IOSR Journal of Dental and Medical Sciences (IOSR-JDMS)
e-ISSN: 2279-0853, p-ISSN: 2279-0861.Volume 14, Issue 1 Ver. IV (Jan. 2015), PP 17-21
www.iosrjournals.org

OPD Attendance of Under-5 Children at an Urban Health Training Centre in West Bengal, India: A Time Series Analysis

Dr. Atanu Biswas[1], Dr. Sutanuka Santra[1], Dr. Surajit Lahiri[1],
Prof. Dr.SaibenduKumar Lahiri[2].

[1](3rd year postgraduate trainee, Department of Community Medicine, R. G. Kar Medical College, Kolkata, West Bengal, India)

[2](Professor and Head, Department of Community Medicine, R. G. Kar Medical College, Kolkata, West Bengal, India)

Abstract: Background:*Occurrence of diseases in children varies every month, so does their OPD attendance. This can be attributed to the seasonal variation of certain diseases. Time series analysis can identify the correlation among the numbers of OPD attendance over a period and can also forecast OPD attendance month wise.*

Objectives:i)*To build a time series model of OPD attendance by under 5 children at the Urban Health Training Centre (UTC) of R. G. Kar Medical College and Hospital, Kolkata.* ***ii)*** *To forecast the OPD attendance at the UTC.*

Methodology:*A time series analysis was conducted using 6 years record (from April '07 to March '13) on OPD attendance by under 5 children at the Bagbazar UTC of R. G. Kar Medical College and Hospital, Kolkata. Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF) were used for setting an approximate model of OPD attendance pattern. Multiplicative model was used for forecasting. Statistical analysis was done using SPSS v.16 and Microsoft Excel 2010.*

Results: *The ACF and PACF correlograms suggest an autoregressive (AR) pattern of 2 and moving average (MA) pattern of 2 , without any differencing ; hence ARIMA (2,0,2) model. Forecasting of OPD attendance was estimated using multiplicative model, $Y_t = S_t \times T_t$.*

Conclusion: *OPD attendance is just an approximate reflection of disease incidence, since not all the cases would visit the OPD because of various reasons. Still, a time series analysis can give a fair idea about the probable OPD attendance in future, and hence, can be useful for resource allocation.*

Keywords: *OPD attendance, Under 5 children, Time series analysis, Forecasting.*

I. Introduction

OPD attendance is a fair reflection of incidence and prevalence of disease in the community. Occurrence of diseases varies every month, which may be reflected by varying OPD attendance every month. As incidence of any disease is uncertain, it is also difficult to predict the number of patients attending OPD in future. There can be sudden rise in incidence of a particular disease, which may be due to many factors (e.g. seasonality), and as a result it is also expected to have increased number of patients attending OPD with that particular disease. An outbreak or epidemic will also result in abrupt increase in OPD attendance.

Children, up to the age of 5 years, are vulnerable to different diseases due to inadequately built immunity. Different national programmes have focused on preventive measures or early diagnosis and treatment for under-5 children, so as to decrease morbidity and mortality in this age group. OPD attendance can also be an indicator for functional evaluation of the health care delivery facility.

Time series means a series of observations recorded in accordance with the time. The variables can be observed by days, weeks, months, quarters or years. The changes of observations over time can be secular, seasonal, cyclical or irregular (random).[1] For example, malarial cases are expected to follow seasonality, while a random or irregular rise is expected in case of an outbreak of food poisoning. Time series analysis of OPD attendance can reveal the trend of OPD attendance, and can also predict the estimated OPD attendance in future. This can particularly be helpful as an estimated OPD attendance can give an idea about the required resource and logistic can be managed accordingly.

II. Materials And Methods

A time series analysis on OPD attendance among the under 5 children was conducted at Baghbazar Urban Health Training Centre (UTC), Kolkata. This UTC is the urban field practice area of the Department of Community Medicine, R. G. Kar Medical College, Kolkata, West Bengal, India.

The time seriesanalysis was done using the record of previous 6 years. The monthly OPD attendance, of the under 5 children, at Baghbazar UTC, were noted for last 72 months from the record.

Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF) were used for setting an approximate model of OPD attendance pattern. In a time series data, ACF describes the correlation of an observation with its previous/subsequent n^{th} observation. PACF describes the correlation of an observation with its previous/subsequent n^{th} observation, having the intermediate autocorrelations partialed out. Using the ACF and the PACF, an approximate autoregressive integrated moving average or ARIMA model can be set for a pattern of OPD attendance at Baghbazar UTC. An ARIMA model is expressed as (p,d,q); 'p' being the autoregressive component, 'q' being the moving average component and 'd' being the differencing component. Autoregression means the number of terms in the model that describes the dependency among successive observations. Moving average means the number of terms that describes the persistence of a random shock from one observation to the next. Differencing is calculating differences among pairs of observations at some lag to make a non-stationary series stationary.[2]Box-Ljung statistics was used for verification of model fitness.[3]Multiplicative model was used for forecasting. According to the multiplicative model, the observation at time 't' is expressed as,$\mathbf{Y_t = S_t \times T_t \times I_t}$ (S= seasonal, T = trend, I= irregular).[4]Statistical analysis was done using SPSS v.16 and Microsoft Excel 2010.

III. Results

When the monthly observations, as collected from record review at the UTC, were plotted against time i.e. the OPD attendance shown along vertical axis and the months and the years shown along horizontal axis, the time series data didn't show any trend. There was no regularity found in thepattern of the peaks and trough [Figure 1].

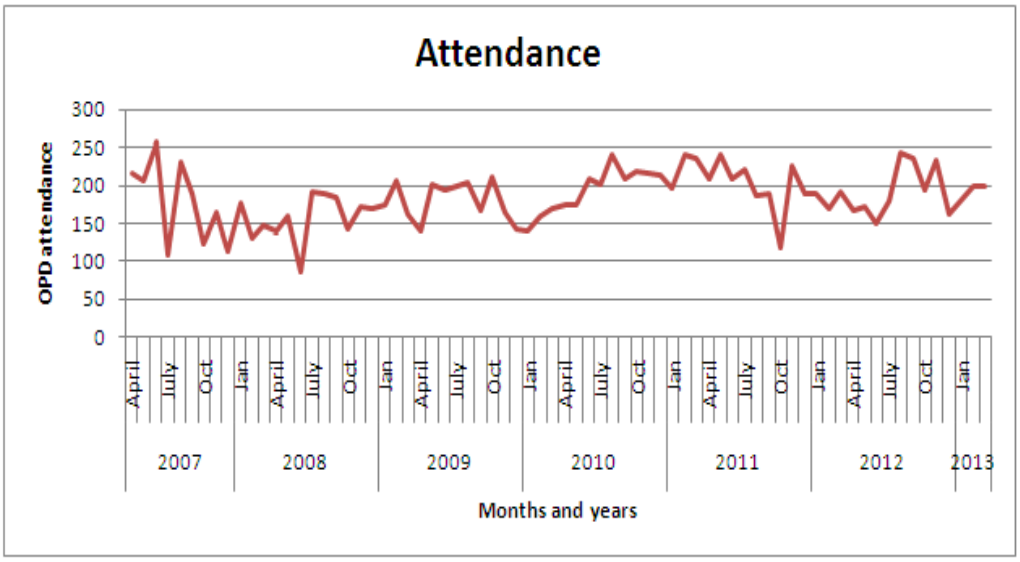


Figure 1: Time series plot of under 5 children attended the OPD.

In Figure 2 the correlogram for autocorrelation function (ACF) shows the autoregressive component of 2 (Autocorrelation at lag 2= 0.375 ; Box-Ljung statistics p =0.001) and it confirms no seasonality of the time series data. Hence, in our study, we accepted a 'p' value of 2.

Similarly, in Figure 3 the correlogram for autocorrelation function (ACF) shows the moving average component of 2 (Autocorrelation at lag 2= 0.337 ; Box-Ljungstatistics p =0.001)and confirms no seasonality of the time series data. Therefore, we have accepted 'q' to be 2 in this study.

Since, for the purpose of making the time series data stationary, differencing was not used, the 'd' value in our study is 0.

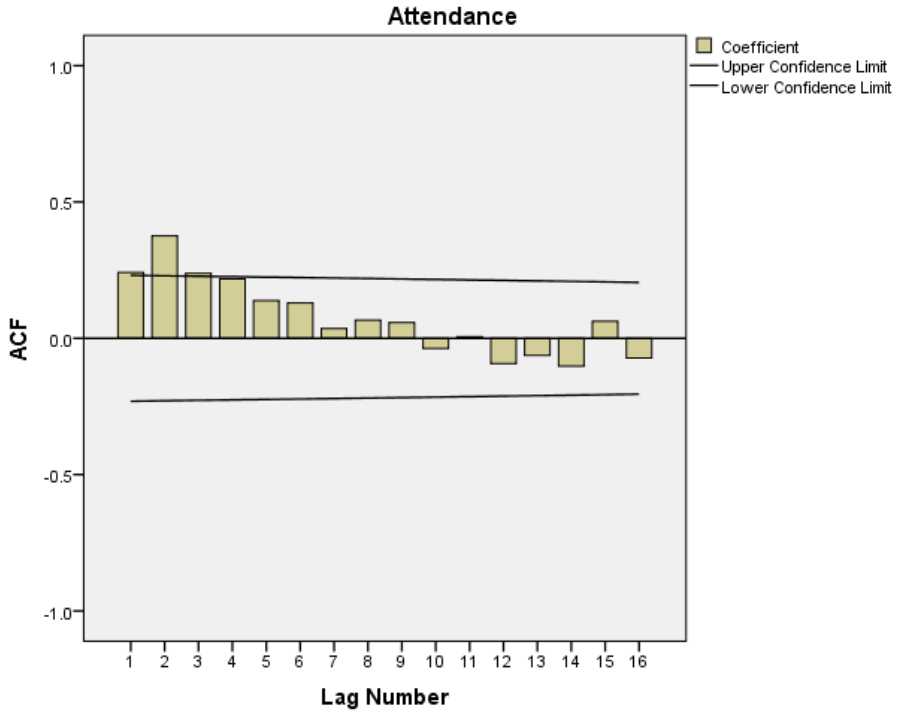


Figure 2: Autocorrelation plot for OPD attendance of under 5 children (Autocorrelation at lag 2= 0.375 ; Box-Ljung statistics p =0.001)

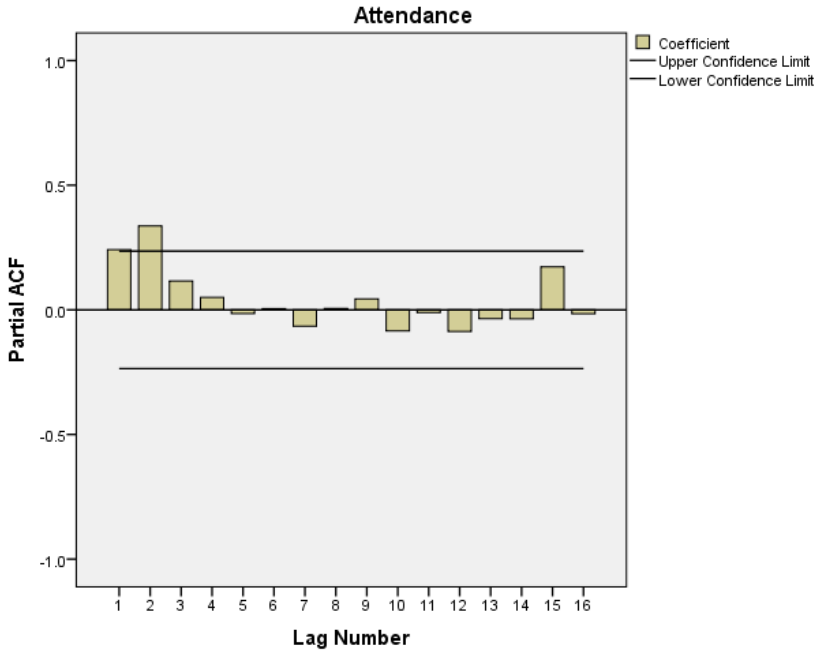


Figure 3: Partial Autocorrelation plot for OPD attendance of under 5 children (Partial autocorrelation at lag 2 = 0.337)

From the correlogram, we have predicted 'p' and 'q' as 2, and since no differencing was used, the predicted value of 'd' was 0. Therefore, the predicted time series model for this study was ARIMA(2,0,2). The ARIMA model was used on the original time series and was found to be the closest among all other options corresponding to the observations in the time series data. Figure 4 shows the observed time series in red line and the predicted model in blue line. Any deviation from the observed value can be attributed to the irregular pattern in the time series data.

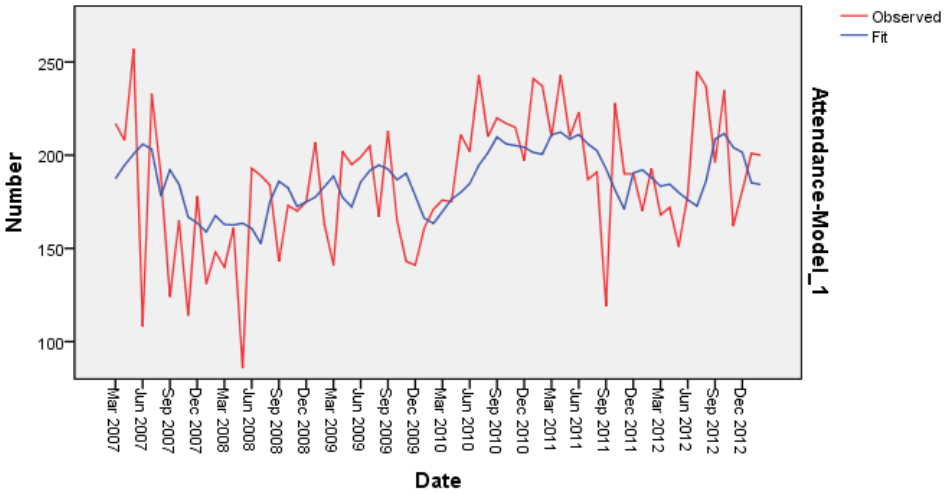


Figure 4: Predicted time series plot using ARIMA (2,0,2) model.

The Ljung-Box statistics, also known as modified Box-Pierce statistic, is an indicator about the fitness of the model. A significant value of less than 0.05 means there are structure in the observed series which are not accounted by the predicted model.[4]Hence, as shown in table 1, the p value of 0.978 confirms the model fitness as there is no structure in the observed series that can't be explained by the model.

Table 1: Model statistics for under 5 children OPD attendance.

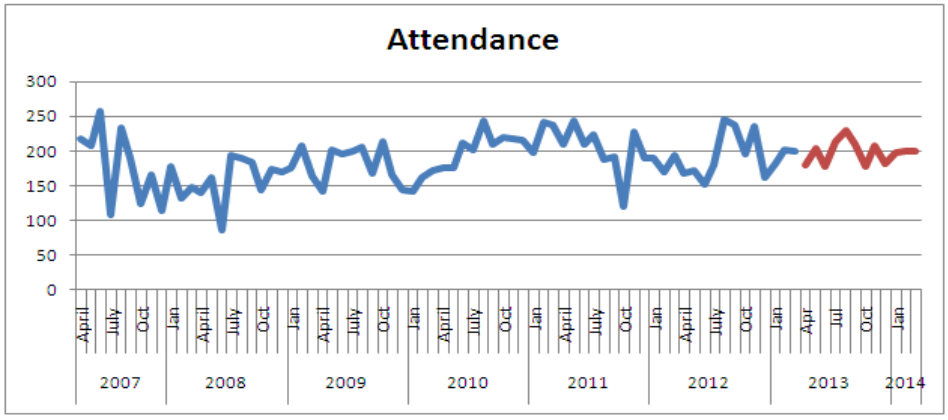
Model	Model Fit statistics	Ljung-Box Q(18)		
	Stationary R-squared	Statistics	DF	Sig.
Attendance-Model_1	0.182	5.500	14	.978

For the purpose of forecasting OPD attendance in future months, we used the multiplicative model of time series which is $Y_t = S_t \times T_t \times I_t$ (S= seasonal, T = trend, I= irregular). But since the irregular pattern can't be predicted and can vary over time, a simplified multiplicative model, $Y_t = S_t \times T_t$, was used. Moving average (MA) and centered moving average (CMA) was estimated for predicting the seasonal pattern.[1]The forecast was predicted for one year, i.e. April 2013 to March 2014. The predicted OPD attendances for next twelve months have been shown in Table 2. Figure 5 shows the predicted forecast in red line along with the observed time series data as shown in blue line.

Table 2: Estimated forecast of under 5 children OPD attendance for next 12 months.

Year	2013									2014		
Months	April	May	June	July	Aug	Sept	Oct	Nov	Dec	Jan	Feb	Mar
Predicted OPD attendance	179	203	178	212	229	209	178	206	183	197	200	200

Figure 5: Forecasting of under-5 OPD cases month wise from April 2013 to March 2014.



Blue line showing the observed trend. Red line showing predicted forecast.

IV. Discussion

The observed time series data shows there was varying number of under five children attending to Bagbazar UTC, Kolkata, over the period of 5 years. The observations didn't show any seasonality or trend due to their irregularity in pattern. This may apparently mean there is no seasonality in the disease occurrence as well as the OPD attendance, which is actually a misinterpretation. Since the OPD attendance for all the diseases taken together, instead individual diseases were considered for the study, seasonality in the OPD attendance on under five children was very unlikely. This is one of the limitations of the study and there remains a scope of further study on OPD attendance for individual disease to predict seasonality, if any. The purpose of the study was to predict a model that corresponds to the time series data as observed, and also to forecast the approximate OPD attendance for a period of time after the study was conducted. Our study predicted ARIMA(2,0,2) model which was statistically proven fit and could exclude the presence of any structure in the observed series that couldn't be explained by the model. The model graphically fits to the observed time series data, and the deviations can be attributed to the irregular trend of the time series, that can't be predicted. One of the major limitations of any time series analysis is that it can not predict, or formulate into model, any irregularity in the trend as these are unexplained deviation for the expected values. Same goes true in case forecasting a time series data. The existing model can only be used to read and analyze the current pattern and thus estimate an approximate series of data for a period of time in future. As any irregular factor can notbe estimated beforehand, the forecasted data is only an estimation that is liable to vary depending on the irregularity.[6] In our study, we have made prediction on the under five OPD attendance for twelve months after the study was conducted. This approximate estimation can be utilized in resource allocation, in terms of doctor, nurse, medicine requirement etc. and it will also help in prevention of valuable resource wastage.

V. Conclusion

Our study has successfully predicted a model that fits the observed time series data and also forecasted an approximate OPD attendance by under five children. The children being the future of the nation, should be given priority and their health care should be focused on. India, like any other developing country, must reinforce the available resourcefor providing health care service to control under five morbidity and mortality. At sub-centre, PHC and block level, where health care delivery is mostly hampered due to inadequate manpower, inappropriate infrastructure or improper inventory management, a proper forecasting can give a fair idea about the approximate number of patients to attend the OPD, and hence the required resource can be estimated and logistic can be managed accordingly, so as to deliver uninterrupted health care to those who need it, without wasting any valuable resource.

References

[1]. Das N G. Statistical Methods. 3rd ed. Delhi: Tata McGraw Hill; 2010.Chapter 6, Time Series; p. 290-343.
[2]. Time-Series Analysis.[Internet] [cited 2014 July 8]. Available from: https://datajobs.com/data-science-repo/Time-Series-Analysis-Guide.pdf
[3]. Chowdhury R, Mukherjee A, Naska S, Adhikary M, Lahiri SK. Seasonality of tuberculosis in rural West Bengal: A time series analysis. Int J Health Allied Sci 2013; 2: 95-8.
[4]. Time Series Analysis: The basics. [Internet] [cited 2014 July 22]. Available from: http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Time+Series+Analysis:+The+Basics
[5]. Ljung-Box Statistics for ARIMA Residuald in R. [Internet] [cited on 2014 June 17]. Available from: http://stats.stackexchange.com/questions/64711/ljung-box-statistics-for-arima-residuals-in-r-confusing-test-results
[6]. Limitations of Time Series Analysis. [Internet] [cited on 2014 July 8]. Available from: http://node101.psych.cornell.edu/Darlington/series/series1.htm.