Al-Qadisiyah Journal for Administrative and Economic Sciences ISSN$_{Online}$: 2312-9883
QJAE, Volume 25, Issue 4 (2023) ISSN$_{Print}$: 1816-9171

Bayesian estimation and variable selection for binary semiparametric model.

Anaam Jameel Motar
Taha Alshaybawee
Al-Qadisiyah University, College of Administration and Economics
Corresponding Author : Anaam Jameel Motar

Abstract **:** In this research, a Bayesian hierarchical model was used to select and estimate variables.
In the context of binary regression, existing approaches to variable selection in the context of binary classification are proposed. The proposed method is that the Laplace probability of the regression parameters is proposed and estimated with a Bayesian Markov chain Monte Carlo. The conceptual result is that by doing so, the regression model is transferred from a Gaussian framework to a full Laplacian framework without sacrificing With a lot of computational efficiency. In addition, the Gibbs sampler is effective for the Parameter estimation of the proposed model and is superior to the Metropolis algorithm Which has been used in previous studies on Bayesian binary regression. Both simulation studies and real data analysis indicate that the proposed method performs well compared to other binary regression methods.

Keywords: Binary analysis, Bayesian Inference, semi-parametric model, Laplace distribution, Variable selection.

INTRODUCTION: Recently, applications of the binary regression model have become widely popular, and binary regression is one of the most well-known models that estimate the conditional mean function ($(Y|X)$). This study discusses the binary regression model proposed by (David, H.A.2003). which explains the dependent variable (response variable) y as a dichotomous variable or dummy variable, meaning that we have a binary response variable. When the response variable has exactly two values (y= 0 or y=1) we have used binary regression model analysis from a Bayesian point of view. (Katrina et, al. 2001) proposed binomial regression and the variable selection problem using the reweighted least squares method. The main objective of this research is to model binary data using the semi-parametric model, the single index model, if the Bayesian estimation method is used to estimate the model parameters and the non-parametric function. The main objective of this research is to model binary data using the semi-parametric model, the SIM if the Bayesian estimation method is used to estimate the model parameters and the non-parametric function. method (adaptive Lasso) proves the performance of Bayesian techniques such as Gibbs sampling algorithm in the prediction accuracy of binary regression. Choosing an appropriate regression model gives more interpretability and more efficient estimates, and results in a full-point estimate in terms of biases and variances of the estimators.
Thus, we can say that binary regression is appropriate for the binary response variable. Dries and Poel (2011) developed a Bayesian Gibbs sampling algorithm for a binary quantile regression model and defined the standard binary regression model as the following simple measurement equation.

$$y_i = \begin{cases} 1 & \text{if } y^* = x_i'\beta + e_i \geq 0 \\ 0 & \text{if } y^* = x_i'\beta + e_i < 0 \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 1)$$

The Single Index Model (SIM) offers an effective method for dealing with high-dimensional nonparametric estimation problems (Hardle et al ., 1993; Yu and Ruppert, 2002) and avoiding the "dimensionality curse" (Bellman et al ., 1966). Nonparametric problems assume that the response is associated with only one linear set of covariates. It is one of the most common and necessary semi-parametric models in statistics as well as applied sciences such as econometrics and psychology due to its ability to reduce dimensions (Ishimura, 1993). In this paper, the semi-parametric single index model will be used due to the importance of this model for modelling binary data and also reducing high dimensions and getting rid of the problem (the curse of dimensions). the Gaussian process will be set as a before the unknown link function and for the selected variable Laplace distribution will be set as a before the parametric index.

Bayesian single index for binary data:

A single index model can be defined as in Ishimura 1993.

$$y_i = g(x_i'\beta) + \in_i \qquad i = 1,2\ldots..n \ \ldots\ldots\ldots\ldots\ldots\ldots 1)$$

Where $y_{1,2}\ldots\ldots\ldots\ldots y_n$ are the response variables and, $\in_i$ denote the term of the error, $x_i'=(\boldsymbol{x_1},\boldsymbol{x_2},\ldots\ldots\boldsymbol{x_p})'$ is p-dimensional of an independent variable, β is the coefficient vector and g (.) is unknown nonparametric function.

The binary single index model is one of the most important tasks to study in this paper. We suppose that the response variable y_i (i=1, 2 ,……,n) are observed variable and take the values (y_i =0 or 1) . Whereas this variable is determined by the unobserved latent variable y_i* so we will rewrite the model above as follows:

$$y_i^* = g(x_i'\beta) + \epsilon_i \qquad i=1,2\ldots..n \ldots\ldots\ldots\ldots\ldots\ldots 2)$$

$y_i = g(y_i^*)$ where $y_i = \begin{cases} 1 \; if \; y^* \geq 0 \\ 0 \; if \; y^* < 0 \end{cases}$

Error term ε_i , i=1,2,…….n assumed to be iid $\sim N(0,\sigma_e^2)$,where the variance σ_e^2 is unknown

$\pi(\varepsilon_i/\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{1}{2\sigma^2}(\varepsilon_i)^2\}$ ……..…3)

There for the joint likelihood function for y_i^*, i=1,2,……..,n given X will be write as follows:

$$\pi(y^*/x,g,\beta,\sigma_e^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_e^2}} \exp\{-\frac{1}{2}\frac{(y^*-g(x'\beta)^2}{\sigma_e^2}\} \ldots\ldots\ldots\ldots\ldots\ldots 4)$$

$$\pi(y^*/x,g,\beta,\sigma_e^2) = (2\pi\sigma_e^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2}\sum_{i=1}^{n} \frac{(y_i^*-g(x\prime\beta)^2}{\sigma_e^2}\} \ldots\ldots\ldots\ldots\ldots 5)$$

Follow (chio et at,(2011)) and Gramacy and lain (2012) Gaussian process distribution is set as a prior distribution to the unknown link unction g(.).Therefore , g(.). function will be distributed as a Gaussian process with a mean zero and square exponential covariance function and can be shown as

$g_{(.)} \sim GP(0, E_{(.,.)})$ ……6)

$E(x_i, x_j) = \delta \exp\{-\frac{(x_i-x_j)^2}{c}\}$ ……7)

δ and c are the hyper parameters.

Based on this ,the prior distribution for the link function and can be written as :

$\pi(g/\beta,\sigma^2,\delta) = \det[E_n]^{-1/2} \exp\{-\frac{gn'En\,gn}{2}\}$ ……8)

E_n is covariance matrix with dimension (n×n) is given

$$(x_i\beta,\ x_j'\beta) = \mathcal{S} exp\ \{-(x_i-x_j)'\ \beta\beta'\ (x_i-x_j)\} \ldots\ldots\ldots\ldots\ldots 9)$$

Follow (Gramacy and Lain (2012)) for identifiable can do by $\frac{\beta}{\sqrt{c}}$ without the condition $|\beta| = 1$.

When Gausses process set as before the unknown link function. The parameter index β will be instead of $\frac{\beta}{\sqrt{c}}$,in the covariance matrix.

$E(x_i'\beta, x_j'\beta) = \mathcal{S}\ exp\ \{-(x_i'\beta - x_j'\beta)'\} \ldots\ldots\ldots\ldots\ldots\ldots 10)$

To the hyper parameters δ The inverse gamma distribution will be set as prior $\delta \sim IG(a_\delta, b_\delta)$ where a$\delta$and b$\delta$are the hyper parameters. also the prior distribution for $\sigma^2$is inverse gamma $\sigma^2 \sim IG(a,b)$

Since we released that the constraint for the parameter index having a unit norm then the prior distribution for β will be easier. therefore, the prior distribution for the coefficient vector β_j, j=1,2,……..,p is the independent Laplace prior distribution :

$(\beta/\lambda,\tau) = \frac{\lambda}{2\tau} \exp\{-\frac{\lambda|\beta|}{\tau}\}$ ……..11)

Where λ> 0 is a penalty parameter. This means l_1-penatly (park and Cosella 2008, Hans (2009)). There are different attractive methods that can be used to represent Laplace distribution, In our study scale mixture of normal distribution and an exponential density that introduced by (Andrews and Mallows1974)

$(\frac{\gamma}{2})\exp\{-\gamma|\beta|\}=\int_0^\infty \frac{1}{\sqrt{2\pi s}} e^{-\frac{\beta^2}{2s}} \frac{\gamma^2}{2}\exp\{\frac{-\gamma^2}{2}s\}ds$ ……..12)

Based on this for will, let assume $\gamma=\frac{\lambda}{\tau}$ then

$\pi(\beta/\gamma)=\prod_{j=1}^{p}\frac{\gamma}{2}\exp\{-\gamma|\beta_j|\}=\prod_{j=1}^{p}\frac{1}{\sqrt{2\pi sj}}\exp\{-\frac{\beta_j^2}{2s_j}\}\frac{\gamma^2}{2}\exp\{-\frac{\gamma^2}{2}s_j\}\,ds_j$ ……..13)

Hierarchical model and MCMC sampler .

As we mentioned above and Based on that assumption about the prior distribution in the last section the hierarchical model for Bayesian Binary for the single index can be emulated as follows:-

$y_i^*=g(x_i'\beta)+\in_i$ i=1,2…..n

$y_i=g(y_i^*)$ where $y_i=\begin{cases}1\ if\ y^*\geq 0\\ 0\ if\ y^*<0\end{cases}$

$\pi(y^*/x,g,\beta,\sigma_e^2)=(2\pi\sigma_e^2)^{-\frac{n}{2}}\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i^*-g(x_i'\beta)\right)^2\}$

$g/\beta,\delta\sim N(0,E)$

$\delta\sim IG(a_\delta,b_\delta)$

$\sigma^2\sim IG(c,d)$

$\beta,S/\gamma\sim\prod_{j=1}^{p}\frac{1}{\sqrt{2\pi s_j}}\exp\{-\frac{\beta_j^2}{2s_j}\}\frac{\gamma^2}{2}\exp\{-\frac{\gamma^2}{2}s_j\}\,ds_j$

$\gamma\sim Gamma(a_1,b_2)$

The full conditional posterior based on the hierarchical model above can be given as follows:

$P(\beta,g,\delta,\sigma^2,\gamma/y^*)\propto(2\pi\sigma^2)^{-\frac{n}{2}}\exp\{-\frac{1}{2\sigma^2}\sum(y_i^*-g(x_i'\beta))^2\}$

$\propto[En]^{-\frac{1}{2}}\exp\{-\frac{gnE^{-1}gn}{2}\}\times\prod_{j=1}^{p}\frac{1}{\sqrt{2\pi s_j}}\exp\{-\frac{\beta_j^2}{2s_j}\}\cdot\frac{\gamma 2}{2}\exp\{\frac{\gamma 2}{2}s_j\}$

$\propto(\sigma^2)^{c-1}\exp(-\frac{d}{\sigma^2})\times(\delta)^{-a-1}\exp\{-\frac{b}{\delta}\}\times(\gamma)^{-a-1}\exp\{-\frac{b1}{\gamma}\}$

Markov chain Mont Carlo (MCMC) algorithm will be used to derive the conditional posterior for all parameters and potential variables

1. The full conditional posterior distribution of y*

$(y^*/y,g,\beta,,s,\gamma,\delta)\begin{cases}N(g(x_i',\beta),\sigma_I^2)I(y_i^*\geq 0) & \text{if } y_i=1\\ N(g(x_i',\beta),\sigma_I^2)\ I(y_i^*<0) & \text{if } y_i=0\end{cases}$

2. The full conditional posterior distribution sampling g_n

$\pi(g_n/\beta,\sigma^2,\gamma,s,\delta,,\gamma^*)\propto p(y^*/x,\sigma^2,g_n,\beta)\times\pi(g_n/\beta,\mathcal{S})\propto[\det(D)]^{-1/2}$

$\{-(y^*-gn'D^{-1}(y*-gn)/2\}\times\det(E_n)^{-1/2}\exp\{-g_n E_n^{-1} g_n/2\}$

$(g_n/\beta,\delta,\gamma^*,\sigma^2,\gamma)\sim N(A_n,B_n)$

Where

$A_n=E(E+D)^{-1}(Y^*)$

$B_n=E(E+D)^{-1}(D)$

3. The full conditional distribution for sampling β

$\pi(\beta/g_n,\delta,\sigma^2,y^*,s,\gamma)\propto p(y^*/g_n,\beta,\sigma^2)\,\pi(g_n/\beta,\delta)\times\pi(\beta/s)$

$\propto det$ ((D+E) $^{-\frac{1}{2}}$] exp[$-\frac{y^{*}(D+E)y^{*}}{2}$]^{x} δ^{-a-1} exp{$-\frac{b}{\delta}$}
Metropolis algorithm will be used.

4. The full conditional distribution for sampling σ^2

$\pi(\sigma^2/g_n,\beta,\gamma,s,y^*)\propto \pi(y^*/g_n,\beta,\gamma,s)^{x}\,\pi(\frac{1}{\sigma^2})$
$\propto(\sigma^2)^{-\frac{n}{2}}$ exp $\{\sum_{i=1}^{n}\frac{(y^*-g_n)^2}{2\sigma^2}\}(\sigma^2)^{-c-1}$ exp$\{-\frac{d}{\sigma^2}\}$
$\propto (\frac{1}{\sigma^2})^{\frac{n}{2}+c+1}$ exp[$-\frac{1}{\sigma^2}$ [$\frac{1}{2}\sum$ (y_i*-g_n)^2+d]
Which is the inverse gamma distribution is the posterior distribution for σ^2

5. The full conditional distribution of s_j

$\pi(s_j/y^*, g_n,\beta,\sigma^2,\delta,\sigma^2)\propto \pi(\beta/s_j)\,\pi(s_{j\,/}\,\delta)$
$\propto \frac{1}{\sqrt{2\pi s_j}}$ exp $(-\frac{\beta_j^2}{2s_j})$ exp $(\frac{\gamma^2}{2}\ s_j)$
$\propto \frac{1}{\sqrt{s_j}}$ exp $\{-\frac{1}{2}(\gamma^2 s_j+\beta_j^2\ s_j)^{-1}\}$
Generalized inverse Gaussian (GIG)is conditional posterior for s_j

6. The full conditional distribution of γ

$\pi(\gamma/g_n,\beta,\delta,\sigma^2)\propto \pi(s_j/\gamma)\,\pi(\gamma)$
$\propto \prod_{j=1}^{p}\frac{\gamma}{2}$ exp $\{-\frac{\gamma^2}{2}\ \sigma_j\}^{x}$ (γ_j) $^{a1-1}$ exp $\{-b_1\gamma_j\}$
Then the conditional posterior for γ is $\pi(\gamma/\sigma_j,\beta)\sim$ Ga (a_1+2p,b_1+$\frac{\gamma}{2}$)

7. The full conditional distribution of δ

$\pi(\delta/\ \beta_j,s_j,y^*)\propto \pi(y^*/g_n,\sigma^2,\beta)\times\pi(g_n/\beta,\delta)\ \times\pi(\delta)$
$\propto$ exp $\{\frac{y^*(E_n+D)y^*}{2}\}$ [det (E_n+D)$^{-1/2}\times(\delta)^{-a-1}$ exp $\{-\frac{b}{\delta}\}$
Metropolis algorithm will be considered to sample δ.

Simulation Study

This chapter part considered the simulation examples, the simulation examples conducted based on proposed method $(BBSI)$ and compared its performance with some other methods $(BLO, BPR, and\ BBQ)$ with MCMC packages and Bayes (OR) package . The first simulation scenario considers that we have Binary response variables in the single index model ,this model implemented by our code packages and the second simulation scenario is about ordinal response variables.

We generate 12000 iterations with Gibbs Sampler algorithm, the first 2000 have burned in. The methods are evaluated based on two measures, the first one is the MSE and the second one is MAD, the formulas for these measures are defined as follows:

$$MSE(\hat{\beta}_j) = var(\hat{\beta}_j) + [Bias(\hat{\beta}_j)]^2$$

and

$$MAD = \sum_{j=1}^{m}\frac{|\hat{\beta}_j - \overline{\beta}_j|}{m}$$

Where $Bias(\hat{\beta}_j) = \overline{\beta J} - \beta_j$ with $\hat{\beta} = \sum_{j=1}^{m}\frac{\widehat{\beta}_j}{m}$

Where $\hat{\beta}$ is an estimated coefficient of β. The following simulation examples illustrate the implementation of the Binary SIM and ordinal SIM.

3.1. Example One:

The dataset in this example is generated from the following model form:

$$y^* = \sin(4z) + 0.1\varepsilon, \quad y_i = \begin{cases} 1 & if \quad y_i^* \geq 0 \\ 0 & otherwise \end{cases}$$

where $z = \boldsymbol{x}'\boldsymbol{\beta}$, $\boldsymbol{x}$ is the design matrix with dimension 5 columns, five independent variables and sample size $n=25,50,100$ *and* 150 $x_i \sim Unif[0,1]$, $(i = 1,2,..,5)$, $\boldsymbol{\beta} = (1,2,0,0,0)/\sqrt{5}$, the quantile coefficient β_τ will be estimated for the quantile levels $\tau = 0.50$ and the error term will be considered with a mixed distribution $\varepsilon \sim 0.90\,\mathrm{N}(0,1) + 0.10 Cauchy\,(0,1)$. In each level 10,000 iterations are run in the MCMC algorithm with 2,000 burn-in. See (Kuruwita, 2015) for more details. Table 1-2 shows a summary of the parameter estimates for simulation example one.

Table 3-1 the parameters optimates of simulation example one in Binary SIM

SAMPLE SIZE	*METHODS*	β_1	β_2	β_3	β_4	β_5
N=25	*BBSI*	*0.37042*	*0.36890*	*0.38545*	*0.29309*	*0.28962*
	BLO	*4.20211*	*-8.23000*	*7.37741*	*2.01792*	*-3.48984*
	BPR	*2.03037*	*-4.07218*	*3.79354*	*1.02767*	*-1.51422*
	BBQ	*3.80452*	*-8.19817*	*7.38383*	*3.05645*	*-3.57842*
N=50	*BBSI*	*0.34708*	*0.28764*	*0.26033*	*0.42862*	*0.36808*
	BLO	*-3.04967*	*-1.58856*	*-0.26717*	*0.06083*	*-2.60920*
	BPR	*-1.72009*	*-0.89350*	*-0.11072*	*-0.04871*	*-1.59950*
	BBQ	*-3.65181*	*-2.00900*	*-0.33803*	*0.38413*	*-2.60402*
N=100	*BBSI*	*0.36882*	*0.31888*	*0.45667*	*0.52216*	*0.35043*
	BLO	*-0.21994*	*-3.70091*	*0.44376*	*-0.82082*	*-0.76824*
	BPR	*-0.12181*	*-2.14240*	*0.23770*	*-0.51648*	*-0.41051*
	BBQ	*-0.48419*	*-4.72397*	*0.43861*	*-1.09455*	*-0.78586*
N=150	*BBSI*	*0.41359*	*0.45278*	*0.47039*	*0.25893*	*0.40063*
	BLO	*-1.75901*	*-3.32187*	*-0.67097*	*-0.22943*	*0.19854*
	BPR	*-0.95669*	*-1.92003*	*-0.41431*	*-0.13456*	*0.14653*
	BBQ	*-2.32390*	*-4.38509*	*-0.82608*	*-0.15068*	*0.39887*

The simulation results in Table (3-1) including the parameter estimate of the Binary single index model (SIM)across four sample sizes$(25,50,100,150)$. It can be observed that the proposed method $(BBSI)$ performs better than other methods (BLO, BPR, BBQ)especially when the sample size getting bigger where the true value of $\beta_1 = 0.45$ and the parameter estimates getting closer to the true value as sample size getting larger, Look at $(\beta_1 = 0.41)$ at sample size (n=150)is close to the true value of $(\beta = 0.45)$ as the comequently ,sample size become more larger, the parameter estimates are close to true values with the proposed method.

Table(3-2) shows the value of estimated bias with different sample size and different method of SIM.

Table 3.2 Estimated Bias for estimated parameters in simulation example one using different methods under Binary SIM

SAMPLE SIZE	*METHODS*	*Bias 1*	*Bias 2*	*Bias 3*	*Bias 4*	*Bias 5*
N=25	*BBSI*	*0.07679*	*0.52553*	*0.38545*	*0.29309*	*0.28962*
	BLO	*3.75489*	*9.12443*	*7.37741*	*2.01792*	*3.48984*
	BPR	*1.58316*	*4.96661*	*3.79354*	*1.02767*	*1.51422*

	BBQ	*3.35731*	*9.09260*	*7.38383*	*3.05645*	*3.57842*
N=50	*BBSI*	*0.10014*	*0.60679*	*0.26033*	*0.42862*	*0.36808*
	BLO	*3.49688*	*2.48298*	*0.26717*	*0.06083*	*2.60920*
	BPR	*2.16730*	*1.78793*	*0.11072*	*0.04871*	*1.59950*
	BBQ	*4.09902*	*2.90343*	*0.33803*	*0.38413*	*2.60402*
N=100	*BBSI*	*0.07840*	*0.57555*	*0.45667*	*0.52216*	*0.35043*
	BLO	*0.66715*	*4.59534*	*0.44376*	*0.82082*	*0.76824*
	BPR	*0.56902*	*3.03683*	*0.23770*	*0.51648*	*0.41051*
	BBQ	*0.93141*	*5.61840*	*0.43861*	*1.09455*	*0.78586*
N=150	*BBSI*	*0.03362*	*0.44165*	*0.47039*	*0.25893*	*0.40063*
	BLO	*2.20622*	*4.21629*	*0.67097*	*0.22943*	*0.19854*
	BPR	*1.40390*	*2.81445*	*0.41431*	*0.13456*	*0.14653*
	BBQ	*2.77111*	*5.27952*	*0.82608*	*0.15068*	*0.39887*

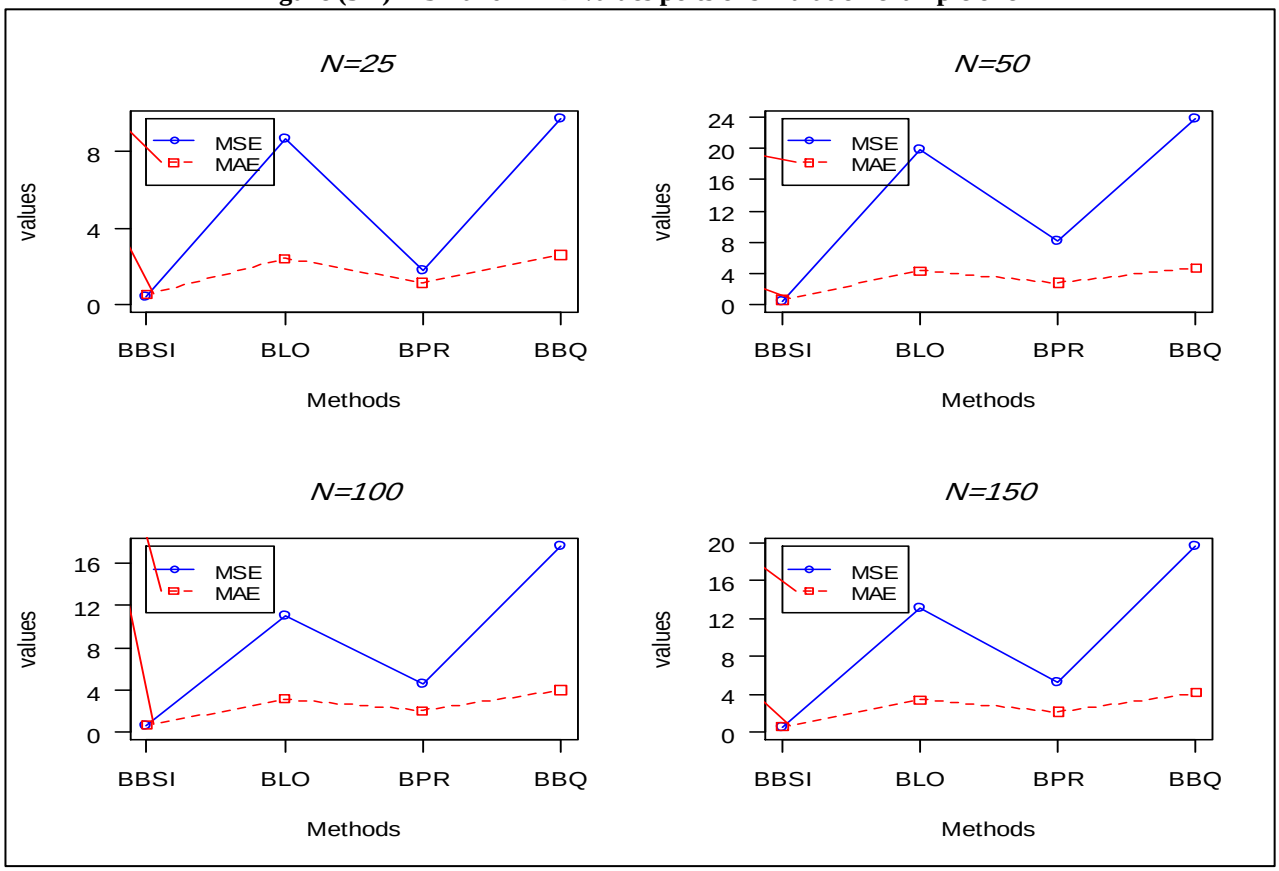
It can be observed from table $(3-2)$ that the obtained bias of our proposed method $(BBSI)$ is much smaller at different sample sizes than the competing methods (BLO, BPR, BBQ) for all the five parameter estimates .Also we can say that the prosed method $(BBSI)$ performs better than other methods .we can see that as the sample size become more larger, the proposed method $(BBSI)$ yields a lower bias values, that suggesting a good performance of $(BBSI)$ $method$.

Furthermore, we calculate the estimated values of the quality criteria ,MSE and MAD. To summary the values in table $(3-3)$ we draw the following figures

Table 3-3 MSE and MAD values of simulation example one in Binary SIM

SAMPLE SIZE	*METHODS*	*MSE*	*MAE*
N=25	*BBSI*	*0.38950*	*0.51204*
	BLO	*8.67103*	*2.39214*
	BPR	*1.78234*	*1.11809*
	BBQ	*9.74059*	*2.58822*
N=50	*BBSI*	*0.46511*	*0.56076*
	BLO	*19.93567*	*4.31102*
	BPR	*8.07966*	*2.75938*
	BBQ	*23.89781*	*4.69343*
N=100	*BBSI*	*0.68680*	*0.68029*
	BLO	*11.01314*	*3.16551*
	BPR	*4.52226*	*2.03575*
	BBQ	*17.56628*	*3.98976*
N=150	*BBSI*	*0.61648*	*0.61262*
	BLO	*13.01494*	*3.45931*
	BPR	*5.25438*	*2.20713*
	BBQ	*19.57248*	*4.20672*

Figure (3-1) MSE and MAD values polts of simulation example one



In Figure $(3-1)$the values of MSE and MAD Criteria are plotted against the methods$(BBSI, BLO, BPR, BBQ)$.It can be observed that the values of MAE criterion with different methods are better than the performance of MSE criterion .The closer the value to zero the better perform.

3.2. Example two:

In this example, the sample size we consider *n=25,50,100* and 150 observations are generated from the regression model:

$$y^* = \sin\left(\frac{\pi(z-\mathrm{A})}{\mathrm{B}-\mathrm{A}}\right) + 0.5\varepsilon\,, \qquad y_i = \begin{cases} 1 & if \quad y_i^* \geq 0 \\ 0 & otherwise \end{cases}$$

where $= \boldsymbol{x}'\boldsymbol{\beta}$, $\boldsymbol{x}$ is the design matrix $\boldsymbol{x} = (x_1, x_2, x_3, x_4, x_5)'$, $\boldsymbol{\beta}$ is the coefficient vector $\boldsymbol{\beta} = (1,1,0,0,1)/\sqrt{3}$, and the x_i $(i = 1,2,3,4,5)$ are i.i.d in a uniform distribution $[0,1]$[5]. A and B are constants that can be shown to be $(\frac{\sqrt{3}}{2} + \frac{1.645}{\sqrt{12}}, \frac{\sqrt{3}}{2} - \frac{1.645}{\sqrt{12}})$ respectively. ε is the error term, and we consider density distributions of the error term to evaluate the robustness of our proposed approach (Benoit et al., 2013).

$$\varepsilon \sim N(0,1),$$

In each level 10,000 iterations are run in the MCMC algorithm with 2,000 burn-in.

Table (3-4) shows a brief summing of the parameter estimates for simulation example two.

Table 3-4 the parameter estimates of simulation example two of Binary SIM

SAMPLE SIZE	*METHODS*	β_1	β_2	β_3	β_4	β_5
N=25	*BBSI*	*0.41886*	*0.33961*	*0.35700*	*0.40631*	*0.39127*
	BLO	*1.59318*	*0.63585*	*5.37369*	*-7.10044*	*-4.91924*
	BPR	*1.03810*	*0.57407*	*2.82198*	*-3.99000*	*-2.44667*
	BBQ	*0.29763*	*-0.23306*	*5.55800*	*-5.45617*	*-4.20753*

N=50	*BBSI*	*0.50271*	*0.35689*	*0.36681*	*0.43270*	*0.37720*
	BLO	*2.66201*	*3.18972*	*-2.01189*	*3.32282*	*-0.47479*
	BPR	*1.43134*	*1.54863*	*-0.75340*	*1.62274*	*-0.51376*
	BBQ	*2.95149*	*2.78922*	*-3.36864*	*3.34923*	*-0.54483*
N=100	*BBSI*	*0.37134*	*0.50746*	*0.24619*	*0.43430*	*0.45289*
	BLO	*1.70353*	*1.68806*	*1.25671*	*0.89153*	*1.34102*
	BPR	*0.72331*	*0.60207*	*0.61526*	*0.26394*	*0.41773*
	BBQ	*3.53048*	*3.37654*	*1.48863*	*1.26423*	*2.37674*
N=150	*BBSI*	*0.47928*	*0.41500*	*0.40537*	*0.32424*	*0.38892*
	BLO	*0.85489*	*1.67401*	*0.62735*	*0.30996*	*1.61626*
	BPR	*0.25907*	*0.78846*	*0.32277*	*0.16483*	*0.64066*
	BBQ	*1.93826*	*2.99887*	*0.50419*	*0.28165*	*3.08130*

Table $(3-4)$ including the parameter estimater of the Binary single index model (SIM)across four sample sizes $(25,50,100,150)$. Obiang that the proposed method $(BBSI)$performs better than other methods (BLO, BPR, BBQ) especially when the sample size getting bigger, In Table (3-5) the value of estimated bias for different sample size with different estimation method of SIM .

Table 3-5 Estimated Bias for simulation example one using different methods

SAMPLE SIZE	METHODS	Bias 1	Bias 2	Bias 3	Bias 4	Bias 5
N=25	BBSI	0.15849	0.23774	0.35700	0.40631	0.18608
	BLO	1.01583	0.05850	5.37369	7.10044	5.49659
	BPR	0.46075	0.00328	2.82198	3.99000	3.02402
	BBQ	0.27972	0.81041	5.55800	5.45617	4.78488
N=50	BBSI	0.07464	0.22046	0.36681	0.43270	0.20015
	BLO	2.08466	2.61236	2.01189	3.32282	1.05214
	BPR	0.85399	0.97128	0.75340	1.62274	1.09111
	BBQ	2.37414	2.21187	3.36864	3.34923	1.12218
N=100	BBSI	0.20601	0.06989	0.24619	0.43430	0.12446
	BLO	1.12617	1.11071	1.25671	0.89153	0.76367
	BPR	0.14596	0.02472	0.61526	0.26394	0.15962
	BBQ	2.95313	2.79919	1.48863	1.26423	1.79939
N=150	BBSI	0.09807	0.16235	0.40537	0.32424	0.18843
	BLO	0.27754	1.09666	0.62735	0.30996	1.03891
	BPR	0.31828	0.21110	0.32277	0.16483	0.06331
	BBQ	1.36091	2.42152	0.50419	0.28165	2.50395

It can be observed from table $(3-5)$ that the bias values from our proposed method ($BBSI$) are the smallest values under different sample size than the competing methods (BLO, BPR, BBQ) for all the five parameter estimates. Also, we can say that the proposed method $(BBSI)$ performs better than other methods, more precisely we can see that as the sample size become larger, the proposed method$(BBSI)$ yields loosest bias values, consequently that suggesting a good performance of $(BBSI)$ method.

Next ,we calculate the estimated values of the quality criteria of the parameters ,MSE and MAD. Table ($3-6$) shows the values of the MSE and MAD Criteria.

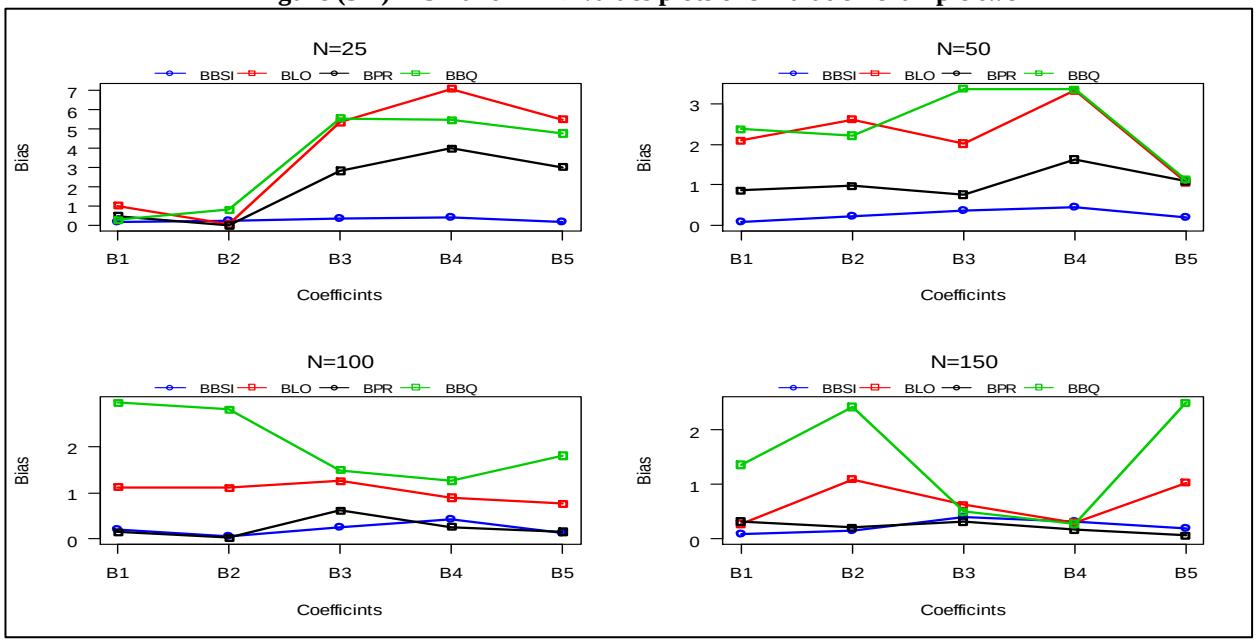
Table 3-6 MSE and MAD for simulation example two Binary SIM

SAMPLE SIZE	METHODS	MSE	MAE
N=25	BBSI	0.21454	0.33259
	BLO	14.83113	3.15272
	BPR	4.86515	1.80305
	BBQ	12.78830	2.93271
N=50	BBSI	0.14627	0.24671
	BLO	8.65291	2.59839
	BPR	1.30228	0.96521
	BBQ	5.51465	1.99139
N=100	BBSI	0.16588	0.26206
	BLO	7.07242	2.48794
	BPR	0.36779	0.45853
	BBQ	28.18380	5.03208
N=150	BBSI	0.20045	0.30176
	BLO	3.36333	1.66280
	BPR	0.24617	0.34495
	BBQ	13.81898	3.43550

The results of MSE and MAD that listed in table $(3-6)$ indicates that the proposed method $(BBSI)$ generally performs better than (BLO, BPR, BBQ) methods overall the different sample sizes. However, we observed that$(BBSI)$ method tends to behove befor than other methods in terms of the values of MSE and MAD Criteria.

Next , we obtained the estimated values of the , MSE and MAD criteria . To summary the values of MSE and MAD criteria, draw the following figures.

Figure (3-2) MSE and MAD values plots of simulation example two



In Figure $(3-2)$the values of MSE and MAD Criteria are plotted along with the methods $(BBSI, BLO, BPR, BBQ)$. It can be observed that the values of MAE criterion with different methods are better than the performance of MSE criterion. The closer value to zero the better performs.

4. Real Data Analysis:

The simulation study has demonstrated that the proposed method has proven its ability to compute with the existing methods, which encouraged the use to find red data to test the performance of the proposed method. In this part of the chapter ,we applied the proposed method and the other existing methods to read data and then summarized the results to analyze it Identification and diagnosing of the real reasons behind the change in the condition of the person infected with COVID-19 is necessary to help the medical staff recover the infected person .Hence , we will employ the propped method to identify the most related predictor variable that effects the response variable (severity of covid-19) the data used in this part is taken from. In this study the dependent variable is binary; it either takes 'zero' in case of major infection or death, or 'one' in case of moderate or minor infection. The independent variables are 14 variables which represent the factors influencing the infection of Coronavirus.

X1: Represents gender, male = 1, female = 2
X2: Represents age
X3: represents the weight
X4: represents pressure, None = 1,Found = 2, Decrease = 3, Medium = 4, Height = 5
X5: represents diabetes, None=1,Found=2,Descending=3 ,Ascending=4
X6: Represents lung problems ,None = 1, found = 2
X7: Represents a weak immune system, None = 1, There is = 2
X8: Represents vitamin D deficiency, None = 1, Fond =2
X9: represents the workplace ,Housewife or not working = 1, Employee =
2, Wage earner = 3 ,Students = 4, Hospital and medical clinics = 5
X10: Represents previous surgical operations, No operations = 1,
Previous operations performed = 2,
X11: represents smoking, Non-smoker = 1, Smoker = 2
X12: Represents the psychological state, Not good = 1, Medium = 2,
Good =3
X13: represents nutrition, Not good = 1, Medium = 2, Good = 3
X14: living status, Poor = 1, Medium = 2, Good or Rich = 3

Some people may experience worsening symptoms that can lead to death Therefore, the researcher tried to shed light on the reasons that lead to the major cases of infections and the minor ones as well. The data was collected by questioning people via Google Forms to measure the impact of the virus on them and the major influencing factors that lead to the infection. The data represent a sample group of 130 infected persons for the summer of 2020. The sample was taken from people in the city of Al-Diwaniyah within four months by using a form.

4.1-Real Data Results:

The following tables summaries the results that have been obtained after we employed the proposed method and the other method based on real data. Table (4-1) shows the parameter estimates of the binary SIM .

Table 4-1 parameter estimates of different method

Methods	*BBSI*	*BLO*	*BPR*	*BBQ*
B1	*0.29014*	*0.31998*	*0.18506*	*0.40531*
B2	*0.04263*	*0.12657*	*0.07744*	*0.16299*
B3	*0.25911*	*-0.20632*	*-0.11840*	*-0.29304*
B4	*0.38520*	*0.05389*	*0.02654*	*0.06903*
B5	*0.33637*	*-0.15886*	*-0.10208*	*-0.19213*
B6	*0.34137*	*-0.07436*	*-0.04022*	*-0.08723*
B7	*0.15425*	*-0.27680*	*-0.15812*	*-0.35888*
B8	*0.44312*	*0.01050*	*0.01533*	*0.03763*
B9	*0.29485*	*-0.13840*	*-0.07825*	*-0.22383*
B10	*0.06082*	*-0.25585*	*-0.14421*	*-0.33602*

B11	*0.32639*	*-0.06998*	*-0.05657*	*-0.12639*
B12	*0.41848*	*-0.23281*	*-0.14321*	*-0.37000*
B13	*0.28595*	*0.09066*	*0.06703*	*0.11334*
B14	*0.26776*	*0.43562*	*0.24214*	*0.69945*

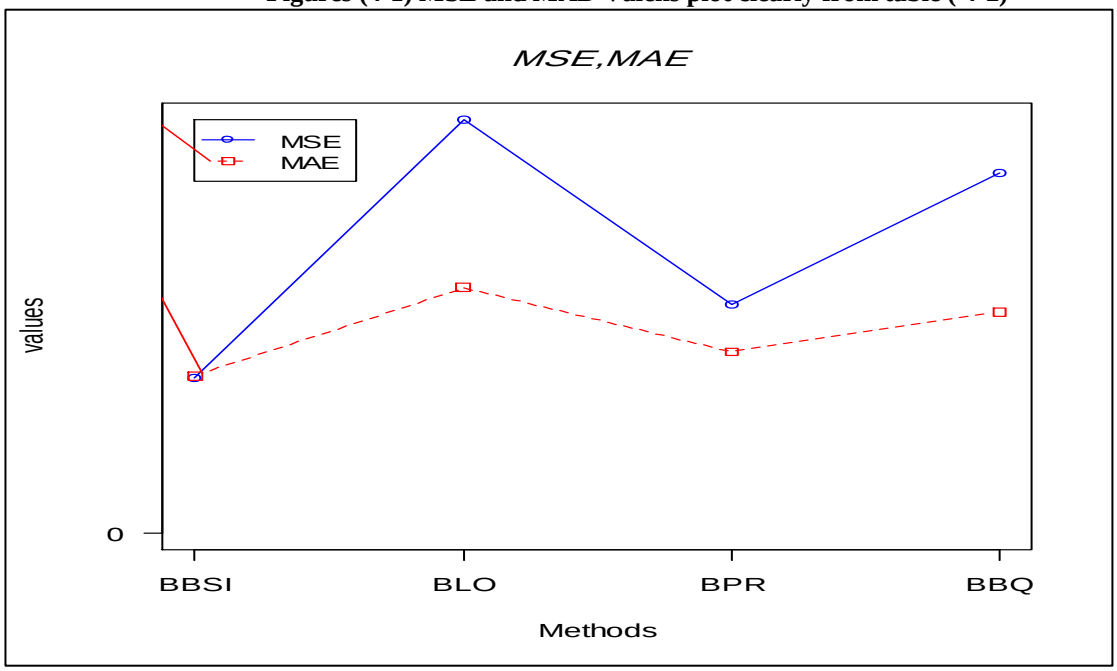
It can be observed from table (4-1) that the most related predictor variable which affects the response variable according to the proposed method $(BBSI)$ is X_8 (vitamin D deficiency) with
(β=0.443),While the variable X_{12} (psychological state) is the second most related pred But variable x_{10} chor variable with (β_{12}=0.418) (previous surgical operator) is the less related predictor variable in response variable with (β_{10}=0.060) Followed by variable x_2 (age)
with (β=0.04).The results of(BLO, BPR, BBQ) are closer to each other , for example , we can see at the variable X_3 (weight) have negative sign which indicates logic result.
Also, the other method indicates y-hat the variables (X_5,X_6,X_7,X_9,X_{10},X_{11},X_{12}) are negatively effects on the response variable which is logic results .Furthermore, the variable X_{13} has fewer effects or response variables.
Next table$(4-2)$ shows the values of MSD and MAD for the real data results.

Table 4-2 shows the values of MSE and MAD for the real data results.

Methods	*MSE*	*MAE*
BBSI	*0.70552*	*0.71607*
BLO	*1.88586*	*1.12075*
BPR	*1.04015*	*0.82519*
BBQ	*1.64210*	*1.00833*

Obviously, from table (4-2) the proposed method ($BBSI$) gives the less values of both MSE and MAD criteria (MSE=0,705 ,MAD=0.71) that indicates the high efficiency of the proposed method in terms of MSE less variance and less bias for estimated parameters . the of MSE and MAD that obtained by BBSI method are the less (closer to zero) compared to the other method.

Figures (4-1) MSE and MAD Valens plot clearly from table (4-2)



In the next figure $(4-1)$, we plotted values of MSE and MAD criterim that illustrated in table$(4-2)$. To summary values of MSE and MAD draw the following figures.

4. Conclusions:

Based on both theoretical, simulation and applied data, we write the following conclusions:

1. A simulation study was conducted for each of the binary chips.
2. We compared the proposed method (BBSI) with other methods (BLO, BPR, BBQ).
3. Real data analysis was used based on the proposed dual SIM methods
4. Based on (2), we find that the (BBSI) method is better than other methods in terms of MSE and MAE values.

References

1. Antoniadis, A., Gr_egorie, G., and McKeague, I. (2004). \Bayesian Estimation of Single-Index Models." Statistica Sinica, 14, 1147{1164}.

2.Andrews DF, Mallows CL (1974) Scale mixtures of normal distributions. Journal of the Royal Statistical Society, Series B 36:99{102} Powell J. 1986. Censored regression quantiles. *Journal of Econometrics* **32**(1): 134–155.

3. Bellman, RE Kalaba, JA Lockett - 1966 - apps.dtic.mil

4.Davis, R. A., P. Zang, and T. Zheng (2016). Sparse vector autoregressive modeling. Journal of Computational and Graphical Statistics 25(4), 1077–1096. Efromovich, S. (2005). Estimation of the density of regression errors. The Annals of Statistics 33(5), 2194–2227.

5.Development Core Team (2009). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

5. Hardle W, Hall P, Ichimura H (1993) Optimal smoothing in single-index models. Ann Stat 21:157–178.

6.Ishimura, H.(1993)."Semi pramaetric least squares (SLS) and weighted SLS estimation of single index models" Journal of econometrics.