RESEARCH ARTICLE

THE COMPARISON BETWEEN ARTIFICIAL NEURAL NETWORKING AND BOX-JENKINS METHODOLOGY FOR FORECASTING OF MORTALITY RATE.

[*]**Asad Ali and Muhammad Zubair.**
Department of Statistics, University of Sargodha, Sargodha, Pakistan.

....

Manuscript Info	*Abstract*
	It is impossible to ignore the demographic changes due to their impact on development of the country as well as effects on all areas of human activities. Forecasting demographic variables help us to control the situation in time. The demographic trends also support in policy making. Mortality rate is an important indicator to evaluate the provision of health facilities by the government. The techniques used for forecasting are also important in this regard. So, we are going to make comparison between two techniques (Box-Jenkins and Artificial Neural Networking (ANN)) which are extensively used for forecasting. For this purpose, we utilize the data related to mortality rate of Pakistan from 1975-2014. Two models (Autoregressive Integrated Moving Average (ARIMA) (15,2,1), and ARIMA (1,2,2)) are selected using Schwarz Criterion (SC) for Box-Jenkins methodology and compared with ANN based on Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The results depict that ANN is more appropriate technique for forecasting as compare to Box-Jenkins.

....

Introduction:-

The mortality rate is playing important role in public health administration to assess the public health reforms, projects and programs. It also provides road map for insurance companies, national planning, epidemiology and medical research. Garenne et al. (1992) found the trend of mortality under 5 years in developing countries due to acute respiratory infection (ARI). They found that after 1950 the trend of mortality was smaller due to ARI in European population and developing countries but there was no real data available to draw exact pattern. Zakir and Wunnava (1999) worked on factors affecting infant mortality rates and found results from correctional data of 117 countries for 1993 years. A regression model developed with exploratory variables fertility rates, female participation in labour force, female literacy rates, per capita gross net production and expenditure of government on health care. They found that fertility rates, female participation in the labour force, female literacy rates and per capita gross net production significantly affect the infant mortality rates but government expenditure on health care does not had significant role.

Adetunji (2000) investigated the relationship of mortality due to HIV under 5 and its symptoms in adults. The result showed that there was direct relationship between both. Black et al. (2003) described that most deaths of children were in developing countries, from which 50% of all deaths were coming from 6 countries and 42 countries claims for almost 90%. They suggested to fight against mortality under 5 years at country level instead of geopolitical

---

Corresponding Author:-Asad Ali.
Address:-Department of Statistics, University of Sargodha, Sargodha, Pakistan.

areas. Rajaratnam et al. (2010) described that there were 11.9 million deaths under 5 years in 1990 and reduces to 7.7 million in 2010. They concluded that the developmental process was needed to be continued to control the mortality rate and reduce it more. Wang et al. (2015) stated that china decreased the mortality rate with 8.8% by applying good strategies about socio-economic factors. Ali et al. (2016) used Box-Jenkins methodology for forecasting and comparing the mortality rate under five years between Pakistan, Bangladesh and India for the period of 1975-2015. Kipp et al. (2016) showed that development in governance, health factors and socio-demographic factors as supporting tool to decrease the mortality rate under 5 in 46 African countries from 2000-2013.

After the comprehensive study of several research papers, we see that area of forecasting mortality rate has few studies and appropriate forecasting technique is desired for it. So, in this study we are going to compare two forecasting techniques Box-Jenkins and Artificial Neural Networking (ANN).

Material and Methodology:-

For forecasting purpose, secondary data is collected about mortality rate (per 1000) of Pakistan from the site of World Bank for the period of 1975-2014. In Box-Jenkins technique, augmented dickey fuller (ADF) test is utilize to check stationarity of the data. By drawing correlogram of stationary time series, possible orders of autoregressive integrated moving average (ARIMA) model are selected. From all possible orders, two appropriate orders are selected based on schwarz criterion (SC). Both ARIMA models have utilized to forecast mortality rates, and their forecasting errors (Mean Absolute Error (MAE) and Root Mean Square Error (RMSE)) are calculated. In ANN technique, 12 input and hidden layers are selected with 1 output layer of neurons. Bipolar Sigmoid function has utilized as an activation function for forecasting mortality rate. Forecasting errors MAE as well as RMSE have also calculated for comparison. Residual analysis for both techniques have also been carried out.

Results and Discussion:-
Box-Jenkins Technique:-
Table 1:- Augmented Dickey-Fuller Test

	ADF-statistic	Critical Value		
		1%	5%	10%
Level	-1.1311	-3.6117	-2.9399	-2.6080
1st diff	-1.4508	-3.6171	-2.9422	-2.6092
2nd diff	-2.6857	-3.6228	-2.9446	-2.6105*

*stationary

In Table 1 ADF test shows that mortality rate of Pakistan is stationary series at 2nd difference with 10% level of significance. So, we draw correlogram of mortality rate at 2nd difference in Figure 1
.
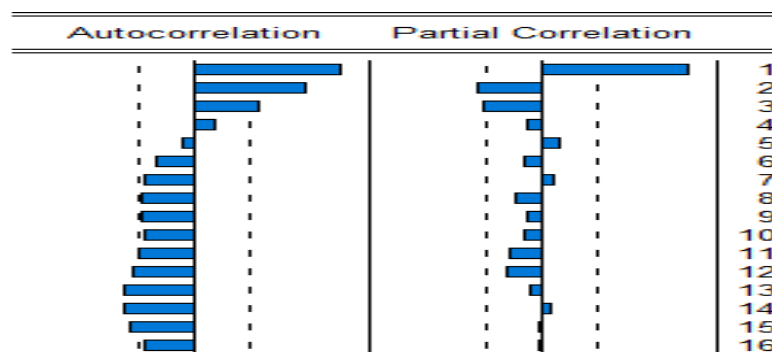
Fig. 1:-Correlogram of Pakistan's mortality rate at 2nd difference

With the help of this correlogram, possible ARIMA models with their SC are obtained in Table 2.

Table 2:- All Possible models and their SC value

| Model	SC	Model	SC
0,2,0	-6.336405	3,2,2	-7.279289
1,2,0	-7.984004	12,2,2	-7.181191
0,2,1	-7.518724	13,2,2	-7.394418
1,2,1	-8.431708	14,2,2	-7.738162
2,2,1	-8.029926	15,2,2	-7.905808
3,2,1	-7.539235	1,2,3	-8.238624
12,2,1	-7.523323	2,2,3	-7.233540
13,2,1	-7.753442	3,2,3	-6.902611
14,2,1	-8.017225	12,2,3	-7.084326
15,2,1	-8.527978*	13,2,3	-7.449583
1,2,2	-8.489277**	14,2,3	-7.510704
2,2,2	-7.606180	15,2,3	-7.640836

*,** lowest and second lowest respectively

Both possible models are used to forecast mortality rate for next six years (2015 to 2020) and their result are as follows:

Table 3:-Forecast Results of ARIMA (1,2,2) and ARIMA (15,2,1)

year	ARIMA (1,2,2)	ARIMA (15,2,1)
2015	10.5934269245	7.17487311073
2016	10.6045745465	7.12240022546
2017	10.6191425181	7.07275513657
2018	10.637128126	7.02593784406
2019	10.6585290851	6.98210750235
2020	10.6833434714	6.94126411145

Residual Analysis:-

Correlogram of squared residuals is utilized to check that residuals of these models are uncorrelated and independent from each other or not. Figure 2 and 3 show the correlogram of square residuals for ARIMA (1,2,2) and ARIMA (15,2,1) respectively.
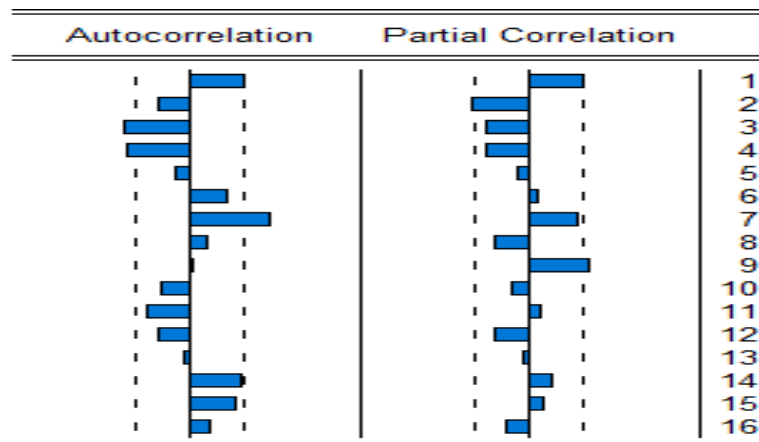
Fig. 2:- Correlogram of Squared residuals ARIMA (1,2,2)
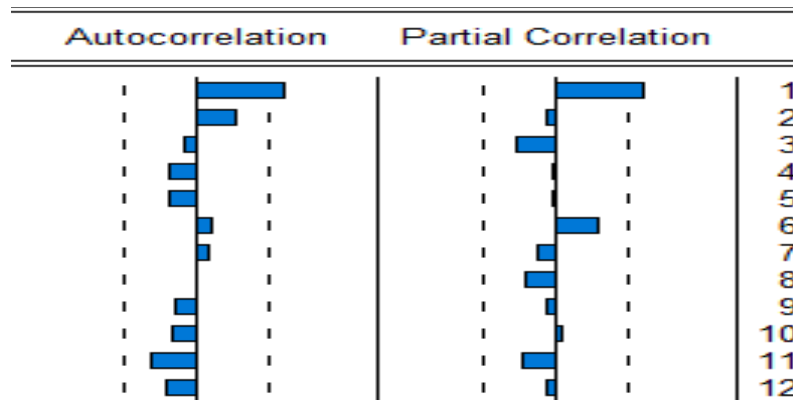
Fig. 3:-Correlogram of Squared residuals ARIMA (15,2,1)

Both figures show that most of the spikes of squared residuals are lies within the marginal lines. So, we can conclude that residuals of both models are uncorrelated and independent from each other and satisfies the assumption of Box-Jenkins technique.

Artificial Neural Networking (ANN):-
In ANN, 12 input, 12 hidden and 1 output layer is selected appropriate with bipolar sigmoid function as an activation function. Forecasted results of ANN are as follows:

**Table 4:-**Forecast Results of ANN

year	ANN
2015	7.5672
2016	7.5539
2017	7.5381
2018	7.5233
2019	7.5129
2020	7.5036

In Table 4, forecasted results of ANN are shown. It shows that at the end of 2020, mortality rate will have decreased and reached at 7.5036 per 1000 persons.
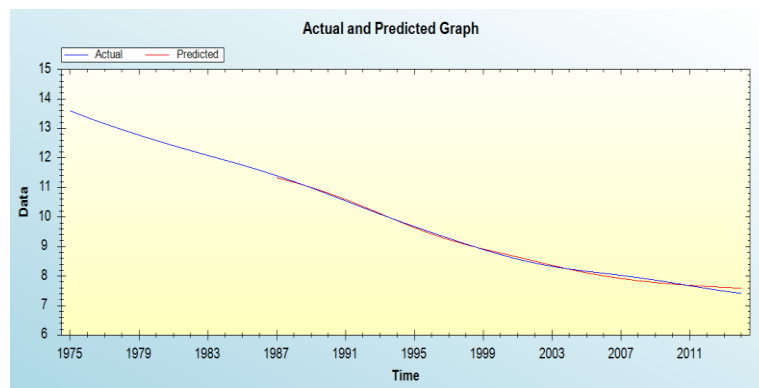

Fig. 4:-Actual and Predicted Graph.

Figure 4 represents the graph of actual mortality rate and predicted mortality rate for the period of 1975 to 2014. Graph shows that ANN predict mortality rate with great accuracy as the lines of both series are overlapped with each other.
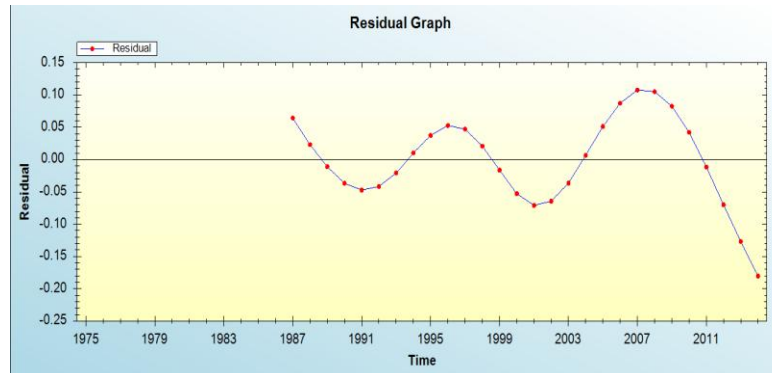
Fig. 5:-Residuals Graph

Figure 5 shows that residuals of ANN are white noise because residuals have zero mean and variance interval 0.15 to -0.20. Residuals also shows the cyclic pattern as we checked in Figure 2.
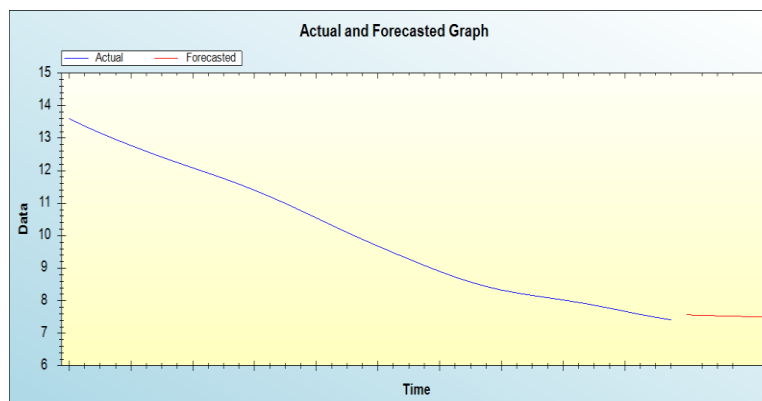
Fig. 6:-Actual and Forecasted Graph.

For checking forecasting accuracy, we draw graph between actual morality rate (1975 to 2014) and forecasted mortality rate (2015 to 2020) in Figure 6. Graph shows that forecasted mortality rate extends the same pattern as seen in actual mortality rate.

For accuracy comparison of both ARIMA models and ANN we measure MAE and RMSE in Table 5.
Table 5:-Forecasting Checks on ARIMA and ANN

Model	MAE	RMSE
ANN	0.054475*	0.0671*
ARIMA (15,2,1)	0.1232	0.1615
ARIMA (1,2,2)	1.4957	1.8496

*smallest value

As Table 5 shows that results of MAE and RMSE for ANN are smallest from both ARIMA models. So, it leads to a better performance of ANN technique over Box-Jenkins technique for mortality rate of Pakistan.

Conclusion:-

The study is carried out for comparison between two forecasting techniques (Box-Jenkins and ANN) based on mortality rate (per 1000) of Pakistan. For this purpose, data of mortality rate is collected from site of World Bank for the period of 1975-2014. By utilizing both techniques, mortality rate of next six years (2015-2020) have been forecasted. Results of MAE and RMSE depict that ANN technique has less forecasting error than Box-Jenkins. So, we recommend to use ANN for forecasting mortality rate in future studies.

References:-

1. Adetunji, J. (2000). Trends in under-5 mortality rates and the HIV/AIDS epidemic. *World Health Organization,* 78(10), 1200-1206.
2. Ali, A., Qamar, S., Zubair, M., Akhtar, N., and Iqbal, M. (2016). A Comparative Study Of Pakistan, Bangladesh And India On The Basis Of Mortality Rate Under Five Years. *Journal of ISOSS,* 2(1), 79-84.
3. Black, R.E., Morris, S.S. and Bryce, J. (2003). Where and why are 10 million children dying every year? *The Lancet,* 361, 2226-2234.
4. Garenne, M., Ronsmans, C. and Campbell, H. (1992). The Magnitude of mortality from acute respiratory infections in childern under 5 years in developing countries. *World Health Statistics Quarterly*, 45, 180-191.
5. Kipp, A.M., Blevins, M., Haley, C.A., Mwinga, K., Habimana, P., Shepherd, B.E., Aliyu, M.H., Ketsela, T. and Vermund, S.H. (2016). Factors associated with declining under-five mortality rates from 2000 to 2013: an ecological analysis of 46 African countries. *BMJ Open,* 6, 1-11.
6. Rajaratnam, J.K., Marcus, J.R., Haxman, A.D., Wang, H., Rector, A.L., Costa, M., Lopez, A.D. and Murray, C.J.L. (2010). Neonatal, postneonatal, childood, and under-5 mortality for 187 countries, 1970-2010: a systematic analysis of progress towards Millennium Development Goal 4. *The Lancet,* 375, 1988-2008.
7. Wang, Y., Li, X., Zhou, M., Luo, S., Liang, J., Liddell, C.A., Coates, M.M., Gao, Y., Wang, L., He, C., Kang, C., Liu, S., Dai, L., Schumacher, A.E., Fraser, M.S., Wolock, T.M., Pain, A., Levitz, C.E., Singh, L., Coggeshall, M., Lind, M., Li, Y., Li, Q., Deng, K., Mu, Y., Deng, C., Yi, L., Liu, Z., Ma, X., Li, H., Mu, D., Zhut, J., Murrayt, C.J.L. and Wang, H. (2015). Under-5 mortality in 2851 Chinese countries, 1996-2012: a subnational assessment of achieving MDG 4 goals in China. *The Lancet*, 1-11.
8. Zakir, M. and Wunnava, P. V. (1999). Factors affecting infant mortality rates: evidence from cross-sectional data. *Applied Economics Letters*, 6, 271–273.