Appendix 3: SARIMA Model Details

University of Manitoba | Rady Faculty of Health Sciences

Introduction

Times series data are data collected at multiple and equally spaced time points (daily, monthly, etc). Such a series can exhibit seasonality, trend over time, and serial correlation between observations. Time series data can be classified into stationary and non-stationary series.

A stationary time series is one whose statistical properties such as mean, variance, and autocorrelation are all constant over time. In other words, a stationary time series is one whose statistical properties do not depend on the time at which the series is observed. Such a series is relatively easy to predict because we simply assume that the statistical properties will remain the same in the future as it was in the past. When the series exhibits some trend, or has means or variances that change over time, then they are said to be non-stationary. Non-stationary data are unpredictable and hence cannot be modelled or forecasted. To model such a series, we first transform them into stationary series. Transforming non-stationary data into stationary can be done by differencing, if the non-stationary data is due to trend or seasonality.

ARIMA and SARIMA models

ARIMA stands for Autoregressive (AR), Integrated (I) Moving Average (MA) models. These models are usually denoted as ARIMS(p,d,q), where p is the order (number) of the Autoregressive (AR) process, d is the order (number) of the differencing required to make the non-stationary process stationary, and q is the order (number) of the Moving Average (MA) process. If seasonality is present, then we have SARIMA process which is usually denoted as ARIMA(p,d,q)(P,D,Q) s, where s represents the seasonal length in the data (e.g. 12 for monthly data), and the lowercase and uppercase notations represent the non-seasonal and seasonal components of the model respectively. For instance, p and seasonal P indicate the number of autoregressive terms (lags of the stationary series), d and stationary D indicate the number of differencing that is done to make the time series stationary; q and seasonal Q indicate the number of moving average terms (lags of the forecast errors).

ARIMA/SARIMA models are commonly used for prediction or forecasting series that exhibit trend/seasonality. The forecasts in these time periods give an idea of the expected need, and from policy perspectives, can be used for planning purposes. The difference between the expected needs (from forecasts) and the actual surgeries can give an idea of the unmet needs.

Modeling Process Employed: Identifying ARIMA/SARIMA Model

We created a monthly time series data for both orthopedic and ophthalmology procedures. Our study period is from April 1, 2004, to March 31, 2020. Our training dataset consists of April 1, 2004, to March 31, 2018, and validation dataset is from April 1, 2018, to March 31, 2020.

There are generally three stages in the modeling of time series data. Identification stage, estimation and diagnostic checking stage, and the forecasting stage.

Identification Stage

This is the stage in which the right candidate model to be fitted is identified. To do this, the following steps are followed

Step 1: Plot the time series

Plot the series to see if there is any trend or seasonality in the data. This can be done using PROC TIMESERIES. If only trend is present without seasonality, then consider fitting ARIMA models. Otherwise, if data has seasonality, then consider fitting SARIMA models. By using Box-Cox transformation method, determine if any transformations of the data are necessary.

Step 2: Difference the series

If there is evidence of trend or seasonality, then difference the series. For instance, if the data is of a monthly series, then take the 12th difference to difference the series. If there is a linear trend, then take the 1st difference. If there is both linear trend and seasonality on monthly data, then take both 1st and 12th difference. Examine the Autocorrelation Function (ACF)/Partial Autocorrelation Function (PACF) to see if data is now stationary.

Step 3: Examine the ACF/PACF plots of the differenced data

Examining the plots of the ACF and PACF is useful in determining the possible models to be estimated. This step can involve some educated guesses.

Estimation and Diagnostic Checking Stage

Model estimation at this stage is based on the differenced data. The Akaike Information Criterion (AIC) and Schwartz Bayesian Criterion (SBC) are used to select reasonable models. The model with the least AIC or SBC is selected. Check also to see if the autocorrelation of the residuals is white noise. Additional statistics that could be employed in model selection are Mean Absolute Percentage Error (MAPE), the Root Mean Squared Error (RMSE), and the Mean Absolute Error. The selected models for the different outcomes considered in this project are presented at the end of this appendix.

Forecasting Stage

This is the stage where the selected model is employed in forecasting future surgeries. It is important to reiterate that these models are best or most reliable for forecasting short time periods.

Population Adjustment

Our initial intent was to include the projected population estimates for our forecast period in the modeling process. The estimates we have were done about 20 years ago. Any forecasts that long are usually unreliable. Hence, this project uses a multiplier that is based on the actual trend of Manitoba population growth over time. We focused on the trend for those 60 and over in the period 2012 to 2020. We regressed the current population on the past to determine the rate of growth. This rate of growth was used to adjust the forecasts to obtain what we call population adjusted forecasts.

Models Selected for the Outcomes

Orthopedic Procedures	
Outcome	**SARIMA Models**
Ortho_Overall	ARIMA(1,1,1)(1,1,1)12
Ortho_Winnipeg	ARIMA(0,1,1)(0,1,1)12
Ortho-West Eastman	ARIMA(1,1,1)(1,1,1)12
Ortho-Southern	ARIMA(0,1,1)(0,1,1)12
Ortho-Interlake	ARIMA(0,1,1)(0,1,1)12
Ortho-Northern	ARIMA(0,0,1)(0,1,1)12
Hospital	
Ortho_Overall	ARIMA(1,0,1)(1,1,1)12
Ortho_Winnipeg	ARIMA(0,1,1)(0,1,1)12
Ortho-West Eastman	ARIMA(1,0,1)(1,1,1)12
Ortho-Southern	ARIMA(0,1,1)(0,1,1)12
Ortho-Interlake	ARIMA(0,1,1)(0,1,1)12
Ortho-Northern	ARIMA(0,0,1)(0,1,1)12
Clinic	
Ortho_Overall	ARIMA(1,1,1)(1,1,1)12
Ortho_Winnipeg	ARIMA(0,1,1)(0,1,1)12
Ortho-West Eastman	ARIMA(1,0,1)(1,1,1)12
Ortho-Southern	ARIMA(0,1,1)(0,1,1)12
Ortho-Interlake	ARIMA(0,1,1)(0,1,1)12
Ortho-Northern	ARIMA(0,0,1)(0,1,1)12

Ophthalomology Procedures	
Outcome	**SARIMA Models**
Ophthal_Overall	ARIMA(1,1,1)(1,1,1)$_{12}$
Ophthal_Winnipeg	ARIMA(3,1,1)(1,1,0)$_{12}$
Ophthal-West Eastman	ARIMA(3,1,1)(1,1,0)$_{12}$
Ophthal-Southern	ARIMA(3,1,1)(1,1,0)$_{12}$
Ophthal-Interlake	ARIMA(0,1,1)(0,1,0)$_{12}$
Ophthal-Northern	ARIMA(1,0,1)(0,1,0)$_{12}$
Hospital	
Ophthal_Overall	ARIMA(1,0,1)(1,1,1)$_{12}$
Ophthal_Winnipeg	ARIMA(3,1,1)(1,1,0)$_{12}$
Ophthal-West Eastman	ARIMA(3,1,1)(1,1,0)$_{12}$
Ophthal-Southern	ARIMA(3,1,1)(1,1,0)$_{12}$
Ophthal-Interlake	ARIMA(0,1,1)(0,1,0)$_{12}$
Ophthal-Northern	ARIMA(1,0,1)(0,1,0)$_{12}$
Clinic	
Ophthal_Overall	ARIMA(1,0,1)(1,1,1)$_{12}$
Ophthal_Winnipeg	ARIMA(3,1,1)(1,1,0)$_{12}$
Ophthal-West Eastman	ARIMA(3,1,1)(1,1,0)$_{12}$
Ophthal-Southern	ARIMA(3,1,1)(3,1,0)$_{12}$
Ophthal-Interlake	ARIMA(0,1,1)(0,1,0)$_{12}$
Ophthal-Northern	ARIMA(1,0,1)(0,1,0)$_{12}$

University of Manitoba | Rady Faculty of Health Sciences

Manitoba Centre for Health Policy
Data | Insight | Informing Solutions

University of Manitoba
Max Rady College of Medicine
Rady Faculty of Health Sciences

408-727 McDermot Avenue
Winnipeg, Manitoba, Canada
R3E 3P5

Tel: (204) 789-3819
Fax: (204) 789-3910
Email: reports@cpe.umanitoba.ca

www.mchp.ca