TIME SERIES ANALYSIS - LAB SESSION 5
ESILV : Master 1 - 2023/2024
Dynamic model : Unit root tests and ARIMA(p,d,q)

Instructions : Your work must be sent to your tutor 7 days after your in-class lecture (deadline t+14 at 10.00 p.m). Any work handed in late, or in word format, will not be considered and read. Your submission will include an R-Markdown file, its compiled version in HTML or PDF. Note that a Tex version is appreciated (it is not mandatory). Particular care should be taken in the presentation of your results (avoid displaying a 1000 points vector in your report for example), your graphics and especially your comments.

Part 1 : Warm up - Data settings and unit root tests

Exercise 1 : Data settings - 2.5 pts

1. Data importation can be automated using R API. There exists several API facilitating economic and financial data import and update. Load the unemployment rate from the Federal Reserve of Saint Louis [1] ? (0.5 pt)

2. Check the status of the imported data and transform it into a right object if necessary. Plot the data. What do you observe ? What type of seasonal pattern is ? What type of the filters do you propose to clean the retail sales (0.5 pt).

3. Run the required filter to cut the seasonal pattern ? You can choose between the seasonal regression, the moving average filter and the decompose function (this function is based on moving average seasonal filters as well) (1 pt).

4. Grab the filtered data and check using the right tool the seasonal pattern has been deleted. Is this filtered data be modeled using an ARMA(p,q) approach (0.5 pt) ?

Exercise 2 : Unit Root tests - (2.75 pts)

We can notice the seasonally adjusted data (or the filtered one) is still showing a cyclical pattern. We need to characterize this pattern, i.e. whether the time series is stationary or not and what is the nature of the non-stationarity,if so.

1. Load the "urca" package. Why focusing on non-stationarity is crucial in time series analysis. Summarize quickly the step-wise approach of the Dickey Fuller test.

2. Compute the Augmented Dickey Fuller test using the right function on the filtered data (note that the selection of the number of lags can be performed on a discretionary way or automatically).(1 pt)

3. Determine then the integration degree of the data ? (0.25 pts)

1. The data is available under the Fed of Saint Louis website. The code for the unemployment, Monthly, Not Seasonally Adjusted is **UNRATENSA**

4. Compute the Phillips and Perron test on the filtered data. Does it confirm your previous result ? (0.5 pt)

5. Compute the KPSS test using one of the artificial data generated previously. Do we find the same conclusion ? 0.5 pt)

6. Find the degree of integration of the US retail sales using the KPSS test. Does it validates the previous result 0.5 pt) ?

Exercise 3 : Modeling - (2.25 pts)

1. Given the results derived from the previous sections, propose the most relevant ARMA(p,q) framework to model the retail sales dynamics. Is there an alternative to the ARMA(p,q) approach which directly deal with non-stationarity ? (0.5 pt)

2. Having choose the correct specification, justify the relevance of your choice and run the required quality check tests to validate you choice. (1.75 pts).

Exercise 4 : Estimating an ARIMA(p,d,q) - (6 pts)

As seen during the class, ARIMA model were designed to deal with non stationary time series. We propose to use this kind of specification to model the Johnson & Johnson stock price from 1997 until now, on a monthly basis. Prior moving to the stock price modeling, we nee to load the data from yahoo finance website. To do so, run the following code

```
install.packages("tidyquant")
library(tidyquant)
jnj = tq_get("JNJ", get="stock.prices", from="1997-01-01") %>%
tq_transmute(mutate_fun=to.period,period="months")
```

1. Determine the degree of integration of the Johnson & Johnson stock prices. (1 pt)

2. Determine the order of the ARIMA model, i.e the values of p, d, q to be used to model the stock prices. Note first, the assessment of p and q cannot be performed on the data expressed in level. Note besides, the determination of the values of p and q can be performed using different methods (graphical approach vs information criteria). (1.5 pts)

3. Estimate the corresponding ARIMA(p,d,q) model to the values of p,d,q selected previously. Check the estimated coefficients and compute the fit of the model. Plot (within the same chart) the estimated values of the stock price and the observed one. (1.25 pts)

4. Calculate the residual of the model, given as the difference between $J\hat{\&}J$ and $J\&J$. Compute the required quality checks on the residuals (1.75 pts).

5. Using the estimated coefficients, generate a forecast over the next 3 months. Calculate the confidence interval of the forecasted points (0.5 pt) .

Exercise 5 : Unit root test another one - 4 pts

In this exercise, we propose to introduce a new unit root test : the test of Zivot and Andrews (1992). On of the weakness of the ADF unit root tests is their potential confusion of structural breaks in the series as evidence of non-stationarity. In other words, they may fail to reject the unit root hypothesis if the series have a structural break.

Zivot and Andrews (1992) endogenous structural break test is a sequential test which utilizes the full sample and uses a different dummy variable for each possible break date. The break date is selected where the t-statistic from the ADF test of unit root is at a minimum (most negative). Consequently a break date will be chosen where the evidence is least favorable for the unit root null.

The Zivot-Andrews (1992) tests state the null hypothesis is that the series has a unit root with structural break(s) against the alternative hypothesis that they are stationary with break(s). We reject Null if t-value statistic is lower than tabulated critical value

.

1. Present the Zivot and Andrews (1992) test paying attention to the nature of the breaks. Explain the strategy of the test (0.5 pt).

2. Generate 3 new random walks. The first one is a pure random walk, the second is a random walk with a break in level and the third on will be a random walk with both a break in level and in the trend.(0.75 pt)

3. Compute the appropriate Zivot and Andrews (1992) test for the generated random walk. Summarize your output within a table.(1.25 pts)

4. Is it relevant to use such test for the filtered retail sales. Justify. Compute the Zivot and Andrews (1992) unit root test using the US retail sales (1 pt).

Exercise 6 : Modeling the business cycle - 4,5 pts

Based on all your knowledge, propose the most appropriate specification to model the monthly credit spread dynamics in the US. The data is available under the Fed of Saint Louis website. You have to load the Moody's Seasoned Aaa Corporate Bond and the Moody's Seasoned Baa Corporate Bond Yield years interest rate. The credit spread is the difference between the Baa index and the Aaa one.

Produce a complete analysis (seasonality, stationarity, modeling, residual checking and forecasting of the credit spread. The forecasting horizon is set at three months. The study should incorporate illustrating charts, detailed and motivated comments and relevant results.