STAT 443: Time Series and Forecasting

Lab 9: Forecasting - Part II

- The lab must be completed in `R Markdown`. Display all the `R` code used to perform your analysis.
- Create a pdf or html file and use it as your lab submission.
- Please ensure that the file you submit is in good order (e.g., not corrupted and contains the work you intend to submit). No late (re-)submissions will be accepted.

In this lab you will apply the Box–Jenkins forecasting method using a case study. The "lake levels" data is a built–in dataset in R. It measures the annual level of Lake Huron from 1875 to 1972, in feet. The data in `data(LakeHuron)` are already in time series format. Using the `window` command or otherwise, first subset the Lake Huron data into two separate datasets: `LakeHuron.train`, which should contain the first 93 lake levels, and `LakeHuron.test` containing the last 5 lake levels. We will use `LakeHuron.train` to "train" a model and forecast the next three points based on the fitted model.

1. The data may show a slightly decreasing trend. Ignore this possible decreasing trend in what follows. Plot the `LakeHuron.train` data, its acf and pacf. Determine an appropriate ARMA model and explain your choice.
2. Fit the model you chose above using the function `arima()`. Write down the fitted model.
3. Examine appropriate diagnostics for your fitted model. Specifically, first plot the acf of the residuals to see if there are significant autocorrelations after lag 0. Then, use `tsdiag()` to observe the standardized residuals and the p-values of the Ljung-Box version for the portmanteau test. Report what you observe and then comment on the fit.
4. Use the `predict` command in R to forecast the Lake Huron level for the next three years, i.e., 1968, 1969, and 1970. Provide 95% prediction intervals for each forecast.
5. Compare the forecast with the true values (in `LakeHuron.test`). Comment on what you find.