STAT 520 – Homework 5 – Fall 2023

Note: For Homework 5A, do Problems 1 and 2 below.

Homework 5B consists of Problems 3 and 4, but note that Problem 3 is mandatory for graduate students and extra credit for undergraduates.

1) A data set of 57 consecutive measurements from a machine tool are in the `deere3` object in the `TSA` package. Type `library(TSA); data(deere3); print(deere3)` in R to see the data set.
(a) Fit an AR(1) model (which includes an overall mean μ) and write the estimated model equation (specifying the parameter estimates). Use it to forecast the next ten values of the series, and list the forecasted values. Also plot the series along with the forecasted values and 95% prediction limits for the next ten values of the series.

2) A data set of durations until payment for 130 consecutive orders from a Winegrad distributor are in the `days` object in the `TSA` package. Type `library(TSA); data(days); print(days)` in R to see the data set.
(a) Fit an MA(2) model (which includes an overall mean μ) and write the estimated model equation (specifying the parameter estimates). Use it to forecast the next ten values of the series, and list the forecasted values. Also plot the series along with the forecasted values and 95% prediction limits for the next ten values of the series.
(b) There are three clear outliers (at times 63, 106, and 129) which can be seen from the time series plot. Replace these outliers with a "typical duration" value of 35. To do this, you can use the code
`days.adj=days; days.adj[c(63,106,129)]=35; print(days.adj)`
Fit an MA(2) model to this adjusted series (which includes an overall mean μ) and write the estimated model equation (specifying the parameter estimates). Use it to forecast the next ten values of the series, and list the forecasted values. Also plot the series along with the forecasted values and 95% prediction limits for the next ten values of the series.
(c) Comment on any differences between the forecasts based on the original data and the forecasts based on the adjusted data.

3) A data set of 324 measurements of an industrial robot's positions are in the `robot` object in the `TSA` package. Type `library(TSA); data(robot); print(robot)` in R to see the data set.
(a) Fit an IMA(1,1) model and use it to forecast the next five values of the series, and list the forecasted values. [You can verify that for this data set, you do NOT need to include a constant (intercept) term in the IMA model.] Also plot *the last ten observed values of the series* along with the forecasted values and 95% prediction limits for the next five values of the series. [Hint: type `help(plot.Arima)` and look at the `n1` argument of the `plot` function.]
(b) Fit an ARMA(1,1) model and use it to forecast the next five values of the series, and list the forecasted values. Also plot *the last ten observed values of the series* along with the forecasted values and 95% prediction limits for the next five values of the series.
(c) Compare the results from parts (a) and (b).

4) A data set of monthly electricity generation values is in the `electricity` object in the `TSA` package. Type `library(TSA); data(electricity); print(electricity)` in R to see the data set.
(a) Fit a seasonal means model which also contains a linear time trend on the (natural) *logarithms* of the data. Some R code to help with this is:

```
data(electricity)
month.=season(electricity)
model.s=lm(log(electricity)~time(electricity) + month.-1)
# -1 removes the intercept term
summary(model.s)
```

Use it to forecast the next three months of the series (January 2006, February 2006, and March 2006), and list the forecasted values. [Hint: This is just a deterministic model; no ARIMA-type forecasting is needed. When predicting using the estimated trend model, you will need to plug in the value of *t* (which may include a fraction of a year) corresponding to the month you are predicting for, and plug in a "1" for the indicator variable for the month of interest and plug in a "0" for the indicator variables for all other months.]
(b) Give the forecasted values for the next three months in terms of the original data, i.e., not in logged values.