International Journal of Scientific Research in Mathematical and Statistical Sciences **Research Paper**
Vol.6, Issue.2, pp.50-55, April (2019) **E-ISSN:** 2348-4519
DOI: https://doi.org/10.26438/ijsrmss/v6i2.5055

Stochastic Time Series Modeling for Rice Production in Andhra Pradesh

S. C. Thasleema[1], Ahammad Basha Shaik[2*]

[1]Department of Statistics, Vikrama Simhapuri University, Nellore, Andhra Pradesh, India
[2]Department of Community Medicine, Narayana Medical College, Nellore, Andhra Pradesh, India

***Corresponding Author:** *ahammadbasha@gmail.com*, Tel.: +91-9985541610

Available online at: www.isroset.org

Received: 06/Mar/2019, Accepted: 08/Apr/2019, Online: 30/Apr/2019

Abstract- Rice is the chief food grain of Indian population. Andhra Pradesh ranks fifth in the rice production in the country. Due to increase in demand of rice over the years, the modeling and forecasting of rice production over the years is very important. An Auto-Regressive Integrated Moving Average (ARIMA) methodology has been successful in describing and forecasting rice production in past studies. In the present study, ARIMA stochastic modeling is used for describing rice production in Andhra Pradesh. The yearly rice production data of Andhra Pradesh from 1980-81 to 2016-2017 has been used for validation of the model. The Statistical software SAS 9.3 is used for analyzing the data. The ARIMA (0, 1, 1) model is selected as the best model based on the minimum values of Root Mean Square Error (RMSE), Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for the rice production of data in Andhra Pradesh. The forecasting of rice production was estimated from 2015-16 to 2026-27 using ARIMA approach.

Keywords- *Rice production, ARIMA, Andhra Pradesh, SAS, Stochastic Time series*

I. INTRODUCTION

Rice is the major food crop and it has the largest cultivation area among food grains in India. India secures the second rank in the world for the production of rice. Rice production of India reaches 104408.2 thousand tons in the year 2015-16. The top five states of rice production in India are West Bengal (15953.9 thousand tons), Uttar Pradesh (12501.0 thousand tons), Punjab (11823.0 thousand tons), Tamil Nadu (7517.1 thousand tons) and Andhra Pradesh (7488.7 thousand tons) in the year 2015-16. Andhra Pradesh state is also known as "*The Rice Bowl of India*" for rice production and it is grown in both Kharif and Rabi seasons in all the districts of Andhra Pradesh. East Godavari district is identified as *"Rice Bowl of Andhra Pradesh"* [1].

A time series is a sequence of observations taken sequentially in time. Time series analysis is concerned with techniques for the analysis of this dependence [2]. This requires the development of stochastic and dynamic models for time series data and the use of such models in important areas of application (Box *et al.*, 2016) [2]. In a stochastic time series models, the popular methodology is Autoregressive Integrated Moving Average (ARIMA) models and is very powerful as they can successfully describe the raw data and forecasting the production of different food grains with minimum forecast error [2].

Several attempts have been made in the past to examine the trends and to develop ARIMA models to forecast the area, production, and productivity of Rice crop. The best fitted ARIMA model that could be used to make efficient forecast boro rice production in Bangladesh from 2008-09 to 2012-13 was examined by Rahman (2010) [3]. The growth pattern and also examine the best ARIMA model to efficiently forecasting Aus, Aman and Boro rice production in Bangladesh was studied by Awal and Siddique (2011) [4]. Badmus and Ariyo (2011) have focused on forecasting the cultivated area and production of maize in Nigeria using Autoregressive Integrated Moving Average (ARIMA) model from 1970-71 to 2005-06 [5]. Zakari and Ying (2012) have studied the Niger grain production and harvested Areas and for the two main staple crops (Millet and Sorghum) using the ARIMA model [6]. Venkatramana *et al.* (2012) studied ARMA models for the selected best model for the rainfall data in India [7]. Jambhulkar (2013 *a*, *b*) has used ARIMA stochastic modeling for studying rice production in West Bengal and Punjab [8,9]. Biswas and Bhattacharyya (2013) have studied the ARIMA modeling to forecast the area and production of rice in West Bengal [10]. Prabakaran and Sivapragasam (2014) have used the ARIMA model for forecasting areas and production of rice in India for the period 1950-51 to 2011-12 [11].

The study attempts to examine the production scenario, growth, sustainability and forecast the area, production and yield of wheat in major growing states of India using ARIMA model, taking into consideration the factors like rainfall, fertilizer was given by Mishra *et al.* (2015) [13]. Srinivasulu and Sarojamma (2018) have fitted the Best Model for Acute Coronary Events (ACEs) Using Interrupted Time Series Data [14].

The main objective of the study is to estimate and forecasts the rice production of Andhra Pradesh using univariate time series modeling techniques.

II. MATERIALS AND METHODS

The time series data of rice production in Andhra Pradesh has been collected from the website of Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Ministry of Agriculture from 1980-81 to 2016-2017. ARIMA stochastic modeling is used on the rice production of Andhra Pradesh for forecasting purpose. SAS Version 9.3 was used to fit the ARIMA models and estimating the forecast values. All the P values are having less than 0.05 are considered as statistical significant.

ARIMA modeling was developed the standard three steps procedures. Namely, Identification of the model, estimation of parameter and diagnostic checking and verification of the model. The identification step determines the process is stationary or not and the form of the process is Autoregressive (AR), Moving Average (MA) or both (ARMA), and its orders. Three parameters are used in summarizing an ARIMA model are the AR parameter *p*, integration parameter *d*, and MA parameter *q*. Parameters *p* and *q* denote the order of AR and MA, while *d* denotes the degree of differencing.

The development of ARIMA models was given by Box and Jenkins in the year 1970. Let us consider the *autoregressive moving average (ARMA)* model, denoted by ARMA (*p,q)*, is given by

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots\ldots + \alpha_p X_{t-p} + Z_t + \beta_1 Z_{t-1} + \ldots.. + \beta_q Z_{t-q} \qquad (1)$$

Using the backward shift operator *B*, Equation (1) can be written as

$$\varphi(B)\, X_t = \Phi(B)\, \nabla^d X_t = \theta_0 + \theta(B) Z_t \qquad (2)$$

Where $\Phi(B)$ and $\theta(B)$ are polynomial operators of *B*, of degree *p* and *q* respectively.

Such that

$$\Phi(B) = 1 - \alpha_1 B - \alpha_1 B^2 - \ldots\ldots - \alpha_p B^p \quad \text{and} \quad \theta(B) = 1 - \beta_1 B - \beta_2 B^2 - \ldots\ldots. - \beta_q B^q$$

Where $\Phi(B)$ is called the *autoregressive operator*, $\varphi(B)\, X_t = \Phi(B)\, \nabla^d$ is called the *generalized autoregressive operator* and $\theta(B)$ is called the *moving average operator*.

Therefore, the process of equation (2) is a combination of AR and MA models with suitable order of differencing then this process is referred as *autoregressive integrated moving average (ARIMA)* process and is denoted as ARIMA (*p*, *d*, *q*) model. The integration parameter *d* is a non-negative integer. The non-stationary time series model is reduced to a stationary model by selecting the appropriate order of differencing. If $d = 0$, then the ARIMA (*p*, *d*, *q*) model is reduced to ARMA (*p*, *q*) model (Box *et al.*, 2016) [2].

The final results included the parameter estimates, standard error, Root Mean Square Error (RMSE), Akaike's Information Criterion (AIC), Schwartz's Bayesian Criterion (SBC) or Bayesian Information Criterion (BIC) and R^2 value. The performance of model selection was based on the minimum (or) least values of AIC and BIC which determines the parameter estimation of tentative models. The statistical analysis has been done by using SAS Version 9.3. All the P values are having less than 0.05 are considered as statistically significant.

III. RESULTS AND DISCUSSION

The time series plot for rice production of Andhra Pradesh from 1980 to 2014 fluctuates from 7011.4 thousand tones and increases up to 11565.4 thousand tones. From the Plots of ACF and PACF (Figure 1), we observed that the data are non-stationary. After one differentiating, the original series has become stationary with the help of Augmented Dickey-Fuller (ADF) unit root test (Figure 2).

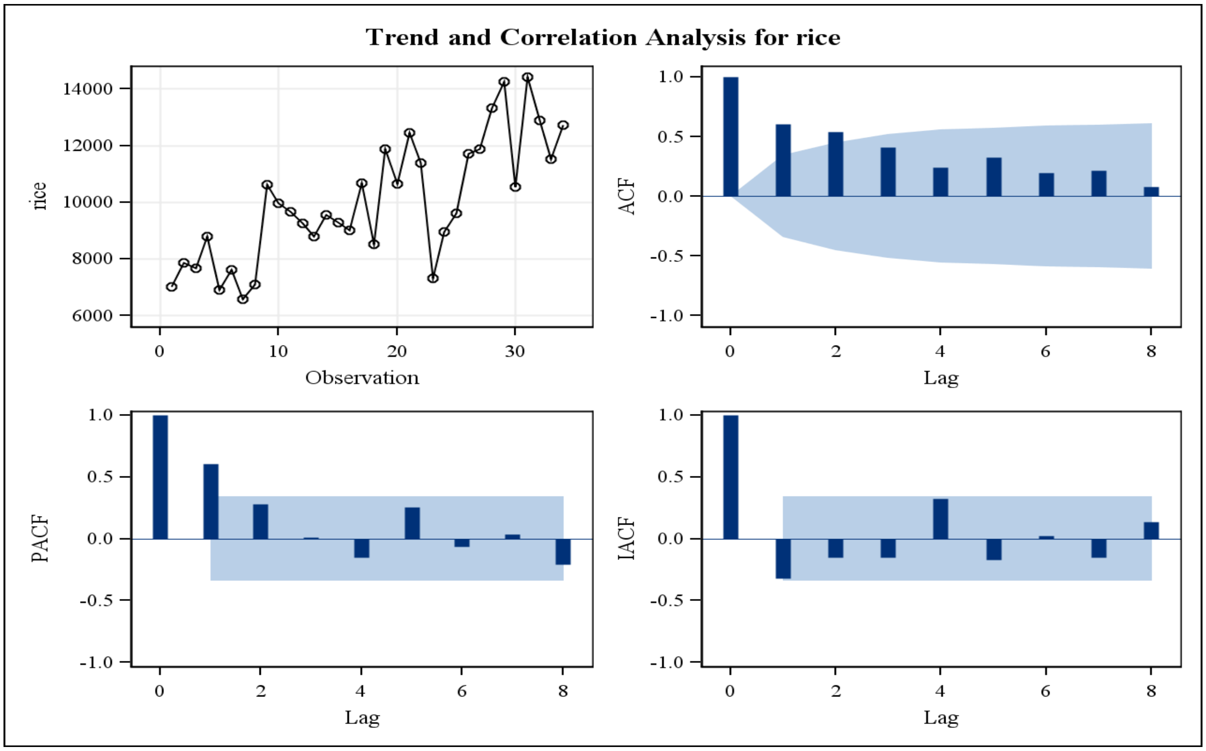

Figure 1: Time series plot and different types autocorrelation function plots of rice production in Andhra Pradesh from 1980-2014

After making the data stationary, the tentative ARIMA models with one differentiation are given in Table 1. For all the models, R^2 is greater than 50%. So the models are a good fit for the data. On the basis of minimum (or) least values of RMSE, AIC and BIC, the ARIMA (0, 1, 1) model is identified as the best model when compared with other models.

Table 1: Tentative ARIMA (*p*, *d*, *q*) Models of rice production in Andhra Pradesh

ARIMA (*p*, *d*, *q*) Models	**RMSE**	**AIC**	**SBC**	**NBIC**	**R^2 Value**
ARIMA (0,1,0)	1813.566	607.6803	609.2066	15.110	0.22
ARIMA (0,1,1)	**1444.054**	**591.1143**	**594.1143**	**14.758**	0.55
ARIMA (0,1,2)	1458.612	593.0546	597.6337	14.882	0.56
ARIMA (1,1,0)	1643.175	601.9196	604.9723	15.016	0.42
ARIMA (1,1,1)	1456.399	593.7190	598.2981	14.879	0.56
ARIMA (1,1,2)	1479.739	602.6406	608.7461	15.014	0.56
ARIMA (2,1,0)	1658.646	603.0481	607.6272	15.139	0.43
ARIMA (2,1,1)	1482.171	597.4003	603.5058	15.017	0.56
ARIMA (2,1,2)	1503.992	611.2445	618.8763	15.150	0.56

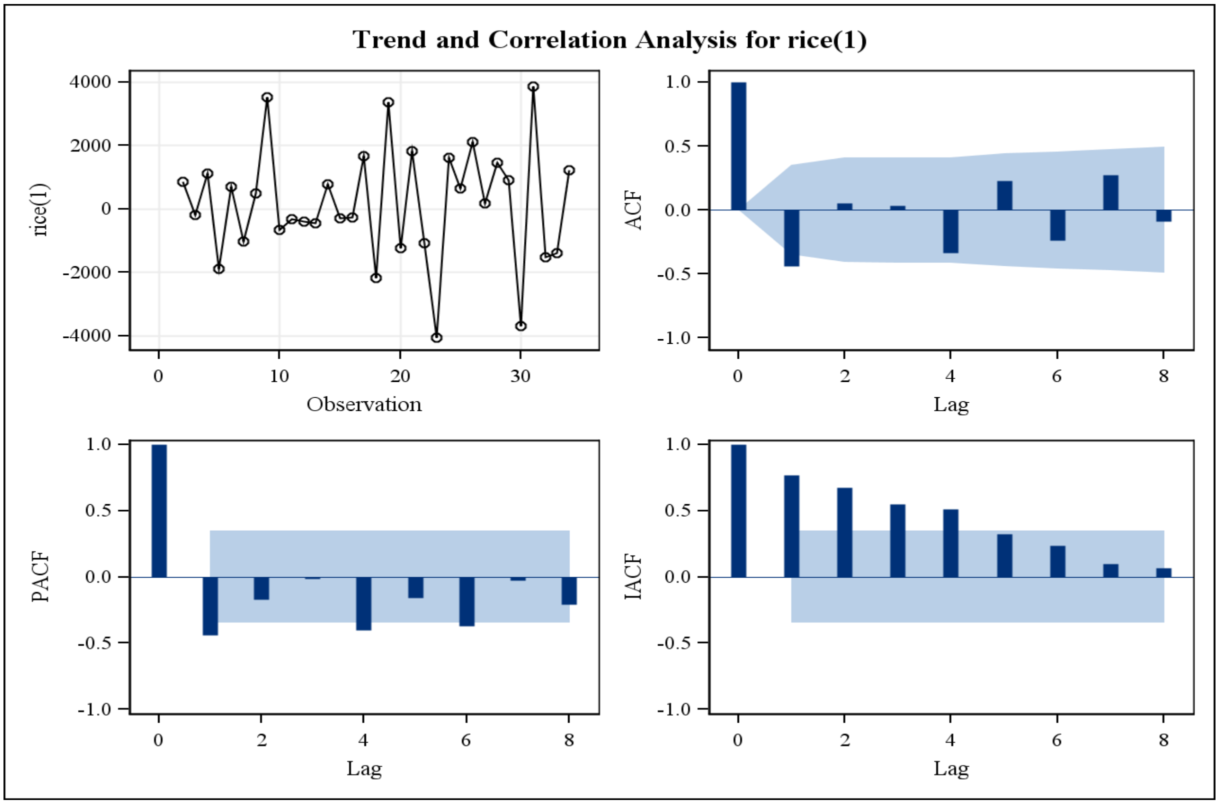

Figure 2: Time series plot, Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF) and Integrated Autocorrelation Function (IACF) with one difference for rice production

Table 2 shows that the estimates of a parameter of ARIMA (0, 1, 1) and it is clearly stated that the MA (1) and the constant term is statistically very high significant with coefficients (0.9999, 176.1107) and standard errors 0.1346 and 18.5880 respectively. R^2 criteria value for this model is 0.552 which means that the series has 55.2% of the variation in the model.

Table 2: Parameter estimates of ARIMA (0, 1, 1) Model for rice production

Parameter (s)	**Coefficie nt**	**Standard Error**	**t Value**	**P Value**
Constant	176.1107	18.5880	9.47	< 0.0001
MA(1)	0.9999	0.1346	7.43	< 0.0001

The forecasted Values of ARIMA (0,1,1) Model for rice production in Andhra Pradesh from the year 2015-16 to 2026-27 is shown in Table 3. It is observed that the year 2015-16 of forecasted value is 13175.2577 thousand tones with their 95% confidence intervals are 10428.0701 thousand tones and 15922.4453 thousand tones, whereas in the year 2020-21 forecasted value reaches 14055.8080 with 95% confidence intervals (11308.6204, 16802.9956) respectively.

Table 3: Forecasting Values of ARIMA (0,1,1) Model for rice production in Andhra Pradesh

Year	Observed values	Forecasting values of ARIMA (0,1,1)	95% Confidence Intervals	
			Lower Limit	Upper Limit
2015-16	7488.7	13175.2577	10428.0701	15922.4453
2016-17	7449.0	13351.3677	10604.1801	16098.5553
2017-18		13527.4778	10780.2902	16274.6654
2018-19		13703.5879	10956.4003	16450.7755
2019-20		13879.6979	11132.5103	16626.8855
2020-21		14055.8080	11308.6204	16802.9956
2021-22		14231.9181	11484.7305	16979.1057
2022-23		14408.0281	11660.8405	17155.2157
2023-24		14584.1382	11836.9506	17331.3258
2024-25		14760.2483	12013.0607	17507.4359
2025-26		14936.3583	12189.1707	17683.5459
2026-27		15112.4684	12365.2808	17859.6560

In a study of Rehman (2000) [3], ARIMA (0,1,0) ARIMA (0,1,3) and ARIMA (0,1,2) are the best for local, modern and total boro rice production in Bangladesh respectively. Awal and Siddique (2011) [4] study revealed that the best models were ARIMA (4,1,4), ARIMA (2,1,1), and ARIMA (2,2,3) for Aus, Aman, and Boro rice production in Bangladesh. A study of Badmus and Ariyo (2011) [5], the model of ARIMA (1,1,1) and ARIMA (2,1,2) were estimated for cultivated area and production of maize in Nigerian. Zakari et al. (2012) [6] were developed ARIMA models for the total harvested area and total grain output and then for the two main staple crops (Millet and Sorghum) cultivated in Niger. They found that ARIMA (2, 2, 1) and ARIMA (1, 1,0) is the best fit models for the total harvested area and total grain production and ARIMA (1, 1, 0) is fit for the Millet Area and ARIMA (1,1,1) is the best fit for the Millet production. Venkatramana *et al.* (2012) [7] were studied the selection of the best model for the rainfall data in India. They fitted two ARMA Models, namely, ARMA (2,3) and ARMA (1,3) and they concluded that ARMA (2,3) model is the best model than ARMA (1,3) model for rainfall data. In a study by Jambhulkar (2013, a) [8], the ARIMA (2,2,0) model is found to be the best model for describing the rice production in West Bengal. Jambhulkar (2013, b) [9] was found that ARIMA (1,1,2) model is the best model for describing the rice production in Punjab. Biswas and Bhattacharyya (2013) [10] studied ARIMA modeling to forecast the area and production of rice in West Bengal. From the results, they observed that ARIMA (2,1,3) model was the best fit model for the series of the gross cultivated area of rice and ARIMA (2,1,1) model was the best fit model for rice production in West Bengal. Prabhakaran and Sivapragasam (2014) [11] were studied forecasting areas and production of rice in India using ARIMA model and they revealed that ARIMA (1,1,1) and ARIMA (1,1,1) model were selected the best fit models for rice area and production of rice in India. Kumari et al. (2014) [12] observed from the eleven ARIMA models and they found that the ARIMA (1,1,1) is the best-fitted model in prediction efficiently the rice yield of India. Mishra *et al.* (2015) [13] were examined the performance of area and production of total food grains production in India using Autoregressive Integrated Moving Average (ARIMA) model. Srinivasulu and Sarojamma (2018) used three models for interrupted time series data and Root Mean Square Error (RMSE) measure was used for fitting the best model for Acute Coronary Events (ACEs). In the present study, RMSE was also used as a measure to select best model among different types of ARIMA models for rice production of data in Andhra Pradesh.

CONCLUSION

It is concluded that the ARIMA (0,1,1) model is selected for the best model on the basis of having the least value of RMSE, AIC, SBC, NBIC among other models. The forecasting of the rice production of Andhra Pradesh is calculated for the year 2015-16 to 2026-27. The forecasting of rice production of Andhra Pradesh can be used by researchers and it can help to the farmers as well as the policymakers for future planning.

REFERENCES

[1]. Reserve Bank of India (RBI). "*Handbook of Statistics on Indian States 2017-18*". Reserve Bank of India, India, pp.196-198, 2018.
[2]. Box, G.E.P., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M. "*Time Series Analysis: Forecasting and control*". 5th Edition, John Wiley & Sons, Inc., USA, 2016.

[3]. Rahman, N. M. F. "Forecasting of boro rice production in Bangladesh: An ARIMA approach", *Journal of Bangladesh Agricultural University*, Vol. 8, Issue 1, pp.103-112, 2010.

[4]. Awal, M. A., and Siddique, M. A. B. "Rice production in Bangladesh employing by ARIMA model", *Bangladesh Journal of Agricultural Research*, Vol. 36, Issue 1, pp. 51-62, 2011

[5]. Badmus, M. A. and Ariyo, O. S. "Forecasting cultivated area and production of Maize in Nigerian using ARIMA Model", *Asian Journal of Agricultural Sciences*, Vol. 3, issue 3, pp. 171-176, 2011.

[6]. Zakari, Seydou and Ying, Liu. "Forecasting of Niger production and harvested area", *Asian Journal of Agricultural Sciences*, Vol. 4, Issue 4, pp.308-313, 2012.

[7]. Venkatramana Reddy, S., Karthick Kumar Reddy, G., Thasleema, SC., Ramkumar, T.K., and Sarojamma, B. "Selection of the best model for the rainfall data in India", *Mathematical Sciences International Research Journal*, Vol. 1, Issue 3, pp.1033-1038, 2012.

[8]. Jambhulkar, N. N. "Modeling of Rice Production in West Bengal", *International Journal of Scientific Research (IJSR)*, Vol. 2, Issue 7, pp. 9-10, 2013a.

[9]. Jambhulkar, N. N. "Modeling of Rice Production in Punjab using ARIMA Model", *International Journal of Scientific Research (IJSR)*, Vol. 2, Issue 8, pp. 1-2, 2013b.

[10]. Biswas, R. and Bhattacharyya, B. "ARIMA modeling to forecast area and production of rice in West Bengal", *Journal of Crop and Weed*, Vol. 9, Issue 2, pp. 26-31, 2013.

[11]. Prabakaran, K. and Sivapragasam, C. "Forecasting areas and production of rice in India using ARIMA model", *International Journal of Farm Sciences*, Vol. 4, Issue 1, pp. 99-106, 2014.

[12]. Kumari P, Mishra GC, Anil Kumar P, Shukla G, Kujur SN. "Autoregressive integrated moving average (ARIMA) approach for prediction of rice (Oryza Sativa L.) Yield in India", *The Bioscan*, Vol. 9, Issue 3, pp. 1063-1066, 2014.

[13]. Mishra P, Sahu PK, Padmanaban K, Vishwajith KP, Dhekale BS. "Study of Instability and Forecasting of Food Grain Production In India", International Journal of Agriculture Sciences, Vol 7, Issue 3, pp. 474-481, 2015.

[14]. Srinivasulu, P. and Sarojamma, B. "Fitting the Best Model for ACEs Using Interrupted Time Series Data", *International Journal of Computer Sciences and Engineering*, Vol. 6, Issue 6, pp. 753-760, 2018.

AUTHORS PROFILE

Dr. S. C. Thasleema received Bachelor of Arts from Sri Venkateswara University, Tirupai, A.P, India 2006, Master of Science from Sri Venkateswara University, Tirupai, A.P, India year 2008 and PhD degree in statistics in 2017 from Sri Venkeswara University, Tirupati. She received several awards and fellowships like Meritorious University Fellowship from SV University, Tirupati for securing highest percentage at PG level during the year 2010-11, UGC-BSR Research Fellowship during 2011-2017 and International Student Travel Grant funded by World Bank for Participating in International Conference at Kualalumpur, Malaysia in the year 2014. She has twelve years of teaching experience and seven years of research experience. She taught the subjects Time series and Forecasting, Statistical Inference, Distribution Theory, Sampling techniques and stochastic process, etc. to the students of M.Sc. Statistics, Applied Statistics, and Statistics with Computer applications. She has published twelve papers in National and International Journals. She has presented nine papers in National and International conferences/workshops/seminars. She is the Life Member in many Professional Bodies like Indian Society for Probability and Statistics (ISPS), International Indian Statistics Association (IISA) and Indian Science Congress Association (ISCA). At present she is working as Guest Faculty in the Department of Statistics, Vikrama Simhapuri University, Nellore, Andhra Pradesh, India.

Dr. Ahammad Basha Shaik received Bachelor of Science from Sri Venkateswara University, Tirupai, A.P, India 2004, Master of Science from Sri Venkateswara University, Tirupai, A.P, India year 2006 and PhD degree in statistics in 2018 from Sri Venkeswara University, Tirupati. He has thirteen years of teaching experience and ten years of research experience. He taught the subjects Biostatistics, Research Methodology, Univariate and Multivariate Statistics, Reliability and Survival Analysis, Distribution Theory, Statistical computing etc. to the students of M.B.B.S., M.D./M.S., D.M./M.Ch., M.Sc. Statistics, Applied Statistics, and Statistics with Computer applications. He has published thirty papers in National and International Journals. He has presented nine papers in National and International conferences/workshops/seminars. He is the Life Member in many Professional Bodies like Indian Society for Probability and Statistics (ISPS), International Indian Statistics Association (IISA) and Indian Science Congress Association (ISCA). At present he is working as an Assistant Professor in the Department of Community Medicine, Narayana Medical College, Nellore, Andhra Pradesh, India.

