ID 357

Forecasting of Female Labour Force Participation Rate Data with Missing Values Imputation, Sri Lanka

NN Ranasinghe[1#] and RAB Abeygunawardana[1]

[1] *Department of Statistics, Faculty of Science, University of Colombo, Sri Lanka*

[#] nuwanir28@gmail.com

Female Labour Force Participation Rate (Female LFPR) is defined as the proportion of the female labour force to the total working age population. This study is based on the female LFPR quarterly data published by the Department of Census and Statistics, Sri Lanka during the period 2004 to 2021. However, it was found that data for eight quarters in the above time period are missing. The main objective of this study is to forecast female LFPR using ARIMA models by imputing the missing values. In the first part of the analysis, missing values were imputed using nine imputation algorithms available in "*imputeTS*" package in R software. Missing values were generated under four missing rates and thirty random seeds. By comparing MAPE and RMSE plots the Exponential Weighted Moving Average (EWMA) method was found to be the best imputation method. In the second part of the analysis female LFPR were forecasted using ARIMA models. In this analysis, the data were divided into two parts as training and test data. In the training data set, trend, seasonal and random components were identified using "*decompose ()*" function in R software. Furthermore, functions "*arima ()*" and "*auto.arima()*" in library "*forecast*" in R software were used to fit ARIMA models. It was found that ARIMA(1,1,1) model without drift was the best model to forecast the female LFPR which has the minimum AIC value. Errors for the fitted values were calculated using the test data. Female LFPR for the next ten quarters were forecasted using the ARIMA(1,1,1) model. Results showed a small increment in female LFPR at the end of 2022.

Keywords: *female LFPR, ARIMA, missing value imputation*