RESEARCH PAPER | Mathematics | Volume : 5 | Issue : 5 | May 2015 | ISSN - 2249-555X

# Bayesian Nonparametric Regression Model Based on Spline

KEYWORDS: Nonparametric Regression , Spline , Bayes Approach , Prior Distribution , Posterior Distribution , Bayes factors

| Abdulhussein Saber AL-Mouel | Ameera Jaber Mohaisen | Jasim Mohammed Ali |
|---|---|---|
| Mathematics Department College of Education for Pure Science AL-Basrah University-Iraq | Mathematics Department College of Education for Pure Science AL-Basrah University-Iraq | Mathematics Department College of Education for Pure Science AL-Basrah University-Iraq |

**ABSTRACT** *In this paper , Bayesian approach to nonparametric regression model described . The nonparametric regression model is assumed to be a smooth spline . Bayes approach based on Markov chain Monte Carlo (MCMC) employed to make inferences on the resulting spline nonparametric model coefficients under some conditions on the prior and design matrix. We investigate the posterior density and identify the analytic form of the Bayes factors.*

**1. Introduction**

In the Bayesian approach to inference the fixed but unknown parameters are viewed as a random variables . It is well known that the Bayes estimate under squared error loss of any subvector of the parameters vector is the mean of its posterior distribution,[10],[12].

Markov Chain Monte Carlo (MCMC) method depends on partition of difficult and compound models into simple ones which can be manipulated and easily analyzed , specially for the posterior distribution which are not easy to find their final formula.

The Bayesian, Bayesian nonparametric and Bayesian Semiparametric regression models , were studied by many researchers for example DeRobertis and Hartign in (1981) discussed the Bayesian inference using intervals of measeres,[7]. Berger in (1985) introduced the statistical decision theory and Bayesian analysis,[4] . Lenk in (1999) presented the Bayesian inference of a Semiparametric regression model using Fourier representation,[11]. Zhao in (2000) studied the Bayesian approach to the nonparametric function estimation problems such as nonparametric regression and signal estimation and he considered the asymptotic of Bayes procedures for conjugate (Gaussian) priors,[15]. Angers and Delampady in (2001) used the Bayesian approach to the nonparametric regression model using a wavelet basis and performed the subsequence estimations,[1]. Dass and Lee in (2002) presented a note on the consistency of Bayes factors for testing point null versus nonparametric alternative,[6].Ghosh, J.K. and Ramamoorthi, R.V. in (2003) studied Bayesian nonparametric,[9]. Angers and Delampady in (2004) studied fuzzy sets in hierarchical Bayes and robust Bayes inference,[2]. Ghosh, J.K. , Delampady M. and samanta, T. in (2006) presented Bayesian analysis and discussed the theory and methods,[8]. Angers and Delampady in (2008) discussed fuzzy sets in nonparametric Bayes regression by using wavelet and membership functions and they treated the membership functions as likelihood functions for the model,[3] . Choi , Lee and Roy in (2008) investigated the large sample property of the Bayes factor for testing the parametric null model against the Semiparametric alternative model,[5] . Under some conditions on the prior and design matrix , and using algebraic smoothing , they identified the analytic form of the Bayes factor and showed that the Bayes factor is consistent. Osaba and Mitaim in (2011) examined Bayesian with adaptive fuzzy priors and the likelihoods member,[13]. Pelenis in (2012) studied the Bayesian Semiparametric regression and considered a Bayesian estimation of restricted conditional moment models with linear regression as a particular example,[14].

This paper came to shed light on the nonparametric regression model which has a nonparametric function is assumed to be a smooth spline , as well as the error term which has normal distribution with mean zero and variance $\sigma_\epsilon^2$.

In this paper , Bayesian approach based on Markov chain Monte Carlo (MCMC) employed to make inferences on the resulting spline nonparametric regression model coefficients under some conditions on the prior distribution and design matrix.

We investigate the posterior density and identify the analytic form of the Bayes factors to choose between a fully Bayesian spline nonparametric regression model with $(p+k+1)$ of parameters against a Bayesian spline nonparametric regression model with $(p+q+1)$ of parameters , where $q<k$ ,

## 2. Description of the problem

Consider the model:

$$y_i = m(x_i) + \varepsilon_i \quad , \quad i = 1,2,...,n \quad . \tag{1}$$

Where the unobserved errors $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$ are known to be i.i.d. normal with mean zero and covariance $\sigma_\varepsilon^2 I_n$ with $\sigma_\varepsilon^2$ unknown and $m(x_i)$ is nonparametric component . By using spline of degree $p$ get :

$$y_i = \beta_0 + \beta_1 x_i + ... + \beta_p x_i^p + \sum_{j=1}^{k} \beta_{j+p} (x_i - t_j)_+^p + \varepsilon_i , \tag{2}$$

where $t_1, t_2, ..., t_k$ are inner knots $a < t_1 < ... < t_k < b$ . The model (2) is rewritten as follows:

$$Y = X\beta + \varepsilon, \tag{3}$$

where

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}_{n\times 1} , \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \\ \beta_{p+1} \\ \vdots \\ \beta_{p+k} \end{bmatrix}_{(p+k+1)\times 1} , \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}, \quad \varepsilon \overset{iid}{\sim} N(0, \sigma_\varepsilon^2 I) ,$$

$$X = \begin{bmatrix} 1 & x_1 & \cdots & x_1^p & (x_1 - t_1)_+^p & \cdots & (x_1 - t_k)_+^p \\ 1 & x_2 & \cdots & x_2^p & (x_2 - t_1)_+^p & \cdots & (x_2 - t_k)_+^p \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^p & (x_n - t_1)_+^p & \cdots & (x_n - t_k)_+^p \end{bmatrix}_{n\times(p+k+1)} .$$

The estimation of the parameters $\beta$ entails minimizing the spline least squares criterion :

$$\|Y - X\beta\|^2 . \tag{4}$$

The least squares estimators from (4) are :

$$\hat{\beta} = (X^T X)^{-1} X^T Y, \tag{5}$$

and the fitted valued are:

$\hat{Y} = X\hat{\beta} = HY$ , where $H$ is the smoothing matrix given by :

Volume : 5 | Issue : 5 | May 2015 | ISSN - 2249-555X

$$H = X(X^T X)^{-1} X^T.$$

## 3. The Prior Distribution

To specify a complete Bayesian model , we need a prior distribution on $(\boldsymbol{\beta},\boldsymbol{\sigma}_\varepsilon^2)$. If a proper prior is desired , one could use a $N(0,\boldsymbol{\sigma}_\beta^2 I)$ prior with $\boldsymbol{\sigma}_\beta^2$ so large that for all intents and purposes , the normal distribution is uniform on the range of $\boldsymbol{\beta}$ . Therefore , we will use $\boldsymbol{\pi}_0(\boldsymbol{\beta}) \equiv 1$. As well as we will assume that the prior on $\boldsymbol{\sigma}_\varepsilon^2$ is inverse gamma with parameters $\boldsymbol{\alpha}_\varepsilon$ and $\boldsymbol{\beta}_\varepsilon$ i.e.

$$\boldsymbol{\pi}_0(\boldsymbol{\sigma}_\varepsilon^2) = \frac{\boldsymbol{\beta}_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\boldsymbol{\alpha}_\varepsilon)}(\boldsymbol{\sigma}_\varepsilon^2)^{-(\alpha_\varepsilon+1)} \exp\left(-\frac{\boldsymbol{\beta}_\varepsilon}{\boldsymbol{\sigma}_\varepsilon^2}\right), \tag{6}$$

where $\boldsymbol{\alpha}_\varepsilon$ and $\boldsymbol{\beta}_\varepsilon$ are hyperparameters that determine the priors and must be chosen by the statistician .

## 4. Posterior Distribution

From the model (3) we have

$$Y \mid \boldsymbol{\beta},\boldsymbol{\sigma}_\varepsilon^2 \sim N(X\boldsymbol{\beta},\boldsymbol{\sigma}_\varepsilon^2 I_n). \tag{7}$$

Then the likelihood function $L(Y \mid \boldsymbol{\beta},\boldsymbol{\sigma}_\varepsilon^2)$ can be expressed as:

$$L(Y \mid \boldsymbol{\beta},\boldsymbol{\sigma}_\varepsilon^2) = \frac{1}{(2\pi\boldsymbol{\sigma}_\varepsilon^2)^{n/2}} \exp\left\{-\frac{1}{2\boldsymbol{\sigma}_\varepsilon^2}(Y - X\boldsymbol{\beta})^T (Y - X\boldsymbol{\beta})\right\}$$

$$= \frac{1}{(2\pi\boldsymbol{\sigma}_\varepsilon^2)^{n/2}} \exp\left\{-\frac{1}{2\boldsymbol{\sigma}_\varepsilon^2}(Y - X\hat{\boldsymbol{\beta}})^T (Y - X\hat{\boldsymbol{\beta}})\right\} \times$$

$$\exp\left\{-\frac{1}{2\boldsymbol{\sigma}_\varepsilon^2}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})^T X^T X(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\right\}. \tag{8}$$

Then the joint posterior density of the coefficients $\boldsymbol{\beta}$ and the error variance $\boldsymbol{\sigma}_\varepsilon^2$ given by the expression

$$\boldsymbol{\pi}_1(\boldsymbol{\beta},\boldsymbol{\sigma}_\varepsilon^2 \mid Y) \propto L(Y \mid \boldsymbol{\beta},\boldsymbol{\sigma}_\varepsilon^2)\boldsymbol{\pi}_0(\boldsymbol{\beta},\boldsymbol{\sigma}_\varepsilon^2) \tag{9}$$

$$\propto (\boldsymbol{\sigma}_\varepsilon^2)^{-n/2} \exp\left\{-\frac{1}{2\boldsymbol{\sigma}_\varepsilon^2}(Y - X\hat{\boldsymbol{\beta}})^T (Y - X\hat{\boldsymbol{\beta}})\right\}$$

$$\times \exp\left\{-\frac{1}{2\boldsymbol{\sigma}_\varepsilon^2}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})^T X^T X(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\right\} \times \frac{\boldsymbol{\beta}_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\boldsymbol{\alpha}_\varepsilon)}(\boldsymbol{\sigma}_\varepsilon^2)^{-(\alpha_\varepsilon+1)} \exp\left\{-\frac{\boldsymbol{\beta}_\varepsilon}{\boldsymbol{\sigma}_\varepsilon^2}\right\}$$

$$\propto (\sigma_\varepsilon^2)^{-(n/2+\alpha_\varepsilon+1)} \exp\left\{-\frac{1}{2\sigma_\varepsilon^2}(\beta-\hat{\beta})^T X^T X(\beta-\hat{\beta})\right\}\times$$

$$\exp-\left\{\frac{\frac{1}{2}(Y-X\hat{\beta})^T(Y-X\hat{\beta})+\beta_\varepsilon}{\sigma_\varepsilon^2}\right\}. \qquad (10)$$

From this expression , we deduce the following conditional and marginal posterior distributions

$$\pi_1(\beta \mid \sigma_\varepsilon^2, Y) \propto \exp\left\{-\frac{1}{2\sigma_\varepsilon^2}(\beta-\hat{\beta})^T X^T X(\beta-\hat{\beta})\right\}, \qquad (11)$$

and

$$\pi_1(\sigma_\varepsilon^2 \mid \beta, Y) \propto (\sigma_\varepsilon^2)^{-(n/2+\alpha_\varepsilon+1)} \exp-\left\{\frac{\frac{1}{2}(Y-X\hat{\beta})^T(Y-X\hat{\beta})+\beta_\varepsilon}{\sigma_\varepsilon^2}\right\}. \qquad (12)$$

Therefore , it follows that

$$\beta \mid \sigma_\varepsilon^2, Y \sim N(\hat{\beta}, \sigma_\varepsilon^2 (X^T X)^{-1}) \qquad (13)$$

$$\sigma_\varepsilon^2 \mid \beta, Y \sim IG\left(\alpha_\varepsilon + \frac{n}{2}, \beta_\varepsilon + \frac{1}{2}(Y-X\hat{\beta})^T(Y-X\hat{\beta})\right) \qquad (14)$$

## 5. Model checking and Bayes factors

We would like to choose between a fully Bayesian spline nonparametric regression model with $(p+k+1)$ of parameters against a Bayesian spline nonparametric regression model with $(p+q+1)$ of parameters , where $q<k$ , by using Bayes factors for two hypotheses

$$\left.\begin{array}{l} H_0 : y_i = \beta_0 + \beta_1 x_i + \ldots + \beta_p x_i^p + \sum_{j=1}^{q} \beta_{j+p}(x_i - t_j)_+^p + \varepsilon_i \quad or \quad Y = X^0 \beta^0 + \varepsilon \\ vesus \\ H_1 : y_i = \beta_0 + \beta_1 x_i + \ldots + \beta_p x_i^p + \sum_{j=1}^{k} \beta_{j+p}(x_i - t_j)_+^p + \varepsilon_i \quad or \quad Y = X\beta + \varepsilon \end{array}\right\}, \qquad (15)$$

where $\beta^0$ is $(p+q+1)\times 1$ vectors of parameters , $X^0$ is an $n\times(p+q+1)$ design matrix and $q<k$. We compute the Bayes factor , $B_{01}$ , of $H_0$ relative to $H_1$ for testing problem (15) as follows

$$B_{01}(Y) = \frac{m(Y \mid H_0)}{m(Y \mid H_1)}, \qquad (16)$$

where $m(Y \mid H_0)$ is the marginal density of $Y$ under model $H_i, i=0,1$.

Volume : 5 | Issue : 5 | May 2015 | ISSN - 2249-555X

We have:

$$m(Y\mid H_0)=\int\int f(Y\mid\beta^c,\sigma_\varepsilon^2)\pi_1(\beta^c\mid\sigma_\varepsilon^2)\pi_2(\sigma_\varepsilon^2)\,d\beta^c\,d\sigma_\varepsilon^2$$

$$=(2\pi)^{-\frac{n}{2}}\frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)}\int(\sigma_\varepsilon^2)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}\exp\left\{-\frac{\left(\frac{1}{2}\right)(Y-x^c\beta^c)^T(Y-x^c\beta^c)+\beta_\varepsilon}{\sigma_\varepsilon^2}\right\}d\sigma_\varepsilon^2$$

$$=(2\pi)^{-\frac{n}{2}}\frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)}\int(\sigma_\varepsilon^2)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}\left(\frac{1}{2}(Y-x^c\beta^c)^T(Y-x^c\beta^c)+\beta_\varepsilon\right)^{\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}\times\left(\frac{1}{2}(Y-x^c\beta^c)^T(Y-x^c\beta^c)+\beta_\varepsilon\right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}\times\exp\left\{-\frac{\left(\frac{1}{2}\right)(Y-x^c\beta^c)^T(Y-x^c\beta^c)+\beta_\varepsilon}{\sigma_\varepsilon^2}\right\}d\sigma_\varepsilon^2$$

$$=(2\pi)^{-\frac{n}{2}}\frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)}\int\frac{\left(\frac{1}{2}(Y-x^c\beta^c)^T(Y-x^c\beta^c)+\beta_\varepsilon\right)^{\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}}{(\sigma_\varepsilon^2)^{\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}}\times\exp\left\{-\frac{\left(\frac{1}{2}\right)(Y-x^c\beta^c)^T(Y-x^c\beta^c)+\beta_\varepsilon}{\sigma_\varepsilon^2}\right\}\times\left(\frac{1}{2}(Y-x^c\beta^c)^T(Y-x^c\beta^c)+\beta_\varepsilon\right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}d\sigma_\varepsilon^2$$

$$=(2\pi)^{-\frac{n}{2}}\frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)}\int\left(\frac{\left(\frac{1}{2}(Y-x^c\beta^c)^T(Y-x^c\beta^c)+\beta_\varepsilon\right)}{(\sigma_\varepsilon^2)}\right)^{\left(\frac{n}{2}+\alpha_\varepsilon+1\right)+1}\times\exp\left\{-\frac{\left(\frac{1}{2}\right)(Y-x^c\beta^c)^T(Y-x^c\beta^c)+\beta_\varepsilon}{\sigma_\varepsilon^2}\right\}\times\left(\frac{1}{2}(Y-x^c\beta^c)^T(Y-x^c\beta^c)+\beta_\varepsilon\right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}d\sigma_\varepsilon^2$$

$$=(2\pi)^{-\frac{n}{2}}\frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)}\Gamma\left(\frac{n}{2}+\alpha_\varepsilon+2\right)\left(\frac{1}{2}(Y-x^c\beta^c)^T(Y-x^c\beta^c)+\beta_\varepsilon\right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}.\quad(17)$$

and

$$m(Y\mid H_1)=\int\int f(Y\mid\beta,\sigma_\varepsilon^2)\pi_1(\beta\mid\sigma_\varepsilon^2)\pi_2(\sigma_\varepsilon^2)\,d\beta\,d\sigma_\varepsilon^2$$

$$=(2\pi)^{-\frac{n}{2}}\frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)}\int(\sigma_\varepsilon^2)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}\exp\left\{-\frac{\left(\frac{1}{2}\right)(Y-x\beta)^T(Y-x\beta)+\beta_\varepsilon}{\sigma_\varepsilon^2}\right\}d\sigma_\varepsilon^2$$

$$=(2\pi)^{-\frac{n}{2}}\frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)}\int(\sigma_\varepsilon^2)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}\left(\frac{1}{2}(Y-x\beta)^T(Y-x\beta)+\beta_\varepsilon\right)^{\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}\times\left(\frac{1}{2}(Y-x\beta)^T(Y-x\beta)+\beta_\varepsilon\right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}\times\exp\left\{-\frac{\left(\frac{1}{2}\right)(Y-x\beta)^T(Y-x\beta)+\beta_\varepsilon}{\sigma_\varepsilon^2}\right\}d\sigma_\varepsilon^2$$

$$=(2\pi)^{-\frac{n}{2}}\frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)}\int\frac{\left(\frac{1}{2}(Y-x\beta)^T(Y-x\beta)+\beta_\varepsilon\right)^{\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}}{(\sigma_\varepsilon^2)^{\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}}\times\exp\left\{-\frac{\left(\frac{1}{2}\right)(Y-x\beta)^T(Y-x\beta)+\beta_\varepsilon}{\sigma_\varepsilon^2}\right\}\times\left(\frac{1}{2}(Y-x\beta)^T(Y-x\beta)+\beta_\varepsilon\right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}d\sigma_\varepsilon^2$$

$$=(2\pi)^{-\frac{n}{2}}\frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)}\int\left(\frac{\left(\frac{1}{2}(Y-x\beta)^T(Y-x\beta)+\beta_\varepsilon\right)}{(\sigma_\varepsilon^2)}\right)^{\left(\frac{n}{2}+\alpha_\varepsilon+1\right)+1}\times\exp\left\{-\frac{\left(\frac{1}{2}\right)(Y-x\beta)^T(Y-x\beta)+\beta_\varepsilon}{\sigma_\varepsilon^2}\right\}\times\left(\frac{1}{2}(Y-x\beta)^T(Y-x\beta)+\beta_\varepsilon\right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}d\sigma_\varepsilon^2$$

$$=(2\pi)^{-\frac{n}{2}}\frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)}\Gamma\left(\frac{n}{2}+\alpha_\varepsilon+2\right)\left(\frac{1}{2}(Y-x\beta)^T(Y-x\beta)+\beta_\varepsilon\right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}.\quad(18)$$

using the above derivations , the Bayes factor for testing problem (15) is then given by:

$$B_{01}(Y)=\frac{\left(\frac{1}{2}(Y-x^c\beta^c)^T(Y-x^c\beta^c)+\beta_\varepsilon\right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}}{\left(\frac{1}{2}(Y-x\beta)^T(Y-x\beta)+\beta_\varepsilon\right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}}\qquad(19)$$

## 6. Simulation Results

In this section , we illustrate the effectiveness of our methodology , we generated observations from the model (1) with the following regression functions :

(i) $y_1=\exp(2\pi x)$,

(ii) $y_2=Sin(2\pi x)+(x-0.5)^2$.

The observations for are generated from uniform distribution on the interval [0,1] . The sample size taken are .

$$AMSE=\frac{1}{N}\sum_{i=1}^{N}MSE(x_i),\qquad(20)$$

$$AMAE=\frac{1}{N}\sum_{i=1}^{N}MAE(x_i),\qquad(21)$$

where and are mean squared error and mean absolute error criterions respectively.

Table(1) presents summary values of the and for the estimation method . From this table we can see that the values of and when are smaller than their values for the first test function , which were (0.0006407081) and (0.000175353) respectively. While the values of and are smaller when for the second test function were (0.0001740030) and (0.000454008) respectively. Figure (1) below shows the number for iterations of Gibbs sampler which used in this paper , which was (10000) iterations for this data . While figure (2) shows density estimates based on (10000) iterations of $\sigma_\varepsilon^2$ .

Volume : 5 | Issue : 5 | May 2015 | ISSN - 2249-555X

**Table(1) results of the $(AMSE)$ and $(AMAE)$ criterions for Bayesian nonparametric regression model**

| Test functions | Sample size | $AMSE$ | $AMAE$ |
|---|---|---|---|
| $y_1$ | 25 | 0.0026328022 | 0.001751402 |
| | 50 | 0.0027315742 | 0.001767563 |
| | 100 | 0.0027701751 | 0.000842594 |
| | 150 | 0.0007417182 | 0.000206646 |
| | 200 | 0.0006407081 | 0.000175353 |
| $y_2$ | 25 | 0.0140061711 | 0.005304077 |
| | 50 | 0.0036242210 | 0.003074504 |
| | 100 | 0.0003320234 | 0.001003220 |
| | 150 | 0.0002111566 | 0.000507663 |
| | 200 | 0.0001740030 | 0.000454008 |

The model checking approach based on Bayes factors has been tested on simulated examples. These Bayes factors are given in table (2) . From this table , it can be seen that the model corresponding to the first test function obtains the largest Bayes factor when $(n = 25)$ followed by that the second test function when $(n = 25)$ , and the Bayes factor favors $H_1$ with strong evidence with all samples sizes for two test functions.

**Table(2) shows the values of Bayes factors**

| Test functions | Sample size | $B_{01}(y)$ | Evidence |
|---|---|---|---|
| $y_1$ | 25 | $2.127525\times10^{-3}$ | Strongly favors $H_1$ |
| | 50 | $2.039452\times10^{-8}$ | Strongly favors $H_1$ |
| | 100 | $3.776421\times10^{-11}$ | Strongly favors $H_1$ |
| | 150 | $1.078844\times10^{-14}$ | Strongly favors $H_1$ |
| | 200 | $4.883221\times10^{-21}$ | Strongly favors $H_1$ |
| $y_2$ | 25 | $6.543244\times10^{-8}$ | Strongly favors $H_1$ |
| | 50 | $7.997665\times10^{-13}$ | Strongly favors $H_1$ |
| | 100 | $4.988876\times10^{-33}$ | Strongly favors $H_1$ |
| | 150 | $9.075544\times10^{-46}$ | Strongly favors $H_1$ |
| | 200 | $3.111275\times10^{-67}$ | Strongly favors $H_1$ |

**Figure (1) shows (10000) iterations of the Gibbs sampler**

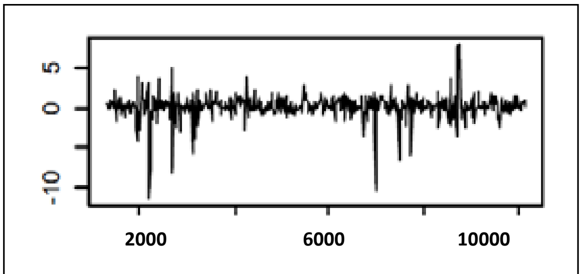
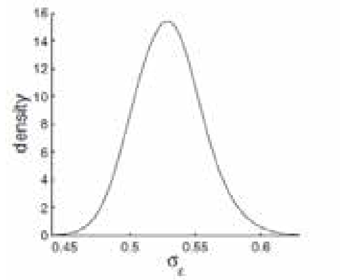



**Figure (2) shows the density estimates based on(10000) iterations of $\sigma_\varepsilon^2$**




## 7. Conclusions

The conclusions obtained throughout this paper are as follows:

(1)The posterior of $\beta$ and $\sigma_\varepsilon^2$ are respectively:

$$\beta \mid \sigma_\varepsilon^2, Y \sim N(\hat{\beta}, \sigma_\varepsilon^2 \quad \text{and} \quad \sigma_\varepsilon^2 \mid \beta, Y \sim IG\left(\alpha_\varepsilon + \frac{n}{2}, \beta_\varepsilon + \frac{1}{2}(Y - X\hat{\beta})^T (Y - X\hat{\beta})\right)$$

(2) The marginal density of $Y$ under model $H_i, i = 0,1$ are :

$$m(Y \mid H_0) = (2\pi)^{-\frac{n}{2}} \frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)} \Gamma\left(\frac{n}{2} + \alpha_\varepsilon + 2\right) \left(\frac{1}{2}\left(Y - X^0\beta^0\right)^T \left(Y - X^0\beta^0\right) + \beta_\varepsilon\right)^{-\left(\frac{n}{2} + \alpha_\varepsilon + 1\right)};$$

and

$$m(Y \mid H_1) = (2\pi)^{-\frac{n}{2}} \frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)} \Gamma\left(\frac{n}{2} + \alpha_\varepsilon + 2\right) \left(\frac{1}{2}(Y - X\beta)^T (Y - X\beta) + \beta_\varepsilon\right)^{-\left(\frac{n}{2} + \alpha_\varepsilon + 1\right)}.$$

(3) The Bayes factor for testing problem (15) is given by the following form:

$$B_{01}(Y) = \frac{\left(\frac{1}{2}\left(Y - X^0\beta^0\right)^T \left(Y - X^0\beta^0\right) + \beta_\varepsilon\right)^{-\left(\frac{n}{2} + \alpha_\varepsilon + 1\right)}}{\left(\frac{1}{2}(Y - X\beta)^T (Y - X\beta) + \beta_\varepsilon\right)^{-\left(\frac{n}{2} + \alpha_\varepsilon + 1\right)}}$$

(4) In the simulation results , we concluded the following:

(a) The values of and when are smaller than their values for the first test function , which were (0.0006407081) and (0.000175353) respectively.

(b) The values of and are smaller when for the second test function were (0.0001740030) and (0.000454008) respectively.

(c) The model corresponding to the first test function obtains the largest Bayes factor when followed by that the second test function when .

(d) The Bayes factor favors with strong evidence with all samples sizes for two test functions.

**REFERENCE**

[1] Angers, J.F. and Delampady, M. ,"Bayesian nonparametric regression | using wavelets ". Sankhy a Ser. B63, 287-308, (2001). | | [2] Angers, J.F. and Delampady, M.,"Fuzzy sets in hierarchical Bayes | and robust Bayes inference ". J. Statist. Research 38,(2004). | | [3] Angers, J.F. and Delampady, M., "Fuzzy sets in nonparametric Bayes | regression ". IMS Collections Pushing the Limits of Contemporary | Statistics : Contributions in Honor of Jayanta K. Ghosh Vol.3 ,(2008). | | [4] Berger , J. O. , "Statistical Decision Theory and Bayesian Analysis". | Springer , New York, (1985). | | [5] Choi, T., Lee ,J. and Roy , A. , "A note on the Bayes Factor in a | Semiparametric regression model", Journal of Multivariate Analysis, | 1316-1327, (2008). | | [6] Dass , S. C. and Lee , J., "A note on the consistency of Bayes factors | for testing point null versus nonparametric alternatives" J. Statist. plan. | Inference 119 (1) 143-152 , (2004). | | [7] Derobertis , L. and Hartigan, J.A., "Bayesian inference using intervals | of measures" Ann. Of Statistics . Vol. 9 , No. 2, 235-244, (1981). | | [8] Ghosh, J.K., Delampady , M. and Samanta, T. , "Introduction to | Bayesian Analysis Theory and Methods", Springer , New York, | (2006). | | [9] Ghosh, J.K. and Ramamoorthi, R.V., "Bayesian Nonparametric ", | Springer , New York , Berlin Heidelberg , Hong Kong, (2003). | | [10] Jon, W. , "Bayesian and frequenst regression methods" , Springer , | New York , Heidelberg Dordrecht London , (2013). | | [11] Lenk,P.J., "Bayesian inference for Semiparametric regression using | a Fourier representation ", J.R. Statist. Soc. Ser. B61, part 4 , (1999). | | [12] Marin, J.M. and Robert, C.P. "Bayesian Essentials with R" , Springer | Text in Statistics , New York Heidelberg Dordrecht London (2014). | | [13] Osoba, O. and Mitaim, S. , :Bayesian Inference with Adaptive Fuzzy | Priors and Likelihoods Member", IEEE, and Bart Kosko, Fellow, | IEEE vol.41, No. 5, October, (2011). | | [14] Pelenis ,J. ,"Bayesian Semiparametric Regression " Institute for | Advanced Studies , Vienna 1-39, (2012). | | [15] Zhao,L.H. "Bayesian aspects of some nonparametric problems", | Ann. Statist. 28, (2000). |