ECOM90024
Forecasting in Economics and Business
Tutorial 8

---

1. Consider the following AR(4) process,

$$Y_t = \mu + \delta t + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \phi_4 Y_{t-4} + \varepsilon_t$$

   a.) Given the above process, derive a regression model that will allow you to test whether the process possesses a unit root. In doing so, make sure to derive expressions of the regression coefficients in terms of the parameters of the above process.

   b.) Having derived the regression equation, explain how it can be used to test for the presence of a unit root. Make sure to write down the elements of the test (i.e., null and alternative hypotheses, test statistic and decision rule).

2. Using the R statistical environment, simulate 1000 observations from an AR(4) (with no mean or deterministic trend) using a set of coefficients generated from the generateAR() function from the DREGAR package (see R-Week5.pdf)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \phi_4 Y_{t-4} + \varepsilon_t$$

   a.) Install and load the urca package. The function ur.df() performs an Augmented Dickey Fuller Test using the following syntax:

```
ur.df(y, type = c("none", "drift", "trend"), lags = k)
```

   Where `y` is the time series data, `type` is the form of the ADF regression where

$$\texttt{"none"} = \Delta Y_t = \rho Y_{t-1} + \sum_{j=2}^{k} \beta_j \Delta Y_{t-j+1} + \varepsilon_t$$

$$\texttt{"drift"} = \Delta Y_t = \mu + \rho Y_{t-1} + \sum_{j=2}^{4} \beta_j \Delta Y_{t-j+1} + \varepsilon_t$$

$$\texttt{"trend"} = \Delta Y_t = \mu + \delta t + \rho Y_{t-1} + \sum_{j=2}^{4} \beta_j \Delta Y_{t-j+1} + \varepsilon_t$$

   and `lags` is the number of lags to include in the ADF regression. Use the ur.df() function to perform an ADF test on your simulated data. Make sure to explain the lag that you've chosen to specify in the function. Do your results conform to your expectations?

   b.) Using the data that you have simulated, construct a new time series Z_t that is defined as:

$$Z_t = \sum_{s=0}^{t} Y_s$$

   (Hint: you can use the cumsum() function).

Generate a plot of Z_t and describe what you see. Use the ur.df() function to perform an ADF test on Z_t. You can use the same lag specification used in part (a). Do your results conform to your expectations?

3.) Download the data contained in the file sp500.csv into your R environment.

a.) Generate a plot of the data and briefly describe the primary visual characteristics of the data?

b.) Generate sample ACF and PACF plots. What are they telling us about the dependence structure of the data?

c.) Perform an ADF test using the ur.df() function in which `type = "drift"`. Set the lag length using the method of Ng & Perron discussed in class. In the output, the relevant test statistic is the first number that is reported. Compare this test statistic with the critical value labelled tau2. Do your results conform with your results in parts (a) and (b)?

d.) Compute the continuously compounded return on the S&P 500 index using the following formula:

$$r_t = 100 \times (\ln(P_t) - \ln(P_{t-1}))$$

Plot the returns and repeat part (c) using the continuously compounded returns that you've just generated and report your results.