INDR 422/522 Spring 2023 Fikri Karaesmen, Bijan Bibak

LAB EXERCISES 2, March 17, 2023

We'll work with different data files for this lab. The files are available on the blackboard page.

1. Transformations: differences

   (a) Download the metro passengers data from the file. The data includes the daily number of passengers that take the Istanbul M2 metro line starting from Haciosman Station. The data is over 45 weeks starting on January 1, 2022 (Saturday) and ending on November 11, 2022.

   (b) Plot the data.

   (c) Take a first difference of the data and plot the differences. Find the mean of the differences. What this does correspond to?

2. Transformations: seasonal differences

   (a) Use the passengers data to investigate seasonality (weekly)

   (b) Plot the data.

   (c) Take a seasonal difference of the data and plot the deasonalized data.

3. To obtain benchmark, perform a naive forecast ($\hat{y}_{t+1} = y_t$) on the data. Assess the error performance.

4. Perform an exponential smoothing forecast (try different values of smoothing constants α) on the differenced and deasonalized data and assess the error performance. Note that we go through the transformations $z_t = y_t - y_{t-1}$ and $u_t = z_t - z_{t-7}$. Our exponential smoothing forecast then obtains $\hat{u}_{t+1}$. From that we have: $\hat{z}_{t+1} = \hat{u}_{t+1} + z_{t-6}$. Finally, $\hat{y}_{t+1} = \hat{z}_{t+1} + y_t$.

5. Transformations: log transformations

   (a) Download the new covid cases data in Italy (from 2020) from the file .

   (b) Plot the data.

   (c) Take a log of the data and plot the log of the data.

(d) Take a seasonal (i.e. 7 day) difference of the log data and plot the desasonalized and logged data.

6. Generate 500 realizations from an AR-1 process $Y_t = 200 + (0.9)Y_{t-1} + \epsilon_t$ where ϵ_t are independent normal random variables with mean zero and standard deviation 20.

   (a) Plot the generated observations.

   (b) Plot the Auto-Correlation Function.

7. Generate 500 realizations from an AR-1 process $Y_t = 200 + -0.9Y_{t-1} + \epsilon_t$ where ϵ_t are independent normal random variables with mean zero and standard deviation 20.

   (a) Plot the generated observations.

   (b) Plot the Auto-Correlation Function.

8. Generate 500 realizations from an MA-1 process $Y_t = 200 + (0.9)\epsilon_{t-1} + \epsilon_t$ where ϵ_t are independent normal random variables with mean zero and standard deviation 20.

   (a) Plot the generated observations.

   (b) Plot the Auto-Correlation Function.

9. Using the monthly Australian Beer Production data:

   (a) Plot the Auto-Correlation Function of the original series.

   (b) Plot the Auto-Correlation Function of the deasonalized series.

   (c) Implement an exponentially smoothed seasonal difference based forecast (why would we want to do that based on the previous analysis?) $\hat{y}_t = \alpha y_{t-12} + (1-\alpha)\hat{y}_{t-12}$. Compute the error performance.

Exercises to Complete Later

10. Generate 500 realizations from an AR-2 process $Y_t = 200 + (0.5)Y_{t-1} + (0.4)Y_{t-2} + \epsilon_t$ where ϵ_t are independent normal random variables with mean zero and standard deviation 20.

    (a) Plot the generated observations.

(b) Plot the Auto-Correlation Function.

11. Generate 500 realizations from an MA process $Y_t = 200 + (0.8)\epsilon_{t-2} + \epsilon_t$ where ϵ_t are independent normal random variables with mean zero and standard deviation 20. Note that this is a special MA-2 process where the coefficient of the first lag $\theta_1 = 0$.

    (a) Plot the generated observations.

    (b) Plot the Auto-Correlation Function.

12. Using the Covid cases data:

    (a) Plot the Auto-Correlation Function of the original series.

    (b) Plot the Auto-Correlation Function of the logged series.

    (c) Plot the Auto-Correlation Function of the deasonalized and logged series.