International Journal of Scientific Research in Mathematical and Statistical Sciences **Research Paper**

Volume-5, Issue-4, pp.198-208, August (2018) **E-ISSN:** 2348-4519

Autoregressive Order Selection Criteria and their Performances at Different Sizes of Time Series

P. Kiran Kumar [1], B. Sarojamma[2*], P. Ramakrishna Reddy[3]

[1, 2, 3] Department of Statistics, Sri Venkateswara University, Tirupati, India.

[*]Corresponding Author: saroja14397@rediffmail.com, Tel.: +91 9492633171

Available online at: www.isroset.org

Accepted 08/Aug/2018, Online 30/Aug/2018

Abstract—In the present study, the most popular model selection criteria are used as the autoregressive order selection criteria and the performances of time series model selection criteria at different sizes of the same time series are observed. In time series analysis and forecasting, selecting the most suitable model for a given time series and size of available time series plays a vital role. We verified that Final Prediction Error Criterion and Akaike's Information Criterion are asymptotically equivalent and Akaike's Information Criterion and Bias-Corrected Akaike's Information Criterion are asymptotically equivalent when the size of time series is large with respect to the dimension of the parameters of the autoregressive process using empirical study. All the time series model selection criteria presented in the paper are evaluated by log-likelihood function.

Keywords—*FPE, AIC, AICc, BIC, HQC, and MDL.*

I. INTRODUCTION

Model section criteria play important role in the selection of most appropriate model or the best model among the candidate models for the given time series. Mainly, we are determining the optimal order of the autoregressive model by these model selection procedures. Selecting the order of autoregressive model is one of the critical issues in time series analysis.

Some model selection procedures for time series data are

A. Final Prediction Error (FPE) Criterion
B. Akaike Information Criterion (AIC)
C. Bias-Corrected Akaike Information Criterion (AICc)
D. Bayes Information Criterion (BIC)
E. Hannan-Quinn Criterion (HQC)
F. Minimum Description Length (MDL)

A review of literature on some time series model selection methods is presented in Section II. A detailed study of the derivation of the above model selection procedures for autoregressive models is presented in Section II given by the different authors. In Section III, the most useful time series model selection procedures by Information Criteria are tabulated and a brief overview of the present study is discussed. In Section IV, we generated the autoregressive process of order 2 and the time series model selection procedures are used for selecting the optimal order of autoregressive process. Mainly, we observed the performance of the time series model selection procedures at different sizes of the same autoregressive process. In the present study, the behaviour of FPE, AIC, AICc, BIC, HQC, and MDL have been studied under standard normal errors. For this study, we used the most popular and powerful R-software. Section V contains the conclusions.

II. SOME TIME SERIES MODEL SELECTION CRITERIA

Most popular and widely used model selection procedures for time series are discussed below

A. Final Prediction Error (FPE) Criterion:

Originally, the FPE was designed for autoregressive time series models. The FPE Criterion was developed by Akaike (1969) to select the appropriate order of the autoregressive process to fit a time series data. The final prediction error (FPE) criterion has been used widely in time series model selection.

Suppose that $y_1, y_2, \cdots\cdots y_n$ is an observed series from AR (p) process and $x_1, x_2, \cdots\cdots x_n$ is an observed series from the same process which is independent of $\{y_t\}$ and $p < n$ [1]. Thus the model is

$$x_t = \sum_{j=1}^{p} a_j x_{t-j} + u_t \qquad (1)$$

where u_t are *i.i.d* $N(0, \sigma^2)$

No intercept included in the model.

Akaike estimated the mean squared prediction error for predicting x_{n+1} by estimating the parameters from the observed series $y_1, y_2, \cdots\cdots y_n$. Then the mean square prediction error is

$$E\left[\left(x_{n+1}-\hat{x}_{n+1}\right)^2\right]=E\left[\left(x_{n+1}-\hat{a}_1x_n-\hat{a}_2x_{n-1}-\cdots\cdots-\hat{a}_px_{n-p+1}\right)^2\right]$$

Where $\hat{a}_1, \hat{a}_2 \cdots\cdots \hat{a}_p$ are the maximum likelihood estimators of the coefficients of the AR (p) model based on $y_1, y_2, \cdots y_n$.

$$E\left[\left(x_{n+1}-\hat{a}_1x_n-\hat{a}_2x_{n-1}-\cdots\cdots-\hat{a}_px_{n-p+1}\right)^2\right]\approx\sigma^2\left(1+\frac{p+1}{n}\right) \quad (2)$$

For large n, $\frac{n\hat{\sigma}^2}{\sigma^2}$ distributed approximately chi-squared with $(n-p-1)$ degrees of freedom. Where $\hat{\sigma}^2$ is the maximum likelihood estimators of σ^2 .

Replace σ^2 in equation (2) by the estimator $\frac{n\hat{\sigma}^2}{(n-p-1)}$ to get the estimated mean square prediction error of x_{n+1} .

$$FPE = \frac{n\hat{\sigma}^2}{(n-p-1)}\left(1+\frac{p+1}{n}\right)$$

$$FPE = \hat{\sigma}^2\frac{n+p+1}{n-p-1}$$

where n is the number of values in the estimation data.

B. Akaike Information Criterion (AIC):

AIC is probably the most commonly used model selection criterion for time series data. The most fundamental model in time series analysis is autoregressive model [2]. In the autoregressive model, the present value of the time series is expressed as a linear combination of past values of the time series and the random component. The AR (p) model is

$$y_t=\sum_{j=1}^{p}a_jy_{t-j}+\varepsilon_t$$

where ε_t are *i.i.d* $N(0,\sigma^2)$

Where p is called the order of the AR model and a_i's are called the AR coefficients.

$$y_t/y_{t-p},\cdots\cdots y_{t-1}\sim N\left(\sum_{j=1}^{p}a_jy_{t-j},\sigma^2\right).$$

Then the conditional density of y_t given $y_{t-p},\cdots\cdots y_{t-1}$ is given by

$$f\left(y_t/y_{t-p},\cdots\cdots y_{t-1}\right)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}\left(y_t-\sum_{j=1}^{p}a_jy_{t-j}\right)^2\right\}$$

The Likelihood of the AR model with order p can be written as

$$L\left(a_1,a_2\cdots\cdots a_p,\sigma^2\right)=\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{n}\left(y_t-\sum_{j=1}^{p}a_jy_{t-j}\right)^2\right\}$$

The log-likelihood of the model can be expressed as

$$l\left(a_1,a_2\cdots\cdots a_p,\sigma^2\right)=-\frac{n}{2}\log\left(2\pi\sigma^2\right)-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_t-\sum_{j=1}^{p}a_jy_{t-j}\right)^2 \quad (3)$$

The maximum likelihood estimators of $a_1, a_2 \cdots\cdots a_p$ and σ^2 are obtained by solving the system of equations.

$$\frac{\delta l}{\delta a_1}=\frac{1}{\sigma^2}\sum_{t=1}^{n}y_{t-1}\left(y_t-\sum_{j=1}^{p}a_jy_{t-j}\right)=0$$

$$\vdots$$

$$\frac{\delta l}{\delta a_p}=\frac{1}{\sigma^2}\sum_{t=1}^{n}y_{t-p}\left(y_t-\sum_{j=1}^{p}a_jy_{t-j}\right)=0$$

and

$$\frac{\delta l}{\delta\sigma^2}=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{t=1}^{n}\left(y_t-\sum_{j=1}^{p}a_jy_{t-j}\right)^2=0$$

The maximum likelihood estimators $\hat{a}_1, \hat{a}_2 \cdots\cdots \hat{a}_p$ are obtained as the solution to the normal equation.

$$\begin{pmatrix} C(1,1) & \cdots & C(1,p) \\ \vdots & \ddots & \vdots \\ C(\mathrm{p},1) & \cdots & C(\mathrm{p},p) \end{pmatrix}\begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix}=\begin{pmatrix} C(1,0) \\ \vdots \\ C(\mathrm{p},0) \end{pmatrix}$$

where $C(\alpha,\beta)=\sum_{t=1}^{n}y_{t-\alpha}y_{t-\beta}$.

The maximum likelihood estimator of σ^2 is

$$\hat{\sigma}^2=\frac{1}{n}\sum_{t=1}^{n}\left(y_t-\sum_{j=1}^{p}a_jy_{t-j}\right)=\frac{1}{n}\left(C(0,0)-\sum_{j=1}^{p}a_jC(j,0)\right)$$

After substitution of this results in equation (3), the maximum log-likelihood is

$$l\left(\hat{a}_1,\hat{a}_2\cdots\cdots\hat{a}_p,\hat{\sigma}^2\right)=-\frac{n}{2}\log\left(2\pi\hat{\sigma}^2\right)-\frac{n}{2}$$

$$-2l\left(\hat{a}_1,\hat{a}_2\cdots\cdots\hat{a}_p,\hat{\sigma}^2\right)=n(\log 2\pi+1)+n\log\left(\hat{\sigma}^2\right)$$

In derivation, Akaike made an assumption that the true model belongs to the set of candidate models.

Akaike (1973) showed that the selection of the best model is determined by AIC score.

$$AIC = -2\ log\left(likelihood\right)+2\times number\ of\ parameters$$

Since the autoregressive model with order p has p+1 free parameters, the AIC is given by

$$AIC = -2\,l\left(\hat{a}_1,\hat{a}_2\cdots\cdots\hat{a}_p,\hat{\sigma}^2\right)+2\left(p+1\right)$$

$$AIC = n\left(\log 2\pi+1\right)+n\log\left(\hat{\sigma}^2\right)+2\left(p+1\right)$$

The constant $n\left(\log 2\pi+1\right)$ play no practical role in the model selection and can be ignored.

$$AIC = \text{n}\log\left(\hat{\sigma}^2\right)+2\left(p+1\right)$$

Where $\hat{\sigma}^2$ is the estimated error or innovation variance for the fitted pth order candidate model.

C. Bias-Corrected Akaike Information Criterion (AICc):

The true auto regressive model with true order p_* is

$$y_t=\mu_{*t}+\varepsilon_{*t}$$

(i.e. $y_t=a_1y_{t-1}+a_2y_{t-2}+\cdots\cdots+a_{p*}y_{t-p*}+\varepsilon_{*t}$)

where ε_{*t} are *i.i.d* $N\left(0,\sigma_*^2\right)$.

Let $a_*=(a_1,a_2,\cdots\cdots,a_{p*},\sigma_*^2)$ be the set of parameters for the true model.

Under the assumption of normality, the likelihood function of the true model is

$$L(a_*)=\left(2\pi\right)^{\frac{-n}{2}}\left(\sigma_*^2\right)^{\frac{-n}{2}}\exp\left\{-\frac{1}{2\sigma_*^2}\sum_{t=1}^{n}\left(y_t-\mu_{*t}\right)^2\right\}$$

The log-likelihood function of the true model is

$$l\left(a_*\right)=-\frac{n}{2}\log\left(2\pi\right)-\frac{n}{2}\log\left(\sigma_*^2\right)-\frac{1}{2\sigma_*^2}\sum_{t=1}^{n}\left(y_t-\mu_{*t}\right)^2$$

The AR (p) models is

$$y_t=a_1y_{t-1}+a_2y_{t-2}+\cdots\cdots+a_py_{t-p}+\varepsilon_t$$

where ε_t are *i.i.d* $N\left(0,\sigma^2\right)$

Let $a=(a_1,a_2,\cdots\cdots,a_p,\sigma_p^2)$ be the set of parameters for a candidate model.

Under the assumption of normality, the likelihood function of the candidate model is

$$L\left(a\right)=\left(2\pi\right)^{\frac{-n}{2}}\left(\sigma_p^2\right)^{\frac{-n}{2}}\exp\left\{-\frac{1}{2\sigma_p^2}\sum_{t=1}^{n}\left(y_t-\left(\sum_{j=1}^{p}a_jy_{t-j}\right)\right)^2\right\}$$

The log-likelihood function of the candidate model is

$$l\left(a\right)=-\frac{n}{2}\log\left(2\pi\right)-\frac{n}{2}\log\left(\sigma_p^2\right)-\frac{1}{2\sigma_p^2}\sum_{t=1}^{n}\left(y_t-\left(\sum_{j=1}^{p}a_jy_{t-j}\right)\right)^2$$

A useful measure of the discrepancy between the true and candidate model is the Kullback-Leibler information.

The discrepancy between the true model and the candidate model is

$$d(a,a_*)=E_*\left\{-2l\left(a\right)\right\}$$

Where $E_*\left(.\right)$ denotes expectation under the true model and $l\left(a\right)$ be the log-likelihood corresponding to the candidate model.

The discrepancy between the true model and the fitted model is

$$d(\hat{a},a_*)=\left[E_*\left\{-2l\left(a\right)\right\}\right]_{a=\hat{a}}$$

Yet evaluation of $d(\hat{a},a_*)$ is not possible, since doing so requires knowledge of a_*

Akaike (1973) noted that $-2l\left(\hat{a}\right)$ is a biased estimator of $d(\hat{a},a_*)$.

Bias adjustment made to obtain unbiased estimator of $d(\hat{a},a_*)$

$$E_*\left\{\left[E_*\left\{-2l\left(a\right)\right\}\right]_{a=\hat{a}}\right\}-E_*\left\{-2l\left(\hat{a}\right)\right\} \quad (4)$$

Equation (4) can be asymptotically estimated by twice of the dimension of $\hat{a}$.

In derivation, Akaike made an assumption that the true model belongs to the set of candidate models.

Under the appropriate conditions, the expected value of

$$AIC = -2l(\hat{a}) + 2(p+1)$$

should be asymptotically nearer to the expected value of $d(\hat{a}, a_*)$.

where $\hat{a} = (\hat{a}_1, \hat{a}_2 \cdots\cdots \hat{a}_p, \hat{\sigma}^2)$ are the parameter values that maximize the likelihood function.

When n is large and the dimension of $\hat{a}$ is comparatively small, AIC is approximately unbiased estimator for $E_*\{d(\hat{a}, a_*)\}$.

Twice of the dimension of $\hat{a}$ may be much smaller than the bias adjustment (4), so AIC is negatively biased estimator of $E_*\{d(\hat{a}, a_*)\}$ .To correct this negative bias, Hurvich and Tsai (1989) proposed AICc for linear regression and autoregressive modeling [3].

Hurvich and Tsai (1989) defined AICc as

$$\text{AICc} = n\log(\hat{\sigma}_p^2) + \frac{n(n+p)}{(n-p-2)} \tag{5}$$

For convenience, we will use the operationally equivalent definition

$$\text{AICc} = n\log(\hat{\sigma}_p^2) + n\log 2\pi + \frac{n(n+p)}{(n-p-2)}$$

$$\text{AICc} = n\log(\hat{\sigma}_p^2) + n + n\log 2\pi + \frac{n(n+p)}{(n-p-2)} - n$$

$$\text{AICc} = \{n\log(\hat{\sigma}_p^2) + n(1+\log 2\pi)\} + \frac{2n(p+1)}{(n-p-2)} \tag{6}$$

(6) is differ from (5), but it has no impact on selection behavior of criterion, by including additive constant.

Derivation of AICc for Autoregressive Models

Expectation of (–2 log- likelihood) of the true model is

$$E_*\{-2l(a_*)\} = E_*\left\{n\log(2\pi) + n\log(\sigma_*^2) + \frac{1}{\sigma_*^2}\sum_{t=1}^{n}(y_t - \mu_{*t})^2\right\}$$

$$E_*\{-2l(a_*)\} = n\log(2\pi) + n\log(\sigma_*^2) + E_*\left\{\frac{1}{\sigma_*^2}\sum_{t=1}^{n}(y_t - \mu_{*t})^2\right\}$$

$$E_*\{-2l(a_*)\} = n\log(2\pi) + n\log(\sigma_*^2) + n$$

$$E_*\{-2l(a_*)\} = n\log(\sigma_*^2) + n(1+\log(2\pi)) \tag{7}$$

Expectation of (–2 log- likelihood) of the fitted model is

$$E_*\{-2l(\hat{a})\} = E_*\left\{n\log(2\pi) + n\log(\hat{\sigma}_p^2) + \frac{1}{\hat{\sigma}_p^2}\sum_{t=1}^{n}\left(y_t - \left(\sum_{j=1}^{p}\hat{a}_j y_{t-j}\right)\right)^2\right\}$$

$$E_*\{-2l(\hat{a})\} = n\log(2\pi) + E_*\{n\log(\hat{\sigma}_p^2)\} + E_*\left\{\frac{1}{\hat{\sigma}_p^2}\sum_{t=1}^{n}\left(y_t - \left(\sum_{j=1}^{p}\hat{a}_j y_{t-j}\right)\right)^2\right\}$$

$$E_*\{-2l(\hat{a})\} = E_*\{n\log(\hat{\sigma}_p^2)\} + n(1+\log(2\pi)) \tag{8}$$

$(7)-(8)$ implies that

$$E_*\{-2l(a_*)\} - E_*\{-2l(\hat{a})\} = n\log(\sigma_*^2) - E_*\{n\log(\hat{\sigma}_p^2)\}$$

$$E_*\{-2l(a_*)\} - E_*\{-2l(\hat{a})\} = E_*\{n\log(\sigma_*^2)\} - E_*\{n\log(\hat{\sigma}_p^2)\}$$

$$E_*\{-2l(a_*)\} - E_*\{-2l(\hat{a})\} = -E_*\left\{n\frac{\log(\hat{\sigma}_p^2)}{\log(\sigma_*^2)}\right\} \tag{9}$$

From equation (9), we have

$$E_*\{-2l(a_*)\} = E_*\{-2l(\hat{a})\} - E_*\left\{n\frac{\log(\hat{\sigma}_p^2)}{\log(\sigma_*^2)}\right\} \tag{10}$$

$$\left[E_*\{-2l(a)\}\right]_{a=\hat{a}}$$

$$= \left[E_*\left\{n\log(2\pi) + \text{n}\log(\sigma_p^2) + \frac{1}{\sigma_p^2}\sum_{t=1}^{n}\left(y_t - \left(\sum_{j=1}^{p}a_j y_{t-j}\right)\right)^2\right\}\right]_{a=\hat{a}}$$

$$= n\log(2\pi) + \mathrm{n}\log(\hat{\sigma}_p^2) + \left[E_*\left\{\frac{1}{\sigma_p^2}\sum_{t=1}^{n}\left(\mu_{*t} + \varepsilon_{*t} - \left(\sum_{j=1}^{p} a_j y_{t-j}\right)\right)^2\right\}\right]_{a=\hat{a}}$$

$$= n\log(2\pi) + \mathrm{n}\log(\hat{\sigma}_p^2) + \frac{n\sigma_*^2}{\hat{\sigma}_p^2} + \frac{1}{\hat{\sigma}_p^2}\sum_{t=1}^{n}\left(\mu_{*t} - \left(\sum_{j=1}^{p}\hat{a}_j y_{t-j}\right)\right)^2$$

$$= n\log(2\pi) + \mathrm{n}\log(\hat{\sigma}_p^2) + \frac{n\sigma_*^2}{\hat{\sigma}_p^2} + \frac{1}{\hat{\sigma}_p^2}\sum_{t=1}^{n}\left(\left(\sum_{j=1}^{p*} a_j y_{t-j}\right) - \left(\sum_{j=1}^{p}\hat{a}_j y_{t-j}\right)\right)^2$$

$$= n\log(2\pi) + \mathrm{n}\log(\hat{\sigma}_p^2) + \frac{n\sigma_*^2}{\hat{\sigma}_p^2} + \frac{\sigma_*^2}{n\hat{\sigma}_p^2} n \frac{1}{\sigma_*^2}\sum_{t=1}^{n}\left(\left(\sum_{j=1}^{p*} a_j y_{t-j}\right) - \left(\sum_{j=1}^{p}\hat{a}_j y_{t-j}\right)\right)^2 \quad (11)$$

$\frac{n\hat{\sigma}_p^2}{\sigma_*^2} \sim \chi_{n-p}^2$ Its reciprocal follows inverse chi-square distribution. $\frac{\sigma_*^2}{n\hat{\sigma}_p^2} \sim \frac{1}{\chi_{n-p}^2}$ .The expectation of the reciprocal of the chi-square random variable with $n-p$ degrees of freedom is $\frac{1}{n-p-2}$.

The quadratic form $\frac{1}{\sigma_*^2}\sum_{t=1}^{n}\left(\left(\sum_{j=1}^{p*} a_j y_{t-j}\right) - \left(\sum_{j=1}^{p}\hat{a}_j y_{t-j}\right)\right)^2 \sim \chi_p^2$

$$E_*\left\{\left[E_*\{-2l(a)\}\right]_{a=\hat{a}}\right\}$$

$$= n\log(2\pi) + nE_*\left\{\log(\hat{\sigma}_p^2)\right\} + n^2\frac{1}{(n-p-2)} + \frac{1}{(n-p-2)}np$$

$$= n\log(2\pi) + nE_*\left\{\log(\hat{\sigma}_p^2)\right\} + \frac{n^2}{(n-p-2)} + \frac{np}{(n-p-2)} \quad (12)$$

$(12)-(7)$ implies that

$$E_*\left\{\left[E_*\{-2l(a)\}\right]_{a=\hat{a}}\right\} - E_*\{-2l(a_*)\}$$

$$= nE_*\left\{\log(\hat{\sigma}_p^2)\right\} + \frac{n^2}{(n-p-2)} + \frac{np}{(n-p-2)} - n\log(\sigma_*^2) - n$$

$$= E_*\left\{n\log(\hat{\sigma}_p^2)\right\} - E_*\left\{n\log(\sigma_*^2)\right\} + \frac{n^2}{(n-p-2)} + \frac{np}{(n-p-2)} - n$$

$$= E_*\left\{n\log(\hat{\sigma}_p^2)\right\} - E_*\left\{n\log(\sigma_*^2)\right\} + \frac{2np+2n}{(n-p-2)}$$

$$= E_*\left\{n\frac{\log(\hat{\sigma}_p^2)}{\log(\sigma_*^2)}\right\} + \frac{2n(p+1)}{(n-p-2)} \quad (13)$$

From equation (13), we have

$$E_*\left\{\left[E_*\{-2l(a)\}\right]_{a=\hat{a}}\right\}$$
$$= E_*\{-2l(a_*)\} + E_*\left\{n\frac{\log(\hat{\sigma}_p^2)}{\log(\sigma_*^2)}\right\} + \frac{2n(p+1)}{(n-p-2)}$$

Substituting (10) in the above equation, we have

$$= E_*\{-2l(\hat{a})\} - E_*\left\{n\frac{\log(\hat{\sigma}_p^2)}{\log(\sigma_*^2)}\right\} + E_*\left\{n\frac{\log(\hat{\sigma}_p^2)}{\log(\sigma_*^2)}\right\} + \frac{2n(p+1)}{(n-p-2)}$$
$$= E_*\{-2l(\hat{a})\} + \frac{2n(p+1)}{(n-p-2)}$$

Under the appropriate conditions, the expected value of

$$AICc = -2l(\hat{a}) + \frac{2n(p+1)}{(n-p-2)}$$

is exactly unbiased to the expected value of $d(\hat{a}, a_*)$.

The relationship between AIC and AICC is

$$AICc = AIC + \frac{2(p+1)(p+2)}{n-p-2}$$

When n is large with respect to the dimension of $\hat{a}$, AICc and AIC are asymptotically equivalent and hence AICc is asymptotically efficient but not consistent.

D. Bayesian Information Criterion (BIC):

BIC has been widely used for model identification in time series and linear regression analysis.

Schwarz (1978) developed a model selection criterion that was derived from a Bayesian modification of the AIC. It is also known as Schwarz Bayesian information criterion (SBC).

Let $M_1, M_2 \cdots\cdots M_r$ be r candidate models and assume that each model M_i is characterized by a parametric distribution $f\left(y/\theta_i\right)$ and the prior distribution of parameter vector θ_i [4].

The marginal likelihood of the i[th] model (when the data $y=\{y_1, y_2, \cdots\cdots y_n\}$ is given) is

$$p_i\left(y\right)=\int f_i\left(y/\theta_i\right)\,\pi_i\left(\theta_i\right)\,d\theta_i \qquad (14)$$

where θ_i is the vector of parameters in the i[th] model.

Applying Bayes theorem to calculate the posterior probability of the i[th] model given the data $y=\{y_1, y_2, \cdots\cdots y_n\}$ is

$$P\left(M_i/y\right)=\frac{p_i\left(y/M_i\right)\,P\left(M_i\right)}{\sum_{i=1}^{r} p_i\left(y/M_i\right)\,P\left(M_i\right)}\;;\; i=1,2\cdots\cdots r$$

Bayes factor $B_{12}\left(y\right)$ is used to choose between two models 1 and 2.

$$B_{12}\left(y\right)=\frac{P\left(M_1/y\right)}{P\left(M_2/y\right)}=\frac{p_1\left(y\right)P\left(M_1\right)}{p_2\left(y\right)P\left(M_2\right)}$$

If all the candidate models are equally likely then the Bayes factor $B_{12}\left(y\right)$ is

$$B_{12}\left(y\right)=\frac{p_1\left(y\right)}{p_2\left(y\right)}=\frac{\int f_1\left(y/\theta_1\right)\,\pi_1\left(\theta_1\right)\,d\theta_1}{\int f_2\left(y/\theta_2\right)\,\pi_2\left(\theta_2\right)\,d\theta_2}$$

Laplace approximation of integrals: let $q(\theta)$ be a real valued function of a p-dimensional parameter vector θ and $\hat{\theta}$ be the mode of $q(\theta)$ [5]. Then the Laplace approximation of the integral is

$$\int \exp\left\{n\,q\left(\theta\right)\right\}d\theta \approx \frac{\left(2\pi\right)^{p/2}}{n^{p/2}\left|J_q\left(\hat{\theta}\right)\right|^{1/2}}\exp\left\{n\,q\left(\theta\right)\right\}$$

Where $J_q\left(\hat{\theta}\right)=-\left.\frac{\partial^2 q\left(\theta\right)}{\partial\theta\,\partial\theta'}\right|_{\theta=\hat{\theta}}$

The marginal likelihood or the marginal distribution of data $y=\{y_1, y_2, \cdots\cdots y_n\}$ can be approximated by using Laplace's method for integrals.

The marginal likelihood (14) can be written as

$$p\left(y\right)=\int \exp\left(\log f\left(y/\theta\right)\right)\,\pi\left(\theta\right)\,d\theta$$

$$p\left(y\right)=\int \exp\left(l\left(\theta\right)\right)\,\pi\left(\theta\right)\,d\theta \qquad (15)$$

where $l\left(\theta\right)$ is the log-likelihood function.

The Taylor expansion of the log-likelihood $l\left(\theta\right)$ around $\hat{\theta}$ is

$$l\left(\theta\right)=l\left(\hat{\theta}\right)-\frac{n}{2}\left(\theta-\hat{\theta}\right)'\,J\left(\hat{\theta}\right)\left(\theta-\hat{\theta}\right)+\cdots\cdots \qquad (16)$$

Where $\hat{\theta}$ is the solution of $\frac{\delta\, l\left(\theta\right)}{\delta\theta}=0$ (i.e. $\hat{\theta}$ is maximum likelihood estimator for θ) and $J_q\left(\hat{\theta}\right)=-\left.\frac{\partial^2 q\left(\theta\right)}{\partial\theta\,\partial\theta'}\right|_{\theta=\hat{\theta}}$

The Taylor expansion of the prior distribution $\pi\left(\theta\right)$ around $\hat{\theta}$ is

$$\pi\left(\theta\right)=\pi\left(\hat{\theta}\right)-\frac{n}{2}\left(\theta-\hat{\theta}\right)'\left.\frac{\delta\, l\left(\theta\right)}{\delta\theta}\right|_{\theta=\hat{\theta}}+\cdots\cdots \qquad (17)$$

Substituting equations (16) & (17) in equation (15) and simplifying the result leads to the approximation of the marginal likelihood.

$$p\left(y\right)=\int \exp\left\{l\left(\hat{\theta}\right)-\frac{n}{2}\left(\theta-\hat{\theta}\right)'\,J\left(\hat{\theta}\right)\left(\theta-\hat{\theta}\right)+\cdots\right\}\left\{\pi\left(\hat{\theta}\right)-\frac{n}{2}\left(\theta-\hat{\theta}\right)'\left.\frac{\delta\, l\left(\theta\right)}{\delta\theta}\right|_{\theta=\hat{\theta}}+\cdots\right\}d\theta$$

$$p(y)\approx \exp\left\{l\left(\hat{\theta}\right)\right\}\,\pi\left(\hat{\theta}\right)\int \exp\left\{-\frac{n}{2}\left(\theta-\hat{\theta}\right)'\,J\left(\hat{\theta}\right)\left(\theta-\hat{\theta}\right)\right\}d\theta \qquad (18)$$

Where p-dimensional parameter vector θ follows p-variate normal distribution with mean vector $\hat{\theta}$ and variance-covariance matrix $\frac{1}{n\,J(\hat{\theta})}$.

Using Laplace approximation of the integral to the integral in equation (18), we get

$$\int \exp\left\{-\frac{n}{2}(\theta-\hat{\theta})'\,J(\hat{\theta})(\theta-\hat{\theta})\right\}d\theta \approx (2\pi)^{p/2}\,n^{-p/2}\left|J(\hat{\theta})\right|^{-1/2} \quad (19)$$

When sample size n is large, the marginal likelihood can be approximated as follows

Substituting equation (19) in equation (18), we get

$$p(y) \approx \exp\left\{l(\hat{\theta})\right\}\,\pi(\hat{\theta})(2\pi)^{p/2}\,n^{-p/2}\left|J(\hat{\theta})\right|^{-1/2}$$

Taking logarithm and multiplying with –2, we get

$$-2\log p(y) = -2\log\left\{\int f(y/\theta)\,\pi(\theta)\,d\theta\right\}$$

$$\approx -2l(\hat{\theta}) + (p+1)\log n + \log\left|J(\hat{\theta})\right| - p\log(2\pi) - n\log\pi(\hat{\theta})$$

$\log\left|J(\hat{\theta})\right| - p\log(2\pi) - n\log\pi(\hat{\theta})$ play no practical role in model selection and can be ignored.

Then the Bayesian information criterion BIC is given by

$$BIC = -2l(\hat{\theta}) + (p+1)\log n$$

$$BIC = -2\,\log f\left(y/\hat{\theta}\right) + (p+1)\log n$$

where $f\left(y/\hat{\theta}\right)$ is the statistical model estimated by maximum likelihood method.

Several authors have pointed out AIC's inconsistency that may lead to an overestimate of the true order. To overcome this inconsistency, the BIC was introduced with the penalty term on the sample size and it is a consistent estimator for large samples [6].

E. Hannan-Quinn's Criterion (HQC):

The autoregressive time series model for order p is

$$y_t = \sum_{j=1}^{p} a_j y_{t-j} + \varepsilon_t \;;\; \text{where } \varepsilon_t \text{ are } i.i.d\; N(0,\sigma^2)$$

They made the assumptions about ε_t are

(i). $E(\varepsilon_n/F_{n-1}) = 0$

(ii). $E(\varepsilon_n^2/F_{n-1}) = \sigma^2$

(iii). $E(\varepsilon_n^4/F_{n-1}) < \infty$

F_n is the $\sigma-$ algebra generated by $\{\varepsilon_n, \varepsilon_{n-1}, \cdots\cdots\cdots, \varepsilon_1\}$ [7].

Hannan-Quinn (1979) proposed an order selection criterion of the form

$$\log\hat{\sigma}_p^2 + n^{-1}2(p+1)\,c\,\log(\log n)$$

It provides a consistent estimator of order p.

Where n is the number of observations (large) and c is an arbitrary real number greater than 1.

The information criteria suggested by Hannan-Quinn is

$$IC_{HQ} = n\log\hat{\sigma}_p^2 + 2(p+1)\,c\,\log(\log \mathrm{n}) \quad (20)$$

Levinson's formula for estimating the variance $\hat{\sigma}_p^2$ of the AR model is

$$\hat{\sigma}_p^2 = \hat{\sigma}_{p-1}^2\left(1-\hat{b}_p^2\right)$$

Where $\hat{b}_p^2 = \hat{a}_{p,\mathrm{p}}^2$

$\hat{a}_{p,j}^2 = \mathrm{j}^{\mathrm{th}}$ AR coefficient of the fitted AR(p) model

$$a_{p,p} = \left[\gamma(p) - \sum_{k=1}^{p-1} a_{n-1,\mathrm{k}}\gamma(n-k)\right]\left(\sigma_{p-1}^2\right)^{-1}$$

and $\hat{\sigma}_0^2 = \gamma(0)$

$\hat{\sigma}_p^2$ is the noise variance of the fitted AR(p) model

Levinson's formula is

$$\hat{\sigma}_p^2 = \left(1-\hat{b}_p^2\right)\hat{\sigma}_{p-1}^2$$

$$= \left(1-\hat{b}_p^2\right)\left(1-\hat{b}_{p-1}^2\right)\hat{\sigma}_{p-2}^2$$

$$= \left(1-\hat{b}_p^2\right)\left(1-\hat{b}_{p-1}^2\right)\left(1-\hat{b}_{p-2}^2\right)\hat{\sigma}_{p-3}^2$$

Finally, we get by repeating the recurrence relation of variance

$$\hat{\sigma}_p^2 = \left(1-\hat{b}_p^2\right)\left(1-\hat{b}_{p-1}^2\right)\left(1-\hat{b}_{p-2}^2\right)\cdots\cdots\left(1-\hat{b}_1^2\right)\hat{\sigma}_0^2 \quad (21)$$

Substituting equation (21) in equation (20), we obtain Hannan-Quinn Information Criteria

$$IC_{HQ} = n\log\hat{\sigma}_0^2 + n\sum_{j=1}^{p}\log\left(1-\hat{b}_j^2\right) + 2(p+1)\,c\log(\log \mathrm{n})$$

Where $\hat{\sigma}_0^2$ is the variance of the AR model of order 0 (i.e. the variance of the time series $\{y_t\}$.)

Hannan and Quinn wanted a model selection criterion of a form similar to AIC yet still strongly consistent for the order p [8].

$$IC_{HQ} = n\log\hat{\sigma}_p^2 + 2(p+1)c\log(\log n);\mathrm{c}>1$$

Strong consistency hold for c = 1, then we obtain Hannan and Quinn Criterion.

$$HQC = n\log\hat{\sigma}_p^2 + 2(p+1)\log(\log \mathrm{n})$$

For convenience, we will use the operationally equivalent definition

$$HQC = \left\{n\log\hat{\sigma}_p^2 + n(1+\log 2\pi)\right\} + 2(p+1)\log(\log \mathrm{n})$$

$$HQC = -2l(\hat{a}) + 2(p+1)\log(\log \mathrm{n})$$

Hannan and Quinn criterion is not asymptotically efficient and is strongly consistent. The behavior of HQC is asymptotically very well.

Of these criteria, HQC and BIC are consistent, and AICc, AIC, FPE are asymptotically efficient [9].

F. Minimum Description Length (MDL):

Minimum Description Length (MDL) was introduced by Rissanen in 1978. Let $\left\{f(y/\theta);\theta\in\Theta\subset R^p\right\}$ is a family of probability models. Assume that the data $y=\{y_1,y_2,\cdots\cdots y_n\}$ are obtained from $f(y/\theta)$.

The Total Description Length (TDL) is defined as

$$\mathrm{TDL} = -\log f(y/\theta) + \mathrm{DL}(\text{probability distribution model})$$

The probability distribution model that minimizes this total description length is such a model that can encode the data $y=\{y_1,y_2,\cdots\cdots y_n\}$ in minimum length.

Dividing the parametric space $\Theta\subset R^p$ into infinitesimal cubes of size δ . Then the TDL depends on δ and its minimum can be approximated as

$$l(y) = -\log f\left(y/\hat{\theta}\right) + \frac{p+1}{2}\log n - \frac{p+1}{2}\log 2\pi + \log\int\sqrt{|J(\theta)|}\,d\theta + O\left(n^{-1/2}\right)$$

Where $J(\theta)$ is Fisher's information matrix [5, 10].

The Minimum Description length is defined as

$$MDL = -\log f\left(y/\hat{\theta}\right) + \frac{p+1}{2}\log n$$

$$MDL = -\log f\left(y/\hat{\theta}\right) + (p+1)\log\sqrt{n}$$

Where $\hat{\theta}$ is maximum likelihood estimator for θ.

III. METHODOLOGY

The true auto regressive model with true order p_* is

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \cdots\cdots + a_{p*} y_{t-p*} + \varepsilon_{*t}$$

Where ε_{*t} are $i.i.d\ N(0,\sigma^2)$

The candidate models are

$$\mathrm{M}_1 : y_t = a_1 y_{t-1} + \varepsilon_t$$
$$\mathrm{M}_2 : y_t = a_1 y_{t-1} + a_2 y_{t-2} + \varepsilon_t$$
$$\mathrm{M}_3 : y_t = a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + \varepsilon_t$$
$$\mathrm{M}_4 : y_t = a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + a_4 y_{t-4} + \varepsilon_t$$
$$\mathrm{M}_5 : y_t = a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + a_4 y_{t-4} + a_5 y_{t-5} + \varepsilon_t$$

The main objective of the study is to determine the true order p_* of the true auto regressive model using autoregressive order selection criterion like FPE, AIC, AICc, BIC, HQC, and MDL.

Information Criterion to choose the best model amongst the candidate model is defined as

$$IC = -2l\left(\hat{\theta}\right) + 2f(n,p)$$

Where $l\left(\hat{\theta}\right)$ is the maximized log-likelihood function.

θ is the (p+1)-vector of unknown free parameters.

$f(n,p)$ is the penalty function.

The model with the smallest value of IC is the chosen model [11]. The penalty functions of commonly used information criteria are as follows

Table 1. Penalty Functions of Some Commonly Used IC

Criterion	Penalty function $f(n,p)$
FPE	$\frac{n\log(n+p+1)-n\log(n-p-1)}{2}$
AIC	$(p+1)$
AICc	$\frac{n(p+1)}{n-p-2}$
BIC	$\frac{(p+1)}{2}\log n$
HQC	$(p+1)\log(\log n)$

The Minimum Description length is

$$MDL = -l\left(\hat{\theta}\right)+\frac{(p+1)}{2}\log n$$

The selected model M_{p^*} can be obtained by minimizing the value of the autoregressive order selection criterion.

IV. RESULTS AND DISCUSSION

The true autoregressive model used in the present study is

$y_t = 0.58 y_{t-1} - 0.65 y_{t-2} + \varepsilon_t$, where ε_t are *i.i.d* $N(0,1)$.

We generated different sizes of the same time series (i.e. AR (2) model with coefficients 0.58 and –0.65). Then autoregressive models with orders 1,2,3,4 and 5 are constructed.

Autoregressive order selection criterion like FPE, AIC, AICc, BIC, HQC, and MDL is used to select the best model. To fit the models and to find the optimal order of autoregressive model, we used packages "fpp", "forecast", "lmtest", "zoo", "fma", "expsmooth", "tseries" in R-Software.

The model with the minimum FPE, AIC, AICc, BIC, HQC, and MDL is highlighted with gray colour in Table 2 to 7.

FPE for each candidate model is determined by the following formula

$$FPE = -2\log(likelihood) + 2\times\left(\frac{n\log(n+p+1)-n\log(n-p-1)}{2}\right)$$

using R Software and these values are presented in table 2.

Table 2. Final Prediction Error

Size of Time Series	Order of AR Model				
	1	2	3	4	5
15	53.702	54.608	56.614	53.312	54.757
30	97.905	96.813	98.618	98.718	100.481
50	157.817	144.544	146.184	144.669	145.854
70	228.470	205.687	207.654	206.283	208.291
100	337.534	289.534	291.429	292.466	294.296
200	651.020	563.014	564.731	566.610	568.595
400	1318.666	1120.588	1122.403	1124.091	1125.040
800	2630.590	2289.880	2291.812	2291.356	2292.550
1600	5347.701	4592.093	4594.086	4594.327	4595.140

Akaike's final prediction error criterion is not performing well for n =15.

AIC for each candidate model for different sizes of the same time series are obtained using R Software and these values are presented in table 3.

Table 3. Akaike's Information Criterion

Size of Time Series	Order of AR Model				
	1	2	3	4	5
15	53.678	54.526	56.416	52.915	54.048
30	97.900	96.793	98.570	98.624	100.317
50	157.815	144.537	146.167	144.635	145.796
70	228.469	205.683	207.645	206.266	208.261
100	337.533	289.532	291.424	292.457	294.282
200	651.020	563.014	564.730	566.608	568.592
400	1318.666	1120.588	1122.403	1124.090	1125.039
800	2630.590	2289.880	2291.812	2291.356	2292.550
1600	5347.701	4592.093	4594.086	4594.327	4595.140

Akaike's information criterion is not performing well for n =15. FPE and AIC values are approximately equal for n exceeding 100.We observed that FPE and AIC are asymptotically equivalent from table 2 and 3.

AICc for each candidate model for different sizes of the same time series are obtained using R Software and these values are presented in table 4.

Table 4. Bias-Corrected Akaike's Information Criterion

Size of Time Series	Order of AR Model				
	1	2	3	4	5
15	54.678	56.708	60.416	59.581	64.548
30	98.344	97.716	100.170	101.124	103.969
50	158.070	145.058	147.056	145.999	147.749
70	228.648	206.047	208.260	207.204	209.595
100	337.657	289.782	291.845	293.096	295.185
200	651.081	563.136	564.935	566.918	569.027
400	1318.696	1120.648	1122.504	1124.242	1125.253
800	2630.605	2289.911	2291.862	2291.431	2292.656
1600	5347.709	4592.108	4594.111	4594.364	4595.193

Bias-Corrected Akaike's information criterion is not performing well for n =15. AIC and AICc values are approximately equal for n exceeding 100. We observed that AIC and AICc are asymptotically equivalent from table 3 and 4. We observed that AICc values are slightly higher than AIC.

BIC for each candidate model for different sizes of the same time series are obtained using R Software and these values are presented in table 5.

Table 5. Bayesian Information Criterion

Size of Time Series	**Order of AR Model**				
	1	***2***	***3***	***4***	***5***
15	55.094	56.650	59.248	56.455	58.296
30	100.702	100.997	104.175	105.630	108.724
50	161.639	150.273	153.815	154.195	157.268
70	232.966	212.429	216.639	217.509	221.752
100	342.743	297.347	301.845	305.483	309.913
200	657.616	572.909	577.923	583.100	588.382
400	1326.649	1132.562	1138.369	1144.047	1148.988
800	2639.959	2303.934	2310.550	2314.779	2320.658
1600	5358.457	4608.226	4615.597	4621.215	4627.407

Bayesian information criterion is not performing well for n =15 and 30.

HQC for each candidate model for different sizes of the same time series are obtained using R Software and these values are presented in table 6.

Table 6. Hannan-Quinn Criterion

Size of Time Series	**Order of AR Model**				
	1	***2***	***3***	***4***	***5***
15	53.663	54.503	56.385	52.877	54.002
30	98.796	98.138	100.363	100.865	103.006
50	159.271	146.721	149.079	148.276	150.165
70	230.255	208.362	211.217	210.732	213.620
100	339.642	292.695	295.642	297.729	300.608
200	653.689	567.018	570.069	573.282	576.600
400	1321.827	1125.330	1128.726	1131.993	1134.523
800	2634.189	2295.279	2299.010	2300.354	2303.348
1600	5351.695	4598.084	4602.073	4604.311	4607.122

Hannan-Quinn criterion is not performing well for n =15.

MDL for each candidate model for different sizes of the same time series are obtained using R Software and these values are presented in table 7.

Table 7. Minimum Description Length

Size of Time Series	**Order of AR Model**				
	1	***2***	***3***	***4***	***5***
15	27.547	28.325	29.624	28.228	29.148
30	50.351	50.498	52.088	52.815	52.661
50	80.819	75.136	76.908	77.098	78.634
70	116.483	106.214	108.319	108.754	110.876
100	171.372	148.674	150.923	152.742	154.956
200	328.808	286.454	288.962	291.550	294.191
400	663.324	566.281	569.185	572.024	574.494
800	1319.980	1151.967	1155.275	1157.389	1160.329
1600	2679.228	2304.113	2307.798	2310.608	2313.703

Minimum description length is not performing well for n =15, 30.The performance of minimum description length is similar as the performance of Bayesian information criterion in model selection.

V. CONCLUSION and Future Scope

In the present work, we have investigated the use of autoregressive order selection criteria for determining the order of autoregressive model. These autoregressive order selection procedures useful for obtaining best model among the candidate models in time series and forecasting. Final Prediction Error (FPE) Criterion, Akaike Information Criterion (AIC), Bias-Corrected Akaike Information Criterion (AICc), Bayesian Information Criterion (BIC), Hannan and Quinn Criterion (HQC) and Minimum Description Length (MDL) are also useful for determining best ARMA model among the candidate ARMA models.

REFERENCES

[1] Brockwell, P. J. , Davis, R. A., "*Introduction to Time Series and Forecasting*", Springer-Verlag New York Publishing, **USA**, pp.169-171, 1991.

[2] Allan D R McQuarrie, Chin-Ling Tsai, "*Regression and Time Series Model Selection*", World Scientific Publishing, **Singapore**, pp.89-97, 1998.

[3] Cavanaugh, J. E., "Unifying the derivations for the Akaike and corrected Akaike information criteria", Statistics & Probability Letters, Vol. **33**, No. **2**, pp. **201-208**, **1997**.

[4] Schwarz, G., "*Estimating the Dimension of a Model* ", Annals of Statistics, Vol. **6**, No. **2**, pp.**461-464**, **1978**.

[5] Sadanori Konish.i, Genshiro Kitagawa, "*Information Criteria and Statistical Modeling*", Springer-Verlag New York Publishing, **USA**, pp.211-218, 2008.

[6] Chang C.Y. Dorea, Catia R. Goncalves, Paulo A.A. Resende, "Simulation Results for Markov Model Seletion : AIC, BIC and EDC", In the Proceedings of the World Congress on Engineering and Computer Science (Vol II WCECS 2014), San Francisco, **USA**, pp.**22-24** ,**2014.**

[7] J. K. Ghosh, M. Delampady, and T. Samanta, "*An Introduction to Bayesian Analysis: Theory and Methods*", Springer-Verlag New York Publishing, **USA**, pp.114-115, 2006.

[8] Hannan, E. J., and B. G. Quinn, "*The Determination of the Order of an Autoregression* ", Journal of the Royal Statistical Society, Series B, Vol. **41**, No. **2**, pp.**190-195**, **1979**.

[9] Ritei Shibata, "An Optimal Selection of Regression Variables", *Biometrika*, Vol. **68**, No. **1**, pp.**45-54**, **1981**.

[10] Rissanen, J, "*Modeling by shortest data description* ", Automatica, Vol. **14**, Issue.,**5**, pp.**465–471**, **1978**.

[11] Baki Billah, Rob J. Hyndman, Anne B. Koehler, "*Empirical information criteria for time series forecasting model selection*", Journal of Statistical Computation and Simulation, Vol. **75**, No. **10**, pp.**831-840**, **2005**.

AUTHORS PROFILE

Mr. P. Kiran Kumar is pursuing Ph.D. in Statistics from Sri Venkateswara University, Tirupathi, Andhra Pradesh, India. His research topic is "Regression Aspects of Time Series Analysis". He has 15 years of experience in teaching Statistics and Mathematics for Pre University Course, Graduates, Post Graduates, Management and Chartered Accountant-Common Proficiency Test students. He has presented 3 papers in conferences and published 2 research papers in reputed international journals.

Dr. B. Sarojamma received Ph.D. degree in statistics in 2006 from Sri Venkateswara University, Tirupati, Andhra Pradesh, India. She taught the subjects Time series and Forecasting, Operations Research, Distribution Theory, Sampling Techniques and Biostatistics, etc. to the students of M.Sc. Statistics, Applied Statistics, Statistics with Computer applications, M.Sc. Home science, M.Sc. Aquaculture, Biotechnology and Bio-informatics. She has given several invited lectures in national/International Workshops/Conferences. She has published more than 62 research papers in international reputed journals and presented more than 83 papers at National/International Conferences/Symposia. She is the Life Member in many Professional Bodies like Indian Society for Probability and Statistics (ISPS), Andhra Pradesh Society for Mathematical Sciences Hyderabad, International Indian Statistics Association (IISA), Cochin, Indian Mathematical Society (IMS), Aurangabad and Indian Statistical Association (AISA), Pune. At present, she is working as an Assistant Professor in the Department of Statistics, S.V.University, Tirupati.

Dr. P. Ramakrishna Reddy received Ph.D. degree in statistics at 2015 from Sri Venkateswara University, Tirupati, Andhra Pradesh, India. He taught the subjects Time series and forecasting, Operations Research, Distribution Theory, Statistical Inference etc. to the students of M.Sc. statistics, M.Sc. Mathematics and B.Sc. Statistics students. He has presented more than 20 papers in national/international workshops/conferences. He was published more than 13 papers in international reputed journals. At present he is working as an Assistant Professor in the Department of Statistics, SDHR Degree & PG College, Tirupati. He has 10 years of experience in teaching.