IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

BAYESIAN ESTIMATION OF MEAN AND VARIANCE CHANGES IN AGRICULTURAL DATA

[1]M.Vijayakumar and [2]K. Poovizhi

[1]Associate Professor, [2]Research Scholar

Department of Statistics, Annamalai University, Annamalai Nagar, 608-002, India.

Abstract: A Bayesian Approach is considered to the problem of making inference about the point in a sequence of random variable at which the underlying distribution changes over time. In this paper, an attempt is made to identify the suitable Time Series Model and in particularly Autoregressive model in order to make a very good forecast using Bayesian Inference. Based on the literature, it is also proposed to detect the change in the AR model, which will be evaluated using Agriculture data set.

Index Terms **- AR model, Bayesian Estimation, Change of Mean and Variance.**

I. INTRODUCTION

Time series analysis is one such model which have been used in many applications including budgetary analysis, census analysis, economic forecasting, inventory studies, process and quality control, stock market analysis, utility studies, workload projections, and yield projections. A time series is a sequence of observation equally spaced over a time period. Time series data are naturally generated when a population or other phenomenon is monitored over time. It has been studied for many years and they belonged to the domain of theoretical statistics. The main aim of time series modelling is to carefully collect and rigorously study the past observations of a time series to develop an appropriate model which describes the inherent structure of the series. This model has been used in forecasting the future values for the series.

The class of autoregressive models is a rather general set of models largely used to represent stationary time series. However, time series often present change points in their dynamic structure, which may have a serious impact on the analysis and lead to misleading conclusions. A change point, which is generally the effect of an external event on the phenomenon of interest, may be represented by a change in the structure of the model or simply by a change of the value of some of the parameters. In the literature various models have been proposed to detect different kinds of change points in a time series and studied the Stationary autoregressive moving average model for detection of a change in the expected value of a process, whose behavior before and after the change follows a Bayesian model to specify prior distributions for the parameters of the models. Change in the value of the expected value is the identification of an autoregressive model to be considered.

The procedure proposed may be used also when the prior distributions of the parameters are proper; the comparison between the models may be performed using the standard definition of Bayes factor. Moreover, the methodology may be easily generalized to other frameworks: for instance, to detect changes of the variance, of the autoregressive parameters, of a deterministic trend, or the emergence of a unit root in the autoregressive polynomial. Prediction models have become an integral part of medical practice, providing information for the Clinicians, policy makers and public health system by providing support for their decision making. Note that the problem of detecting a change point in autoregressive time series under the assumption of vague prior information for the parameters within each model has been already considered. Their solution uses the device introduced based on the idea of an imaginary training sample. This study has been concerned with a problem of order change in the autoregressive model and which could be follows normal distribution. The stationary of the time series data could be same behavior and the characters with the constant probability distribution in time. Estimating the mean and variance of the order of an autoregressive model is subject to order change and given to the one changing in the parameter of the distribution (Like binomial, normal, Poisson, Gamma). There are few of the studied has been done on changing point in time series model, therefore, detection of change point in time series is important.

In this paper, to identification of autoregressive model and the contemporary detection of a possible change points in the value of the expected value and we adopt the Bayesian method and assume that the prior information about the parameters of the various models. In Section 2 we introduce the models and the notation, Section 3 The computational results are given and section 4 conclusion are presented.

II. METHODS

The autoregressive model is the most widely used time series model for predicting the future events. Auto regression is a time series model in which the observations from previous time points are considered as input to a regression equation to predict the value of the next time point. The model is characterised by the components with the corresponding parameters p. The autoregressive process threats each observation in the time series as a linear function of past value (p).

III. AUTOREGRESSIVE MODEL

Suppose observe the time series $\{y_1, y_2, \dots, y_n\}$ the autoregressive moving average model of order p can be represented as follow,

$$Y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} \; \dots\dots (1)$$

Where $\varphi_1, \varphi_2, \dots \varphi_p$ are parameters, p denotes the order of autoregressive process. The time series that has a constant mean, variance and covariance structure depends only on the difference between two points called stationary. For instance, an AR (1) autoregressive process is one in which the current value is based on the immediately preceding value, while an AR (2) process is one in which the current value is based on the previous two values. An AR (0) process is used for white noise and has no dependence between the terms. We would formulate the problem using a Bayesian approach and applied Gibb's sampler to obtain marginal posterior distribution for making decision. The main advantage of the Bayesian approach is no need to consider the number of point change in the data series. The Gibbs sampler is convenient for the traditional posterior distributions of the parameters involved are relatively simple, even though their joint posterior distribution is complicated for manually calculated. Suppose that there are n observations, let say $y_1, y_2, \dots, y_n$. We assume that $X_1, X_2, \dots, X_P = 0$. The likelihood function resulting from n observation is

$$P(X|\emptyset) \propto \prod_{i=1}^{m+1} (\sigma_i^2)^{-\frac{n_i - n_{i-1}}{2}} . \exp\left\{\frac{-S_i(\beta)}{2\sigma_i^2}\right\} \dots\dots (2)$$

Using the Gibbs sampler to find the conditional posterior distributions of subsets of the unknown parameters ($\phi, \delta, \varepsilon, \sigma^2$). the density of conditional probability of w is given g by p(w/g), and using bayesian techniques, we get the following results,

$$p = (\varphi / y,\ \delta, \beta, \varepsilon, \sigma^2) \sim N(\emptyset^*, A_*^{-1})$$

where, $x_t = y_t - \mu_t$,

where μ_t can computed form μ_p,δ and β,

$$x_t = (x_{t-1}, \dots, x_{t-p}), \hat{\emptyset} = \left(\sum x_{t-1} x_t'\right)^{-1} \sum x_{t-1} x_{t-1},$$

and the summation over t from p+1 to n.

$p\left(\frac{\sigma_a^2}{y}, \delta, \beta, \emptyset, \varepsilon\right) \sim inverted\ chi - squared$. This is the standard conjugate result, the relation between gamma distributions.

The conditional posterior probability function of δ_j.

$$P(\emptyset) \propto \left[\prod_{i=1}^{m+1} (\sigma_i^2)^{-\left(\frac{n_i - n_{i-1}}{2} + \delta_1 + 1\right)}\right] e^{-\sum_{i=1}^{m+1}\left(\frac{S_i(\beta)}{2\sigma_i^2} + \frac{\gamma_i}{\sigma_i^2}\right)} \dots\dots (3)$$

$$P(\beta, n_1, n_2, \dots, n_m) \propto e^{\frac{1}{2}(\beta - a)^2 \prod_{i=1}^{m+1} \frac{n_i - n_{i-1} + \delta_i}{\left[\frac{S_i^*(\beta)}{2}\right]^{\frac{-n_i - n_{i-1}}{2} + \delta_i}}} \dots (4)$$

IV. COMPUTATIONAL RESULTS

In the current study, we have incorporated the date for Bajira production 9in million) of agriculture data from 2000 to 2017 was used to fit the model (Agricultural Statistics at Glance 2018). The main objective is to evaluate the developed model parameter estimation. In this estimation, we used Gibbs Sampler methods is used to compute the marginal posterior distribution and Bayes estimates of the parameters. It is a powerful technique for performing the model estimation. Python Software was used to perform the parameter estimation and the results are given below.

Table 1: Bajira Production in India from 2000 to 2017.

Year	production (in million)
2000	12.04
2001	13.16
2002	11.15
2003	14.98
2004	14.17
2005	14.71
2006	15.1
2007	18.96
2008	19.73
2009	16.72
2010	21.73
2011	21.76
2012	22.26
2013	24.26
2014	24.17
2015	22.57
2016	25.9
2017	28.72

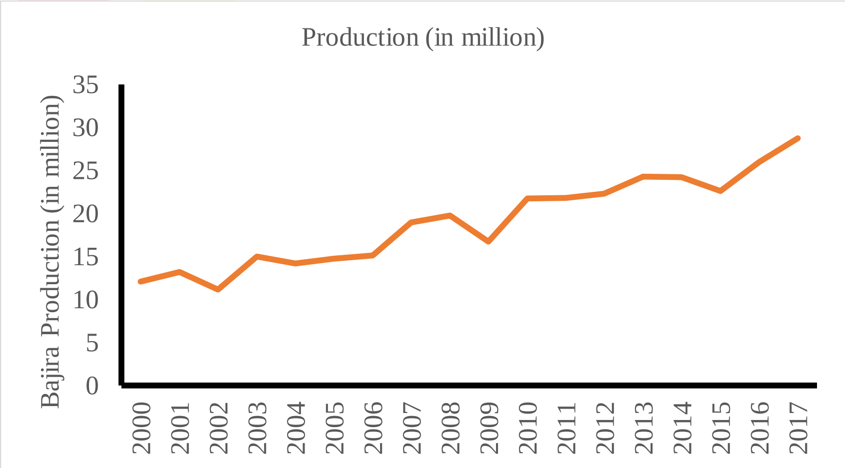
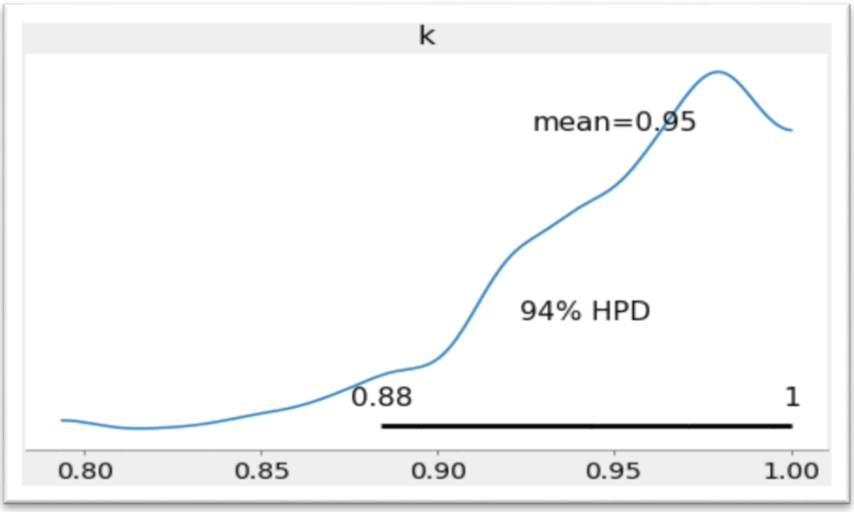


Fig. 1: Time series plot of Bajira Production in India (2000-2017)

The above figure 1 shows that the Bajira production in India is gradually increasing, but it has some changes in between the years. it clearly reveals that the data has suitable for fitting the model of Bayesian Autoregressive model. Different model has been fitted and the final model estimated values are given in the equation (5) and figure 2.

$$\hat{\alpha} = 0.002, 10^4\, Var(e_t) = \begin{cases} 0.037 & ; t < 15 \\ 0.012 & ; 15 \le t \le 21.7 \\ 0.020 & ; 21.7 \le t \le 28.7 \end{cases} \quad \dots (5)$$



V. RESULTS AND DISCUSSION

This paper the Bayesian Autoregressive model was used to detect the structural changes in time series models. It has been used in making decision about the parameters estimation of the suitable model. It has been numerically illustrated using agriculture data of Bajira Production in India and also evaluate the varaince of model selection. Python software was used to estimate the parameters. The estimates are quite close to the true values when the magnitude of the switch is large relative to the variance.

REFERENCES

[1] Albert, J.H. and Chib, S. (1993) “Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shift”, J. Business and Econ. Statist., 11,1-15.
[2] Box, G.E.P., and Jenkins, G.M. (1970). Time Series Analysis, Forecasting and Control, Holden Day, San Francisco.
[3] Broemeling, L.D. (1972). Bayesian Procedure for Detecting a Change in a Sequence of Random Variables. Metron, XXX-N-1-4, 1-14.Leamer, E.E., 1978. Specication Searches: Ad Hoc Inference with Nonexperimental Data. Wiley, New York.
[4] Moulton, B.R., 1991. A Bayesian Approach to Regression Selection and Estimation with Application to a Price Index for Radio Services. Journal of Econometrics 49, 169-193.
[5] Hodges, J.S., 1987. Uncertainty, Policy Analysis, and Statistics. Statistical Science 2, 259-291.
[6] Costa PJ, Dunyak JP, Mohtashemi M. Models, prediction, and estimation of outbreaks of infectious disease. Conf Proc - IEEE SOUTHEASTCON. 2005;174–8.
[7] Venkatesan, D., and P. Arumugam (2007).“Bayesian Analysis of Structural Changes in Autoregressive Models,” American Journal of Mathematical and Management Science, Vol.27, No.1&2, 153-162.
[8] Venkatesan, D., P.Arumugam and M.Vijayakumar(2009a),” Bayesian Analysis of Change Point Problem in Autoregressive Model with Autocorrelated Errors”, Int.J.Agricult.Stat.Sci., Vol.5, No.1,June.
[9] Venkatesan, D., P.Arumugam, M.Vijayakumar and M Gallo (2009b),” Bayesian Analysis of Change Point Problem in Autoregressive Model: A Mixture Model Approach”, Statistica and Applications, Vol.v11, 2, 223-230.
[10] Tsay, R.S. (1988) “Outliers, level shift, and variance changes in time series”, J. Forecasting, 7, 1-20.
[11] Spiegel halter, D.J. and Smith, A.F.M. (1982) “Bayes factors for linear and log-linear models with vague prior information”, J. Roy. Statist. Soc., B 44, 377-387.