1 Model Identification

In this practice we will work on the identification of ARMA(p,q) models on stationary series. As has been done in the previous practices, each group of students must carry out the script of the practice below, performing the calculations and graphics needed in the R environment.

2 Questions

In this quiz we continue to work on some of the time series from the previous session that you can find in the series repository in the subject area at ATENEA. We are adding two new series:

- `PIBsp`: (https://www.ine.es) **Gross Domestic Product-GDP) of Spain**.These are quarterly data that represent the chained volume in dollars (2009) in billions and are seasonally adjusted.
- `PreuIndEner`: (https://www.ine.es/jaxiT3/Tabla.htm?t=27065) **Industrial Price Index for energy** based on the year 2015, from January 2000 to December 2018. The Industrial Price Index (IPRI) is a short-term indicator that measures the monthly evolution of the prices of industrial products manufactured and sold in the internal market, in the first step of their commercialization, that is to say, the sale prices at the factory, excluding transport and commercialization expenses and invoiced VAT.

The series to be worked on will be: GDPUSA, EURODOL, PIBsp and PreuIndEner. For each series you must:

1. Upload the file containing the series. Define the read data as an object of type `ts` (time series) indicating the origin and frequency of the series.
2. Graph the time series. Describe the most relevant aspects that can be seen.
3. Apply the appropriate transformations to make the series stationary.
4. For the transformed series, plot the ACF and PACF
5. Based on the sample ACF and PACF, propose at least two models for each series, justifying the proposal.
6. Estimate the proposed models and verify the significance of the coefficients and that the residuals have an ACF compatible with white noise. If there is any non-significant coefficient, remove it from the model.
7. Indicate which model you would propose, using the AIC criterion.

3 Hints

Interpretation of the ACF/PACF

- It is preferable to identify simple models with few parameters (principle of parsimony). It is most common to start with AR(p) or MA(q) type models only. If no suitable models are found, then try ARMA models with few parameters: ARMA(1,1), ARMA(2,1), ARMA(1,2)...

- A delay within the confidence bands can be considered zero or different from zero but with a small magnitude. This makes it possible to interpret a graph where the delays are within the bands from a certain point as if there were a finite number of delays or as if it were a decreasing pattern with infinite non-zero delays
- A delay that lies outside the bands indicates that we can consider it non-zero with 95% confidence. If the delay is far and there is no seasonality, considering it significant may lead to proposing a model with many parameters. In this case it may be preferable to consider that it has left the bands by chance and is part of the 5% spurious significance.

Model Estimation

- Use the `arima` method, indicating in the `order` parameter a vector of the form `c(p,0,q)` to estimate an ARMA(p,q) model on the stationary series .
- By default, if the series is stationary the mean parameter (misnamed `intercept`) is also estimated. If you don't want to estimate, you can use the parameter `include.mean=F`.
- The estimate provides the value of the estimators and their standard errors. This allows solving the significance test of each coefficient with the corresponding t-ratio statistic.
- If a parameter is not significant and it determines the value of p or q, it is only necessary to re-estimate the model by reducing the corresponding order. Instead, if it is a parameter of the characteristic polynomial and does not determine its degree, the `fixed` parameter must be used. This parameter must be a vector of the same length as the number of parameters where the NA value indicates that it is to be estimated and any other value sets the parameter to that value.