

$$y_t = y_{t-1} + u_t \tag{1}$$
$$y_t = 0.5y_{t-1} + u_t \tag{2}$$
$$y_t = 0.8u_{t-1} + u_t \tag{3}$$

   (a) What classes of models are these examples of?

   (b) What would the autocorrelation function for each of these processes look like? (you do not need to calculate the acf, simply consider what shape it might have given the class of the model from which it is drawn)

   (c) Which model is more likely to represent stock market prices from a theoretical perspective, and why? If any of the three models truly represented the way stock market prices move, which could potentially be used to make money by forecasting future values of the series?

   (d) By making a series of successive substitutions or from your knowledge of the behavior of these types of processes, consider the extent of persistence of shocks in the series in each case.

3. (a) You obtain the following sample autocorrelations and partial autocorrelations for a sample of 100 observations from actual data:

| Lag | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| acf | 0.420 | 0.104 | 0.032 | -0.206 | -0.138 | 0.042 | -0.018 | 0.074 |
| pacf | 0.632 | 0.381 | 0.268 | 0.199 | 0.205 | 0.101 | 0.096 | 0.082 |

Table 1: Sample autocorrelations and partial autocorrelations

Can you identify the most appropriate time series process for this data?

(b) Use the Ljung-Box $Q^*$ test to determine whether the first three autocorrelation coefficients taken together are jointly significantly different from zero.

4. You obtain the following estimates for an AR(2) model for some returns data $y_t = 0.803y_{t-1} + 0.682y_{t-2} + u_t$ where $u_t$ is a white noise error process. By examining the characteristic equation, check the estimated model for stationarity.

5. A researcher is trying to determine the appropriate order of an ARMA model to describe some actual data, with 200 observations available. She has the following figures for the log of the estimated residual variance (i.e. $\log(\hat{\sigma}^2)$) for various candidate models. She has assumed that an order greater than (3,3) should not be necessary to model the dynamics of the data. What is the 'optimal' model order?

| ARMA(p,q) | $\log(\hat{\sigma}^2)$ |
|---|---|
| (0,0) | 0.932 |
| (1,0) | 0.864 |
| (0,1) | 0.902 |
| (1,1) | 0.836 |
| (2,1) | 0.801 |
| (1,2) | 0.821 |
| (2,2) | 0.789 |
| (3,2) | 0.773 |
| (2,3) | 0.782 |
| (3,3) | 0.764 |

Table 2: Log of the estimated residual variance for various ARMA model orders

6. You have estimated the following ARMA(1,1) model for some time series data: $y_t = 0.036 + 0.69y_{t-1} + 0.42u_{t-1} + u_t$ Suppose that you have data for time to $t-1$, i.e. you know that $y_{t-1} = 3.4$, and $u_{t-1} = -1.3$.

(a) Obtain forecasts for the series $y$ for times $t$, $t+1$, and $t+2$ using the estimated ARMA model.

(b) If the actual values for the series turned out to be -0.032, 0.961, 0.203 for $t$, $t+1$, $t+2$, calculate the (out-of-sample) mean squared error.

7. Select two of the stock series from the 'CAPM.XLS' Excel file (given in Exercise 1). Construct a set of continuously compounded returns, and then perform a time series analysis of these returns. The analysis should include:

(a) An examination of the autocorrelation and partial autocorrelation functions.

(b) An estimation of the information criteria for each ARMA model order from (0,0) to (5,5).

(c) An estimation of the model that you feel most appropriate given the results that you found from the previous two parts of the question.

(d) Split your data in two parts, such that you keep the five last periods as a holdout sample. Use metrics such as MSE, MAE, MAPE and 'percentage of correct sign' to compare the forecasting accuracy of:

i. Your chosen ARMA model.

ii. An arbitrary ARMA(1,1).

iii. A random walk with drift in the log price levels.

8. (a) What kinds of variables are likely to be non-stationary? How can such variables be made stationary?

(b) Why is it in general important to test for non-stationarity in time series data before attempting to build an empirical model?

(c) Define the following terms and describe the process that they represent:

i. Weak stationarity

ii. Strict stationarity

iii. Deterministic trend

iv. Stochastic trend