ASSIGNMENT 9

STAT 485/685 E100/G100: Applied Time Series Analysis
Fall 2020
Simon Fraser University

This week's assignment will be about model diagnostics and forecasting with time series models. The topics here have been covered in the Videos 32-34, as well as Chapters 8-9 of the textbook *Time Series Analysis with Applications in R (2nd ed.)* by Cryer & Chan.[1]

Due date: **Tuesday, Dec. 8th at 11:59 pm (end of day)** (Pacific Time).

Marks: 10.

Please include your R code for each of questions that requires it below. Some ideas for how you can most easily do this:

- Copy-paste your code/plots into a Word document along with your responses to the questions, and save as a PDF.
- Take images of your code/plots and upload to Crowdmark, along with your responses to the questions.
- Save your code and responses together in an RMarkdown document and save as PDF (if you've worked with RMarkdown before).

Other important policies on assignment submissions:

- Please write each question on a **separate page**!
- Please **show all your code and work**, in order to get full marks.
- Upload your complete answers as PDF files or high-resolution images.
- If you're hand-writing answers, please make sure they are **neat and clearly readable**, and that the photo is high resolution.
- Please **clearly label the question numbers**.

[1] Cryer, J. D., & Chan, K. S. (2008). *Time series analysis: with applications in R.* Springer Science and Business Media.

For this assignment, we will be using some datasets and functions in the `TSA` package in R. For instructions on how to install and load the package, please see the Week 2 module on Canvas.

1. *(4 marks)* The `gold` dataset in the `TSA` package gives the daily price of gold (in $ per troy ounce) for the 252 trading days of 2005.

   (a) Suppose we were to try fitting an AR(1) model to this dataset. Fit this model using the `arima()` function in R. Give the estimates of ϕ and μ.
   (Note: As we saw in Video 32, the coefficient named "intercept" in the ***`arima()`*** *output is actually referring to the mean μ, NOT the intercept θ_0.)*

   (b) Create a plot of the (standardized) residuals vs. time for this model. Interpret what you see in the plot.

   (c) Create a Q-Q plot of the (standardized) residuals for this model. Interpret what you see in the plot.

   (d) Create the sample ACF plot of the (standardized) residuals for this model. Interpret what you see in the plot.

2. *(6 marks)* The `units` dataset in the `TSA` package gives the annual sales of certain large equipment, 1983-2005.

   (a) Fit an MA(2) model (with a potentially non-zero constant mean) to this data using the `arima()` function in R. Give the estimates of the parameters θ_1, θ_2 and μ.
   (IMPORTANT: The way the ***`arima()`*** *function defines the MA model is by placing plus signs, instead of minus signs, in front of the MA parameters. Therefore, the values of the MA parameters given in this output are actually $-\theta_1$ and $-\theta_2$!)*
   (Note: As we saw in Video 32, the coefficient named "intercept" in the ***`arima()`*** *output is actually referring to the mean μ, NOT the intercept θ_0.)*

   (b) Using the methods practiced in Video 33, derive the equation for the forecast of $Y_{t+\ell}$ at any lead time ℓ. Make sure to replace any parameters with the estimates you obtained in part (a).

   (c) Derive the equation for the forecast error variance for $Y_{t+\ell}$, denoted by $Var(e_t(\ell))$. Make sure to include each possible case of values that ℓ can take on.

(d) Using your equation in part (b), obtain the forecast of Y_{t+1}. Show your calculations.
(Note: You can use R's estimates of the noise terms to help you out. The estimates of $e_1, \ldots, e_t$ can be found in the object `name_of_your_ma2_model$residuals`*.)*

(e) Using your forecast in part (d), and the equation in part (c), calculate the the 95% prediction limits for Y_{t+1}.
(Note: You can use R's estimate of the white noise variance if you need it. It can be found in the object `name_of_your_ma2_model$sigma2`*.)*

(f) Create a plot of the predictions of $Y_{t+\ell}$ out to 20 time points in the future. Does the forecast for $\ell = 1$ match your results above?
(Hint: You can directly read the values off the plot or, if you'd like exact values, you can extract them by adding `$pred`*,* `$lpi` *or* `$upi` *after the* `plot()` *function. This will give you the forecasts, and lower and upper 95% prediction limits, respectively.)*