INDR 372 - Spring 2024 Fikri Karaesmen

HOMEWORK 1, Due Date: March 18, 2024

- Please work **in groups of two** or **individually** and submit one file for each group with both names.
- For this homework please perform all computations in a python notebook. Please don't use forecasting packages or functions, you are expected to implement your own forecasts. Please submit one python notebook file that clearly shows all the computations.
- In addition to the the notebook, **submit a short typed summary report** that includes the results (error tables, prediction intervals etc.) of all exercises. Also add a general assessment of the methods (which method is the best, which should be avoided etc.). **The report is part of the overall grade and must be written and formatted clearly.**

Exercises

1. Forecasting sales of Renault vehicles. The data file contains the monthly domestic sales of total monthly sales of Renault brand cars in Turkey from the beginning of 2013 to the end of 2022 (found in the blackboard page).

   (a) Plot the data and visually assess whether there is significant trend and seasonality.

   (b) To obtain a benchmark for errors, implement the following naive forecasts i) $F_t = D_{t-1}$ for $t > 1$ and ii) $F_t = D_{t-12}$ for $t > 12$. Report the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE) of these forecasts for years 2019 until the end of 2022. These error measures constitute a simple benchmark for all other approaches (i.e. hopefully you will obtain lower errors by more sophisticated methods). Note that different methods require different initialization periods. To be consistent, we start forecasting as early as possible in 2013 but start comparing the errors from January 2019 onwards.

(c) Use a 3-period moving average to forecast the one month-ahead monthly demand. Report the MAE, MAPE and RMSE of the forecast for years 2019 through 2022. Report 90 percent prediction intervals (using the RMSE estimated in years 2014 to 2018) for the one-month ahead forecasts for year 20 22.

(d) Comment on the residual diagnostics (i.e. independence and normality of residuals). What is the drawback of this forecast with respect to this data?

(e) Use exponential smoothing to forecast the one period ahead monthly demand. Experiment with at least 10 different smoothing constants ($\alpha = 0.1, 0.2, ..., 1$) and for the best smoothing constant, report the MAE, MAPE and RMSE of the forecast for years 2019 through 2022. Report 90 percent prediction intervals (using the RMSE estimated in years 2014 to 2018) for the one-month ahead forecasts for 2022. How do these compare with MA-3 forecasts?

(f) Use Double Exponential Smoothing to forecast the one period ahead monthly demand. Experiment with different values of the smoothing constants ($\alpha = 0.1, 0.2, ..., 1$, $\beta = 0.1, 0.2, ..., 1$) and for the best smoothing constants report the MAE, MAPE and RMSE of the forecast for years 2019 through 2022. Report 90 percent prediction intervals (using the RMSE estimated in years 2014 to 2018) for the one-month ahead forecasts for 2022.

(g) Comment on the resdual diagnostics (i.e. independence and normality of residuals). Comment on the comparison to the previous forecasts? What is the drawback of this forecast with respect to this data?

(h) In part f, you must have found the smoothing constants, α^* and β^* that leads to the best error performance. Use α^* and β^* you found in part f to find 3-month and 6-month ahead forecasts for year 2022. Report the MAE, MAPE and RMSE of these forecasts for 2022 and comment on the differences.

(i) To take into account the effect of seasonality, perform the following data transformation: $U_t = D_t - D_{t-12}$. Plot the transformed series (U_t) and visually verify whether seasonality is eliminated. Use simple exponential smoothing to find a forecast G_t for U_t (make sure to test different values of the smoothing parameter to minimize the error). To obtain a forecast F_t for D_t, you can then consider $F_t = G_t + D_{t-12}$ or a smoothed version $F_t = G_t + \gamma D_{t-12} + (1 - \gamma)F_{t-12}$ for a smoothing constant

$0 \leq \gamma \leq 1$. For the best one-month ahead forecast, report the MAE, MAPE and RMSE of the forecast for years 2019 through 2022.

(j) Your homework report should include a table similar to the one below.

Method	**Spec.**	**RMSE**	**MAPE**
Benchmark 1	-		
Benchmark 2	-		
MA-3	-		
ES	-		
DES	-		
Seasonal			

Note that the model specification for exponential smoothing is: $\alpha^* = ...$, $\beta^* = ...$.

2. Forecasting sales of domestic sales of beer. The data file contains the total monthly sales of beer in Turkey from the beginning of 2010 to the end of 2014 (found in the blackboard page).

(a) Plot the data and visually assess whether there is significant trend and seasonality.

(b) To obtain a benchmark for errors, implement the following naive forecasts i) $F_t = D_{t-1}$ for $t > 1$ and ii) $F_t = D_{t-12}$ for $t > 12$. Report the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE) of these forecasts for years 2011 until the end of 2014.

(c) Comment on the residual diagnostics (i.e. independence and normality of residuals). What is the drawback of this forecast with respect to this data?

(d) Use Triple Exponential Smoothing to forecast the one period ahead monthly demand. Experiment with different values of the smoothing constants ($\alpha = 0.1, 0.2, ..., 1$, $\beta = 0.1, 0.2, ..., 1$, $\gamma = 0.1, 0.5, 0.9$) and for the best smoothing constants (the ones that lead to the best error) report the MAE, MAPE and RMSE of the forecast for years 2011 through 2014. Report 90 percent prediction intervals (using the RMSE estimated in years 2011 to 2013) for the one-month ahead forecasts for 2014.

(e) Use the best smoothing constants that you found in the previous exercise to find 3-month and 6-month ahead forecasts for year 2014. Report the MAE, MAPE and RMSE of these forecasts for 2014 and comment on the differences.

(f) Your homework report should include a table similar to the one below.

Method	**Spec.**	**RMSE**	**MAPE**
Benchmark 1	-		
Benchmark 2	-		
TES			

Note that the model specification for exponential smoothing is: $\alpha^* =$ ..., $\beta^* =$ ... .