1 Identification of seasonal models

In this practice we will work on the identification of $ARIMA(p,d,q)(P,D,Q)_s$ models on non-stationary series that have a seasonal component. As has been done in the previous practices, each group of students must perform the script of the practice below, performing the necessary calculations and graphics 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:

- `ATUR`: https://expinterweb.mites.gob.es/series/) **individuals registered as unemployed at INEM offices**
- `AIRBCN`: (https://apps.fomento.gob.es/BoletinOnline/?nivel=2orden=03000000) **monthly passengers of international flights in El Prat Airport (BCN)**.

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 source 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 for seasonal models

- It is important to be able to identify lags that are multiples of seasonality, as they are the basis for proposing values for the seasonal $ARMA(P,Q)_s$ part. Therefore, if we use the standard methods `acf` and `pacf` it is convenient to add the parameter `col` which assigns color in order to each bar. Specifically, for the ACF they would use the instruction: `acf(w,col=c(2,rep(1,11)))` and for the PACF, which starts with the first delay the parameter should be:`pacf(w,col=c(rep(1,11),2))`.
- For the seasonal part of $ARMA(P,Q)_s$, we only need to use the multiple lags of the seasonality. If we use the above code, we only need to look at the red delays, and apply the identification criteria used with the delays for stationary ARMA models.

- For the regular part $ARMA(p, q)$ we only need to consider the first lags (if the seasonality is of order 12, we can consider only the first 5 or 6 lags). It must be taken into account that near the delays of order multiple of seasonality (those of order 12, 24, 36,... in this case) there may be significant satellite delays, which should not be considered for identification

Model Estimation

- In the estimation process, we must first state the transformed series into stationary (W_t) to obtain the estimate of the mean:

  `arima(w,order=c(p,0,q),seasonal=list(order=c(P,0,Q)))`

  If the mean is not significant, we re-estimate the model for the original series (X_t or $\log(X_t)$), indicating the differences in the estimation method:

  `arima(x,order=c(p,d,q),seasonal=list(order=c(P,D,Q)))`