Bayesian Sequential Updation and Prediction of Currency in Circulation using a Weighted Prior

Abstract

Aims/ objectives: The objective of the study is to analyse and predict the dependent variable using bayesian sequential updation of priors. An information weighted criterion is constructed from the previous information. A bayesian multiple regression model under student's t distribution and log normal error distribution is performed. Priors are updated on a sequential basis. Consequently, model comparison evaluates the best model and predictions are made for the future period . The illustration is performed on a real dataset of currency in circulation and its related macro economic variables.

Methodology: A weighted prior of the regression estimates is constructed from two mutually exclusive parts of the same data using the Deviance Information Criterion (DIC) obtained after performing the bayesian regression on the two parts under two different likelihoods- student's t and log normal distributions. The prior so constructed is further used and gets updated for prediction purposes.

Conclusion: It was found that the weighted prior thus constructed improved on sequentially after updating the priors and incorporating the previous information into the likelihood. Consequently, since DIC was the lowest in log normal error likelihood, it was concluded to be of the best fit to the dataset.

Keywords: Bayesian Sequential updation of prior; log normal error disturbances; t distribution; DIC; CIC; macroeconomic indicators; Weighted prior

1 Introduction

Bayesian sequential updating, also known as Bayesian sequential analysis or Bayesian updating, is a key concept in Bayesian statistics. It involves updating the posterior distribution sequentially as new data becomes available. This method integrates new information into the analysis, providing updated estimates and uncertainty quantification.

There are several advantages to sequentially updating the prior as new data arrives. First, the prior distribution of the parameters becomes more informative and adaptive to the model being analyzed. Second, it breaks down the computation of analyzing the entire model into manageable parts, which are updated as new data comes in. Third, this approach addresses issues related to both the volume and velocity of data.

In Sequential Bayesian Updation, the focus is on the dynamic nature of data size and online inference. It assumes data accumulate over time, potentially refining inference. Except for the initial data batch, all subsequent data batches are analyzed with informative priors, theoretically enhancing convergence speed compared to parallel techniques. [(4)].

Zellner (1971) explained the concept of sequential updation in the context of control problems, where the future period variable is estimated using minimising the loss function after observing the posterior predictive pdf for the first future variable.Lindley (1993) and Fisher (1935) discussed the concept of sequential Bayesian updation using "The Lady Testing Tea" example.Gordon, Gordon, N.J., Salmond, D.J., Smith, A.F. (1993) describes the Sequential Monte Carlo Methods (SMC) also known as particle filters.Geweke, J. (2005) used the sequential updation in extensively used in finance and econometrics for estimating parameters in time series models and forecasting.Cesa-Bianchi, N., Lugosi, G. (2006) explores the sequential bayesian updation widely used in machine learning. It is often used for the models which get updated continuously as new data arrives.Glasauer, S. (2019). Sequential Bayesian updating as a model for human perception. Progress in brain research, 249, 3-18. Kosikovaa,et. al (2021) introduces a Bayesian model updating approach that uses Gaussian Process (GP) models to describe discrepancies between the model and structural responses. Unlike mainstream methods that overlook temporal correlations in dynamical responses, this approach employs kernel-based GPs for more accurate predictions.Oravecz, Z. et al (2016) explores the big data aspect in bayesian sequential updation. Also, Glasauer, S. (2019) explains the concept of Bayesian updating for the random-change model and demonstrates its implementation using probability distributions, the Kalman filter for Gaussian distributions, and a particle filter for approximate updating.Beck, J. L., Katafygiotis, L. S. (1998) explores updating structural models and their uncertainties using dynamic response data within a Bayesian framework. Bissiri, P. G. et. al (2016) through explains that a prior belief distribution to a posterior can be made for parameters which are connected to observations through a loss function rather than the traditional likelihood function.Jaffray, J. Y. (1992) explores the bayesian updating and states that there exists a true probability that is only known to belong to a certain set of probabilities.

This study deals with the prediction of Currency in Circulation (CIC) that has been used to forecast the future demand. Curency demand is important for economy as it is a representative of various macro economic indicators such as purchasing power which consequently reflects the demand of the economy.It forms an important component of the Reserve money growth of a country. CIC can depend upon various factors which ar etaken into account in this paper. They are Payment system indicators, Weighted average call money rate, Uncertainity Index[1], credit deposit ratio etc. Chandrakar, I. (2017, February)addressess the impact of demonetisation and proposes prediction scheme using Bayesian network.It further predicts the outcome of demonetization, considering several dependencies.Albert L. et.al, (2013) developed a model to predict the CIC using SARIMA model.

[1]Economic Policy Uncertinity Index(EPU) is an index developed from newspaper articles developed by Scott. R. et.al to measure the uncertainity in the economy.

2 Materials and Methods

In this chapter, updation of priors is performed in a sequential manner, as and when new information is attained. The posterior obtained in the previous step is used as the prior for the next step. While predicting the next value of the dataset, the value predicted in the first step is incorporated in the likelihood and the next value is predicted with the help of updated priors and the new likelihood. Consequently, the process of updation of priors and prediction of the future value is performed.

2.1 Bayesian AutoRegressive Distributed Lag Model

Bayesian AutoRegressive Distributed Lag (ARDL) models combine Bayesian inference techniques with ARDL models. ARDL models are used in econometrics to analyze the relationship between variables over time, particularly when the variables might have different orders of integration.Implementing a Bayesian ARDL model involves specifying the model structure (e.g., the lag length of the autoregressive component, the order of differencing for each variable) and choosing prior distributions for the parameters.The ARDL framework can be described as below:

$$y_t = \alpha + \beta_1 y_{t-1} + \beta_2 x_{1,t-1} + \beta_3 x_{2,t-1} + \beta_4 x_{3,t-1} + \beta_5 x_{4,t-1} + \epsilon_t \tag{2.1}$$

where:
y_t: Dependent variable
x_t: Independent variable
α: coefficient of the intercept term
β: coefficient of lag of dependent variable and lag of independent variable.

The likelihood used is log normal distribution and student t distribution.

2.2 Construction of Prior Distribution

The initial prior taken is as follows:

$$\alpha, \beta \sim N(\mu, \tau) \;\; where\ \mu = 0, and \;\; \tau = 0.001\ is; the\ precision\ parameter\ (\tau = \frac{1}{\sigma^2})$$

$$\tau \;\sim\; Gamma(0.001, 0.001)$$

This was the initial prior to get the initial posterior information.

2.3 Weighted average Prior

In order to construct a robust prior, a weighted prior was used. For this, the likelihood was divided into two disjoint parts, for the first set of sample, the prior as suggested by Spiegelhalter et al. (2002) was used and for the other set of sample, the prior obtained from the first step was updated and was used on the second set of likelihood. The second set of sample is taken with a gap of some data points in a time series, to ensure that the inferences are not likelihood dominated. As, the prior obtained from the two sets of data, were not changing significantly, in terms of both direction and magnitude, they were considered to create a weighted prior. A weighted average prior was constructed using the inverse of the Deviance Information criterias (DICs), obtained from the two set of Bayesian ARDL regression models, as the weights. This weighted prior was further used to predict the future data

points.
The weighted prior distribution follows a normal distribution, with :

$$Mean\ of\ the\ weighted\ prior = \frac{\mu_1 * (\frac{1}{DIC_1}) + \mu_2 * (\frac{1}{DIC_2})}{(\frac{1}{DIC_1}) + (\frac{1}{DIC_2})} \tag{2.2}$$

$$Standard\ Deviation\ of\ the\ weighted\ prior = \frac{\sigma_1 * (\frac{1}{DIC_1}) + \sigma_2 * (\frac{1}{DIC_2})}{(\frac{1}{DIC_1}) + (\frac{1}{DIC_2})} \tag{2.3}$$

2.4 Model fit using in-sample predictions

The weighted prior constructed above was used to predict the in-sample predictions. The in-sample predictions by comparing with the actuals and with the help of DIC, assists to gauge the prediction performance of the model. In the further section, out-of–sample predictions are done.

2.5 Out-of-sample prediction along with Sequential prior updation

After creating a weighted prior, out-of-sample predictions were performed by updating the weighted prior through the sequential updation procedure as described below:

Process of sequential updation of priors and prediction of the future value:
1. The initial prior was the weighted prior constructed in the above section. It was used to run the model for the first set of likelihood, say up to, n.
2.The posterior distribution of the parameters obtained in the above model are used as the prior for the next step. Now the next value i.e., $(n+1)^{th}$ value is predicted , say, Y_{t+1}.
3.Once the future predicted value is obtained, this is added in the likelihood, to create an updated likelihood with sample size (n+1), including the new information.
4.This new likelihood and the updated prior, which is the posterior obtained in the previous step is used to predict the next future value, Y_{t+2}.
5. Steps 2 to 4 were repeated to get the future values.

3 Data and Methodology

The dataset comprises of quarterly macroeconomic high frequency indicators starting from 2005 Q1 to 2015 Q4. The dependent variable is the ratio of currency circulation and nominal GDP. The independent variables used lagged value of dependent variable, payment system of indicators, lagged value of weighted average call money rate, lagged value of uncertainty index, lagged value of credit deposit ratio.

The objective is to predict the ratio of currency in circulation in the nominal GDP given the predictor variables taken.

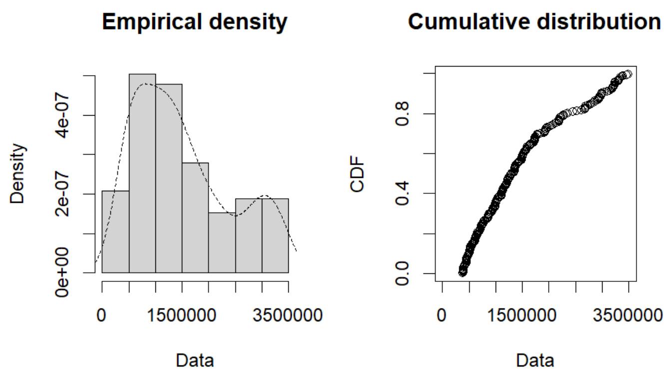


Figure 1: Empirical density and Cumulative distribution of CIC

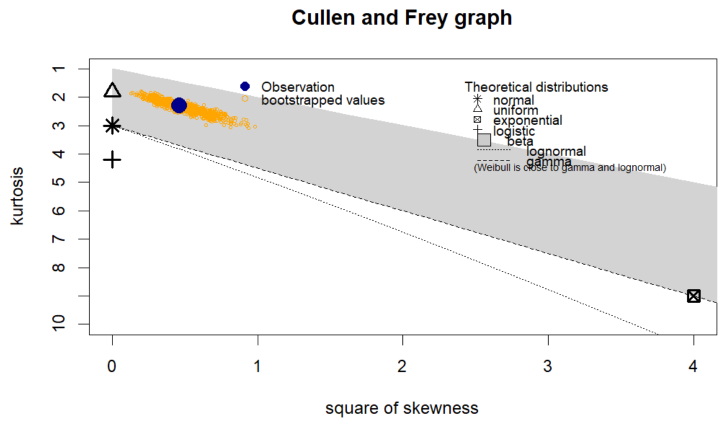


Figure 2: Cullen Frey Graph of CIC

The above graphs clearly depicts that the distribution is skewed and is non normal. Therefore, two likelihoods were taken- t distribution and log normal distribution for the error disturbances.

3.1 Seasonal Adjustment

Since the data is a periodic data, it exhibited seasonality as can be seen in the below chart. To predict the data, it was essential to deseasonalise the data. The seasonal adjustment was done using seas package of R.

ISSN: 2456-9968

(Past name: & , Past ISSN:)

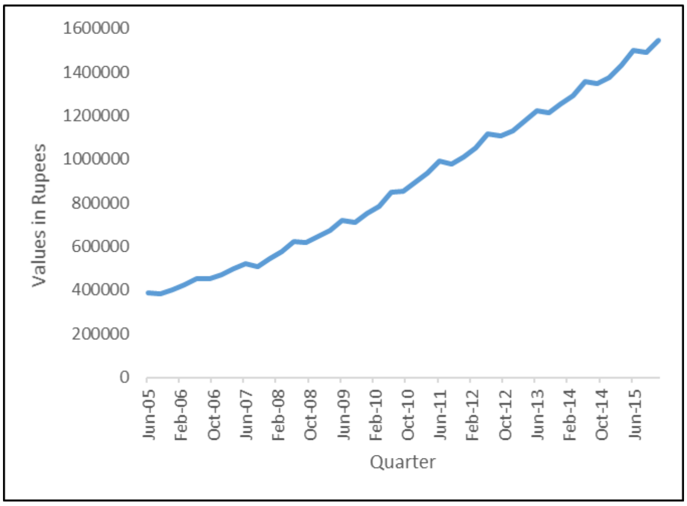


Figure 3: Plot of CIC

3.2 ARDL model

The following ARDL model is considered for the study:

$$\begin{aligned} CICNGDP_t = \alpha + \beta_1\ CICNGDP_{t-1} + \beta_2\ AlternatePayments_{1,t-1} + \beta_3\ WCMR_{2,t-1} \\ + \beta_4\ UncertaintyINDEX_{3,t-1} + \beta_5\ CreditDepositRatio_{4,t-1} + \epsilon_t \end{aligned}$$

Where, α and $\beta's$ are the regression coefficients and errors follow log normally distributed or t distribution.

4 Results and Discussion

This section discusses the results obtained through performing analysis in R software using OpenBUGS software.

4.1 Weighted Prior

The prior was constructed from two disjoint parts: Sample set 1: FY:2005-06 Q2 to FY:2010-11 Q3 and Sample set 2: FY: 2011-12 Q4 to FY:2015-16 Q2.
Prior for Sample set 1: As suggested by **Spiegelhalter et al. (2002)**, the prior distribution for the parameters of the regression model is taken as section (2.2).
Prior for Sample set 2: The posterior obtained from combining the above prior and sample set 1 was taken as the prior for the second set of sample. The posterior thus obtained are as follows:

LogNormal Model:

$$Share\ for\ Sample\ set1 : \frac{DIC_1}{DIC_1 + DIC_2} = \frac{-147.2}{-147.2 + (-103.1)} = 0.59 \quad (4.1)$$

$$Share\ for\ Sample\ set2 : \frac{DIC_2}{DIC_1 + DIC_2} = \frac{-103.1}{-147.2 + (-103.1)} = 0.41 \quad (4.2)$$

Table 1: Weighted Prior

Log normal Model			t distribution Model		
Parameter	Mean	Precision	Parameter	Mean	Precision
α	-0.79	70959.61	α	0.45	173731.99
β_1	0.02	15539.83	β_1	0.01	52616.56
β_2	-0.01	14437.55	β_2	0	45566.03
β_3	0	23588.64	β_3	0	78602.72
β_4	0.02	27747.56	β_4	0.01	64351.55
β_5	0	24299.85	β_5	0	57193.28

t distribution:

$$Share\ for\ Sample\ set1 : \frac{DIC_1}{DIC_1 + DIC_2} = \frac{-130.5}{-130.5 + (-88.1)} = 0.60 \tag{4.3}$$

$$Share\ for\ Sample\ set2 : \frac{DIC_2}{DIC_1 + DIC_2} = \frac{-88.1}{-130.5 + (-88.1)} = 0.40 \tag{4.4}$$

Using the formula described in the Section (2.2) and (2.3), the mean and the precision of the weighted prior is described in the below table. The distribution of the prior is a normal distribution. The weighted normal distribution parameters for the log normal and t distribution error models are listed below:

4.2 Sequential Updation and Prediction

The weighted prior is sequentially updated at each step after attaining posterior from the previous step. The prediction done at each step is included in the new set of likelihood for further updation of priors.
Step 1: The weighted prior constructed above is used for the Bayesian regression under two regression error models-log normal and t distribution, alternatively.
Step 2: Prediction is done for the future period y_{t+1}
Step 3: The posterior so obtained is further used in the new set of likelihood, which includes the new estimated future period, to get the further updated posterior estimates.
Step 4: Further prediction for y_{t+2} , y_{t+3} i.e., two step and three step ahead forecast are obtained by repeating the step 1 to 3.

The posterior estimates obtained after predicting the three step ahead forecast is described below for the two error distribution:

Table 2: Posterior Estimates

	Log normal Model		t distribution Model	
Parameter	**Mean**	**Precision**	**Mean**	**Precision**
α	-0.794	522916.653	0.452	1477580.237
β_1	0.025	246587.875	0.01	711884.057
β_2	-0.009	320264.147	-0.003	943677.683
β_3	-0.002	218819.262	-0.001	636625.945
β_4	0.008	277047.662	0.003	765588.535
β_5	0.004	197432.563	0.001	583783.211
DIC	**-311.2**		**-299.7**	

The DIC obtained from the above steps, in both the regression models, suggest that, log normal distribution has a better fit than the t distribution. As the DIC is reducing as the prior gets updated sequentially for predicting the future values.

The posterior distribution of the parameters under the log normal error distribution of the parameters are drawn below:

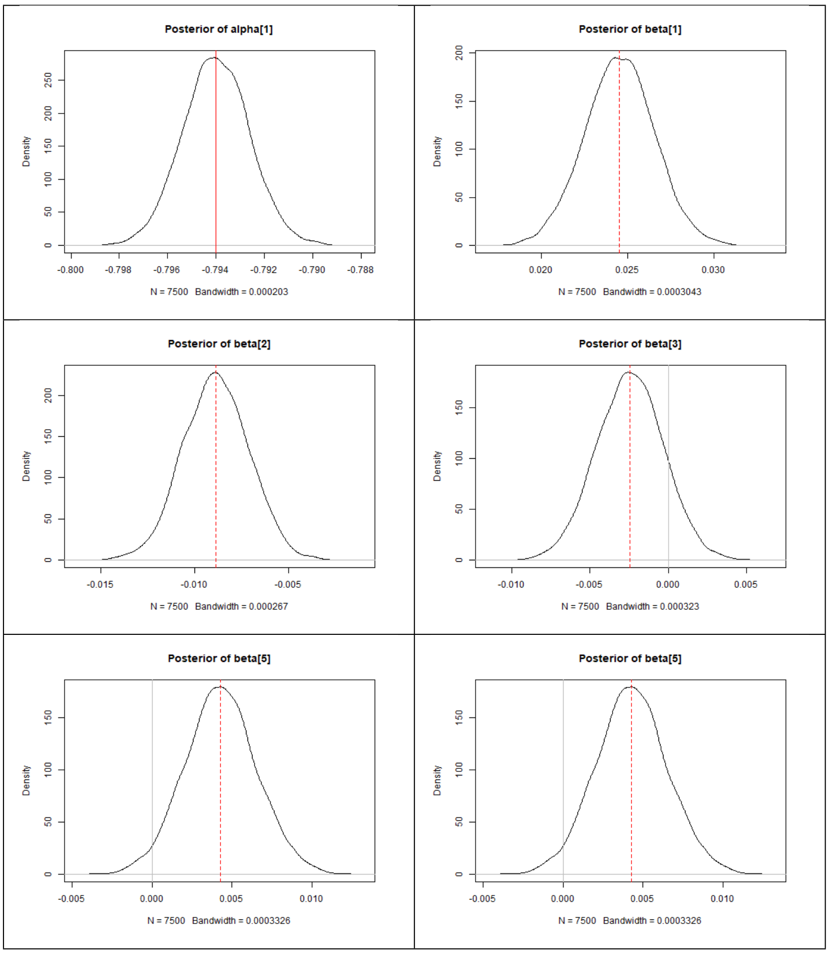


Figure 4: Posterior Estimates of the parameters

The red dotted line in the charts represent the mean of the posteriors and the grey line is the zero line. The more the deviation of the grey line from the mean, the more the parameter is statistically significant.

4.3 Posterior Predictive Probability distribution

The posterior predictive distribution is the distribution of the future values obtained after observing the set of data. It is the likelihood of the future data averaged over the posterior distribution of the parameter.
Prior Predictive distribution:

$$p(y) = \int p(y|\theta)\ p(\theta)\ d\theta \tag{4.5}$$

Posterior Predictive distribution:

$$p(y|\bar{y}) = \int p(\bar{y}|\theta)\ p(\theta|y) d\theta \tag{4.6}$$

The in sample fit was tested for the period 2015,1 i.e., 2015 March Quarter, June Quarter, September Quarter. The out-of-sample prediction, were made for 2023, March quarter, June Quarter and September Quarter. Also, it must be noted, that in out-of-sample predictions were made by following the sequential updation, i.e., updating the prior obtained from the previous model, unlike done in in-sample predictions, where each predictions were made with the same weighted prior. The difference between actual and estimated was not significant, depicting that the model was of acceptable fit.

Table 3: Prediction estimate from log normal distribution

In sample-Prediction	**Actual**	**Estimated**	**DIC**
2015,1	0.440	0.442	-290.2
2015.2	0.444	0.444	-290.2
2015,3	0.453	0.438	-290.2
Out of sample-Prediction	**Actual**	**Estimated**	
Step1- 2023, Mar Qtr	0.476	0.455	-290.2
Step 2-2023, Jun Qtr	0.470	0.454	-301.4
Step 3-2023, Sep Qtr	0.453	0.448	-311.2

Table 4: Prediction estimate from t distribution

In sample-Prediction	**Actual**	**Estimated**	**DIC**
2015,1	0.440	0.443	
2015.2	0.444	0.437	
2015,3	0.453	0.440	
Out of sample-Prediction	**Actual**	**Estimated**	
Step1- 2023, Mar Qtr	0.476	0.459	-278.3
Step 2-2023, Jun Qtr	0.470	0.458	-289.5
Step 3-2023, Sep Qtr	0.453	0.451	-299.7

Also, the DIC for log normal distribution was lesser as compared to the DIC obtained from t distribution. This suggested that log normal error distribution can be a good fit in terms of DIC. Therefore, the posterior predictive pdf for y was derived under log normal error distribution. The new likelihood (including the estimated value of y from previous step) under log normal error model was combined with the posterior obtained from the previous step. This entailed to capture the updated information about the desired parameters and combined it with the new set of information (the predefined regressor values for the corresponding information). The following charts show the posterior predictive pdf under the selected model (log normal) .

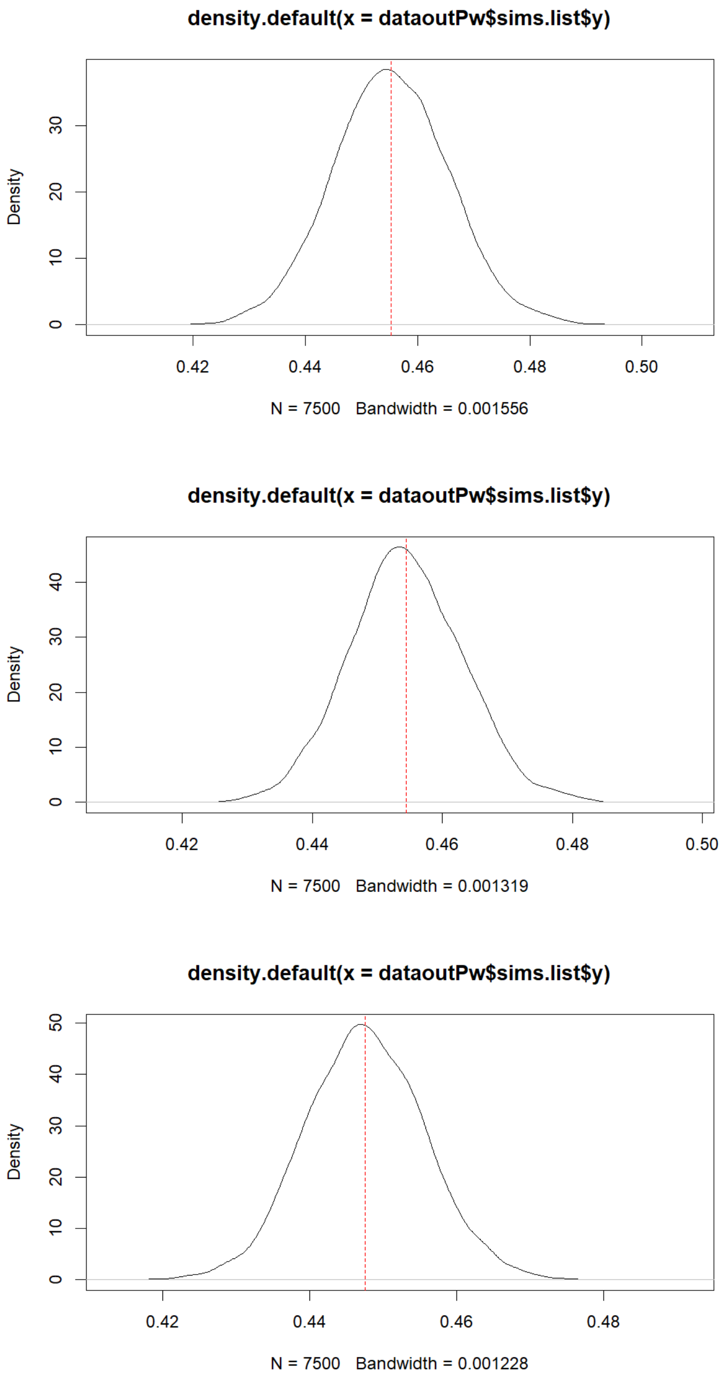


Figure 5: Posterior Predictive distribution from log normal distribution

4.4 Highest Posterior Density (HPD) Interval- LogNormal Error Model

The HPD region also is the shortest possible interval trapping the desired probability.The highest posterior density interval for the predicted value of y under log normal model are as follows:

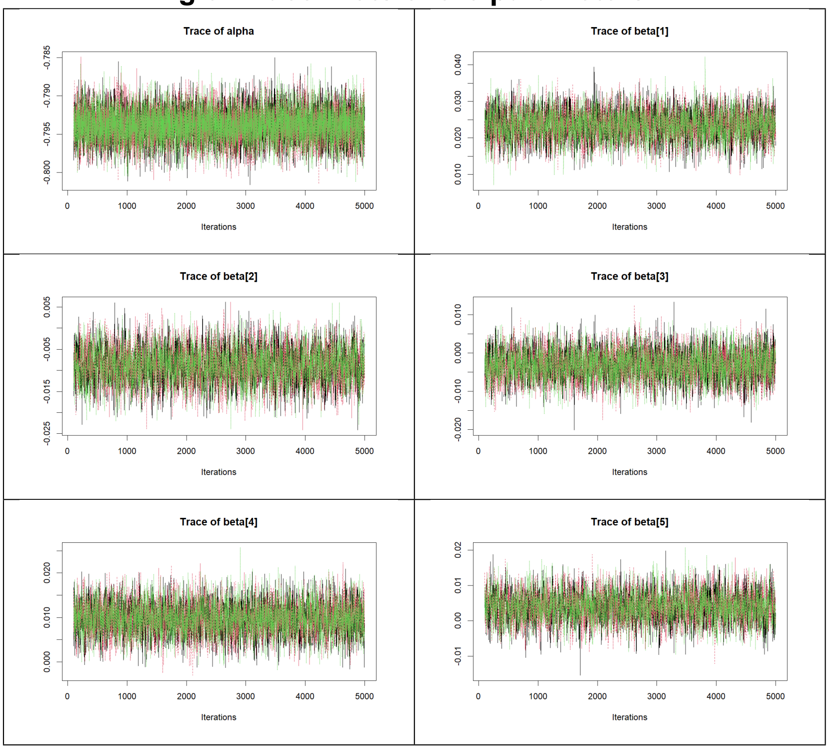
Table 5: HPD for the prediction estimates-LogNormal

	Out of sample-Prediction	HPD Interval
2023,1	0.455	(0.4351,0.476)
2023,2	0.454	(0.4374475,0.4721)
2023,3	0.448	(0.4317,0.4638)

4.5 TRACE PLOTS

Trace plots of the parameters for the first prediction is listed below:

Fig 6: Trace Plots of the parameters



5 CONCLUSION

In this chapter, the dataset of currency in circulation and factors related to it are used to exercise the Bayesian sequential updation of priors under non normal error distribution. Since the DIC got reduced, the prior constructed using the weights taken in proportion of DIC of the two disjoint datasets pertaining to the same likelihood, improved as and when the prior got updated in the light of new information. Although, the difference between actual and predicted values of y, was not much different between log normal error distribution and t error distribution, the model was best fitted to the "log normal distribution" as compared to the "student's t distribution" in terms of DIC. The out of sample predictions were made according to the log normal distrubution. Hence, it was concluded that log normal was the model of best fit and sequentially updating the model can further improve the model as and when new data is incorporated.

References

[1] Zellner, A. (1971). An introduction to Bayesian inference in econometrics. Journal of Economic Literature , 10, 1232–1234.

[2] Lindley, D. (1993). The analysis of experimental data: The appreciation of tea and wine. Teaching Statistics, 15(1), 22–25.

[3] Fisher, R. A. (1935). The design of experiments.

[4] Gordon, N. J., Salmond, D. J., Smith, A. F. (1993, April). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In IEE proceedings F (radar and signal processing), 2, 107-113.

[5] Glasauer, S. (2019). Sequential Bayesian updating as a model for human perception. Progress in brain research, 249, 3-18.

[6] Oravecz, Z., Huentelman, M., Vandekerckhove, J. (2016). Sequential Bayesian updating for big data. In Big data in cognitive science (pp. 22-42). Psychology Press.

[7] Kosikovaa, A. M., Sedehia, O., Katafygiotisa, L. S. (2021). Bayesian Model Updating using Gaussian Process Regression. structure, 1, 2.

[8] Beck, J. L., Katafygiotis, L. S. (1998). Updating models and their uncertainties. I: Bayesian statistical framework. Journal of Engineering Mechanics, 124(4), 455-461.

[9] Bissiri, P. G., Holmes, C. C., Walker, S. G. (2016). A general framework for updating belief distributions. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 78(5), 1103-1130.

[10] Jaffray, J. Y. (1992). Bayesian updating and belief functions. IEEE transactions on systems, man, and cybernetics, 22(5), 1144-1152.

[11] Chandrakar, I. (2017, February). Analysis of the effects of currency demonetization in India using Bayesian prediction. In 2017 2nd International Conference on Computing and Communications Technologies (ICCCT) (pp. 117-122). IEEE.

[12] Luguterah, A., Suleman, N., Anzagra, L. (2013). A predictive model for monthly currency in circulation in Ghana. Mathematical Theory and Modeling, 3(4), 43-52.
