**Periodicals of Engineering and Natural Sciences** ISSN 2303-4521
Vol. 9, No. 1, February 2021, pp.285-292

# New penalized Bayesian adaptive lasso binary regression

**Ameer Musa Imran Alhseeni , Ali Abdulmohsin Abdulraeem Al-rubaye**
Planning Department, General Directorate of Qadisiyah Education, Ministry of Education

**Abstract**

The scale mixture of normal mixing with Rayleigh as representation of Laplace prior of $\beta$ has introduced by Flaih et al. [1]. We employed this new scale mixture for the adaptive lasso Binary regression. New hierarchical model is considering, as well the Gibbs sampler algorithm in introduced. We considering the new penalized Bayesian adaptive lasso in Binary regression as variable selection method in case of presenting they high dimensional data. The new proposed model can overcome the multicollinearity problem in predictor variables. We conducting simulation analysis, as well as real data application to show the performance of the proposed method.

**Keywords**: Bayesian binary regression, adaptive lasso, Gibbs sampling, Variable selection, normal-Rayleigh scale mixture.

***Corresponding Author:***
Ameer Musa Imran Alhseeni
Planning Department, General Directorate of Qadisiyah Education, Ministry of Education
AL-Qadisiyah, Iraq
Email: amer.alhussani@gmail.com

## 1. Introduction

Recently the applications of binary regression model has been widely common , the binary regression is one of the most well know models that estimate the conditional mean function $E\left((Y|X)\right)$. This paper discusses the binary regression model that proposed by David [2] .which is explain the dependent (response) variable $y$ as a dichotomous or dummy variable ,that is means ,we have binary response variable .When the response variable has exactly two values ($y = 0$ $or$ $y = 1$), then we often speak of binary model. We employed the binary regression model analysis through Bayesian point of view. Katerina et al [3] proposed the binary quintile regression and variable selection problem by using reweighted least squares method. The main objective of this paper is to how deal with the variable selection problem easily in Bayesian analysis for the Binary regression model and how the regularization method (adaptive lasso) proves the performance of Bayesian techniques like Gibbs sampler algorithm in prediction accuracy for the binary regression [4]. regression choice is the big deal to be aware about in the analysis of data. Choosing the appropriate regression model gives more interpretability and more efficient estimates and leads meaning full point estimation in terms of bias and variances of estimators. So consequently, we can say that the Binary regression is suitable for the binary response variable [5]. Dries and Poel [6] develop a Bayesian Gibbs sampler algorithm for the Binary quantile regression model and defined the standard Binary regression model as the following simple measurement equation

$$y_i = \begin{cases} 1 \ if \quad y^* = X_i^t\beta + e_i \geq 0 \\ 0 \ if \quad y^* = X_i^t\beta + e_i < 0 \end{cases} \tag{1.1}$$

Model (1.1) is the widely commonly used form for the Binary response regression model. That assume the range of $y^* \in (-\infty, \infty)$ is an observable response variable.

From form (1.1) $y_i$ is the observable variable can be defined as an indicator of the $i^{th}$- response individual ,and its values determined by the unobservable (latent) variable $y^*$ ,$x_i = (x_i, \dots, x_p)$ of explanatory variables , $\beta^t = (\beta_1, \dots \beta_p)$ of coefficients, and $e_i$ is the term of random error $i = 1, \dots, n$. Here $p$ is the number of predictors and $n$ is the sample size of observation. Katerina et al. in [3] assumed that if $e_i \sim N(0, \sigma^2)$, then the binary probit analysis model arises.

In this paper we derive the binary regression model (BRM) by considering that the latent variable (unobserved) can be formulated on the measurement model relating to unobserved variable ($y_i^*$) to the observed ($y_i$) variable, binary outcome [7].

Regression analysis involve examine the interpretability of the model through the correct functional form which the irrelevant explanatory variable that have smaller effect on the response variable must be removed ,because the large number of explanatory variables in the model may be hard to interpret, consequently overfitting can produced invalidate the model by involving too many irrelevant explanatory variables which are not effect on the response variable, therefore, variable selection procedures are very helpful to remove the irrelevant explanatory variables.

Robert [8] produced the lasso and Bayes lasso regression models that can be considering as variable selection method, which give zero for some irrelevant predictor variables and then removed it from the regression model, this new penalized method produced more interpretable and more prediction accuracy model. In 2008 Park and Casella [9] proposed the Bayesian concept for the lasso method through defining the prior distribution of regression parameters as scale mixture of normal mixing with exponential distribution on their variances. Himel and Yi [10] proposed new scale mixture for the prior distribution, which are uniforms mixing with $Gamma(2, \lambda)$, Abbas and Alhamzawi in [5] produced adaptive lasso Binary regression by considering that the prior distribution of $\beta$ is a scale mixture of uniforms mixing with standard exponential.

As seen in [1], introducing the Bayesian lasso and adaptive lasso based on considering that the prior distribution of $\beta$ is normal mixing with Rayleigh distribution, and also work motivated us to utilize the proposed scale mixture in Binary response regression model (1.1).

## 2. Bayesian adaptive lasso binary response regression (BALBRReg)

Zou [11] introduced the frequentist adaptive lasso and defined the adaptive lasso estimators as follows,

$$\hat{\beta}_{Adaptivelasso} = {}^{argmin}_{\beta} = \|Y - X^t\beta\|^2 + \lambda \sum_{j=1}^{p} w_j|\beta_j|, \tag{1.2}$$

where $\lambda \sum_{j=1}^{p} w_j|\beta_j|$ is the penalty function with $\lambda \geq 0$ , the shrinkage parameter and $w_j$ are the weights ($w_j > 0$).

Minh-Ngoc and Nott [12], standard the minimization problem (1.2) in Bayesian point of view and considered the Bayesian regularization adaptive lasso estimator as follows,

$$\hat{\beta}_{BL} = {}^{argmin}_{\beta} = \|Y - X^t\beta\|^2 + \sum_{j=1}^{p} \lambda_j|\beta_j|, \tag{1.3}$$

The minimization problem (1.3) gives different shrinkage, parameter $\lambda_j$ for each $\beta$.

Flaih et a. in [1]. proposed new scale mixture that motivated us to study new Bayesian parameter analysis for the Binary response regression model, their proposed scale mixture takes the following representation form.

$$\pi(\beta|\lambda, \sigma^2) = \frac{\lambda}{2\sigma^2} exp\left[-\frac{\lambda|\beta|}{\sigma^2}\right] = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2\varepsilon}} e^{-\frac{\beta^2}{2\sigma^2\varepsilon}} \frac{\lambda}{2} e^{\frac{-\lambda\varepsilon}{2}} \, d\varepsilon \tag{1.4}$$

The prior $\boldsymbol{\pi(\beta|\lambda, \sigma^2)}$ in (1.4) conditioning on $\boldsymbol{\sigma^2}$ to generate the unimodality of the posterior distribution of $\boldsymbol{\beta}$ , for more details see in [9]. The presentation form **(1.4)** proved that it is a comparative formula in producing stationary Gibbs sampler algorithm.

### 2.1. The BALBRR hierarchical model

The hierarchical model for the Bayesian adaptive lasso Binary response regression model is considering based on the scale mixture form **(1.4)** and the measurement model **(1.1)** and defined as follows,

$$y_i = \begin{cases} 1 \;\, if \;\; y^* = X_i^t\beta + e_i \geq 0 \\ 0 \; if \;\; y^* = X_i^t\beta + e_i < 0 \end{cases}, \tag{2}$$

$$y^*|x,\beta,\sigma^2 \sim N_n(x^t\beta,\sigma^2 I_n)\ , \quad (3)$$
$$\beta\big|\sigma^2,\varepsilon_1,\varepsilon_{2,},\dots\varepsilon_p \sim N_p(0,\sigma^2 v_\varepsilon), \quad (4)$$
$$v_\varepsilon = diag\ (\varepsilon_1,\varepsilon_{2,},\dots\varepsilon_p), \quad (5)$$
$$\varepsilon_1,\varepsilon_{2,},\dots\varepsilon_p \sim \pi_{j=1}^{p} = \frac{\lambda}{2}e^{-\frac{\lambda\varepsilon_j}{2}}\ d\varepsilon_j\ , \quad (6)$$
$$\sigma^2 \sim Inverse\ Gamma$$
$$\lambda \sim Gamma\ (a,b)$$

### 2.2. The full conditional posterior distribution for the BALBRR model

The Gibbs sampling algorithm is implementing with the following full conditional posterior distribution of BALBRR model.

1- The full conditional posterior distribution of the latent response variable $\boldsymbol{y}^*$ is defined of follows.

$$y^* = \begin{cases} y = 1\ if\ \ \sim N_n(X_i^t\beta,\sigma^2 I_n)\pi_{i=1}^{n} \quad I\{y_i > 0\} \\ y = 0\ if\ \ \sim N_n(X_i^t\beta,\sigma^2 I_n)\pi_{i=1}^{n} \quad I\{y_i > 0\} \end{cases} \quad (7)$$

2- The full conditional posterior distribution $\boldsymbol{\beta}$ is

$$\pi(\beta,\lambda,\sigma^2|y^*,x)\alpha\pi(y^*|x,\beta,\sigma^2,\lambda)\pi(\beta|\sigma^2) \quad (8)$$
$$\propto -\frac{1}{2\sigma^2}[(y-X\beta)^t(y-X\beta)] - \frac{1}{2\sigma^2}\beta^t V_c\beta \quad (9)$$
$$\propto -\frac{1}{2\sigma^2}[(\beta - C^{-1}X^t y)^t C(\beta - C^{-1}X^t y)], \quad (10)$$

where is the multivariate normal distribution with mean $\boldsymbol{C^{-1}X^t y}$ and variance $\boldsymbol{\sigma^2 C^{-1}}$ , for more details See In [1].

3- The Full conditional posterior distribution for $\sigma^2$ is the following inverse Gamma distribution

$$\pi(\sigma^2|y^*,x,\beta)\alpha\pi(y^*|x,\beta,\sigma^2)\pi(\sigma^2) \quad (11)$$
$$\propto (\sigma^2)^{\frac{n}{2}} e^{-\frac{1}{2\sigma^2}(y-X\beta)^t(y-X\beta)} \quad (\sigma^2)^{-\alpha-1}\ \ e^{-\frac{r}{\sigma^2}} \quad (12)$$
$$\propto\ (\sigma^2)^{\frac{-p}{2}}\ \ e^{-\frac{1}{2\sigma^2}\beta V_\varepsilon^{-1}\beta} \quad (13)$$
$$\propto (\sigma^2)^{-\frac{-n}{2}-\frac{P}{2}-(\alpha+1)}\ \exp\left[-\left\{\frac{1}{2\sigma^2}(y-X\beta)^t(y-X\beta) + \frac{1}{2\sigma^2}\beta V_\varepsilon^{-1}\beta + \frac{r}{\sigma^2}\right\}\right] \quad (14)$$

Which is the inverse gamma with shape parameter $\boldsymbol{\frac{n}{2} + \frac{P}{2} + \alpha}$ and scale parameter $\boldsymbol{\frac{(y-X\beta)^t(y-X\beta)}{2}} + \boldsymbol{\beta V_\varepsilon^{-1}\beta/(2+r)}$.

4- The full conditional posterior distribution of $\varepsilon$ is inverse Gaussian distribution

$$\pi(\boldsymbol{\varepsilon}|\beta,\sigma^2,\lambda)\alpha\pi(\beta|\boldsymbol{\varepsilon},\sigma^2,\lambda)\pi(\boldsymbol{\varepsilon}) \quad (15)$$
$$\propto (\boldsymbol{\varepsilon_j})^{-\frac{3}{2}}\exp\left(\frac{-\beta_j^2}{2\sigma^2\boldsymbol{\varepsilon_j}}\right)\left(\frac{\lambda\boldsymbol{\varepsilon_j}}{2}\right) \quad (16)$$
$$\propto (\boldsymbol{\varepsilon_j})^{-\frac{3}{2}}\exp\left(\frac{\beta_j^2(\frac{1}{\varepsilon_j} - \sqrt{\lambda\sigma^2/\beta_j^2})}{2\sigma^2(1/\boldsymbol{\varepsilon_j})}\right) \quad (17)$$

Which is inverse Gamma $\left(\sqrt{\frac{\lambda\sigma^2}{\beta_j^2}},\lambda\right)$

5- The full conditional posterior distribution of $\lambda$ is Gamma

$$\pi(\boldsymbol{\varepsilon_j}|\lambda)\alpha\pi(\ \lambda_j|\boldsymbol{\varepsilon_j})\pi(\lambda_j) \quad (18)$$
$$\propto \pi_{j=1}^{p}\frac{\lambda}{2}\ e^{-\frac{\lambda}{2}\varepsilon_j}\ \ (\lambda)^{a-1}e^{-b\lambda} \quad (19)$$

$$\propto \lambda^{p+a-1}\, exp\left[-\lambda(\tfrac{1}{2}\textstyle\sum_{j=1}^{p}\boldsymbol{\varepsilon_j}+1)\right] \tag{20}$$

Which is Gamma $(\mathbf{p+a}\ ,\frac{1}{2}\sum_{j=1}^{p}\boldsymbol{\varepsilon_j}+\mathbf{1})$

By following the hierarchical model and the posterior distributions of the interested parameters, we created an efficient simple fast Gibbs sampler algorithm for finding the interested parameters estimation. For comparting our proposed models performing, we used different exists estimation methods.

## 3. Simulation analysis

In this section, we employed the simulation techniques for comparing the proposed model with the Bayesian binary regression (BBReg) that introduced by Martin Quinn, park and Bayesian lasso Binary quantile regression (BLBQReg) that proposed by D,f. Benot et al In[6], the simulation study conducted based on three samples sizes , low sample size n=15 ,medium sample size n=75, and large sample size n=150. For each sample size, we assumed that the error form distributes as the following different distributions:

1- $\epsilon_i \sim N(0,1)$ , The standard normal distribution,
2- $\epsilon_i \sim N(2,2) + N(3,3)$ , The mixed normal distributions,
3- $\epsilon_i \sim Lap(0,1)$ , The standard Laplace distribution,
4- $\epsilon_i \sim Lap(1,1) + Lap(1,1)$ , The mix Laplace distribution,
5- $\epsilon_i \sim t(5)$ , The t-student distribution with degree freedom 5,

We draw 11000 iterations from the Gibbs sampler algorithm, the first 1000 iterations have been burned in. The mean absolute error (MAE) criterion has used, as well the standard deviation (SD) criterion for comparing between the proposed model and other estimations methods;

$$MAE = \frac{\sum_{i=1}^{p}|X\hat{\beta}-X\beta^{true}|}{p} \quad , \quad SD = \frac{\sum_{i=1}^{p}(X\hat{\beta}-X\beta^{true})^2}{p} \tag{21}$$

1-Simulation case 1 (very sparsity model)

We simulate a random data with seven independent random variables ($p = 7$), each with 100 observations (n=100), the true regression coefficients vector is ,

$$\beta^{true} = (1, 0^6)$$

That is

$$y_i = \sum_{i=1}^{7} X_i\, \beta_i + \epsilon_i \tag{22}$$

The matrix of the predictors variables (observations) $X \sim N(0, \Sigma)$, under different types of error distributions. The following table show the values of MAE and SD for the different scenarios.

Table 1. The MAE and SD values under different seniors, for very sparse model

| | Methods | | $e_i \sim N(0,1))$ | $e_i \sim N(2,2) + N(3,3)$ | $e_i \sim t_{(5)}$ | $e_i \sim Lap(0,1)$ | $Lap(1,1) + Lap(2,2)$ |
|---|---|---|---|---|---|---|---|
| Sim1 | n=15 | BLBQReg | (0.652)(0.442) | (0.784)(0.548) | (0.834)(0.672) | (0.873)(0.508) | (1.231)(1.002) |
| | | BBReg | (0.385)(0.856) | (0.756)(0.955) | (0.568)(0.806) | (0.639)(0.845) | (1.545)(1.305) |
| | | BALBRR eg | (0.723)(0.539) | (0.634)(0.643) | (0.567)(0.743) | (0.683)(0.430) | (1.423)(1.357) |
| | n=75 | BLBQR eg | (0.304)(0.450) | (0.473)(0.505) | (0.530)(0.634) | (0.438)(0.563) | (1.342)(1.211) |
| | | BBReg | (0.532)(0.410) | (0.773)(0.572) | (0.837)(0.655) | (0.765)(0.574) | (1.214)(0.993) |
| | | BALBRR eg | (0.682)(0.451) | (0.845)(0.530) | (0.563)(0.490) | (0.453)(0.404) | (1.340)(1.312) |
| | n=150 | BLBQReg | (0.341)(0.238) | (0.201)(0.263) | (0.456)(0.452) | (0.397)(0.379) | (0.892)(0.859) |
| | | BBReg | (0.324)(0.301) | (0.222)(0.267) | (0.456)(0.344) | (0.305)(0.403) | (0.985)(0.982) |
| | | BALBRReg | (0.242)(0.201) | (0.106)(0.196) | (0.378)(0.435) | (0.388)(0.394) | (0.845)(0.923) |

Table 1, shows the values of MAE and SD at different , sample size and different error distributions for the proposed method (BALBRReg),the (BLBQReg) and the (BBReg), we can conclude that our proposed method is comparable with the different methods and performs better ,based on the values MAE and SD ,where are less than the values of the different methods .that is the proposed method gives more prediction accuracy and variable selection comparing with the different methods, especially at n=150 .

2- Simulation case 2 (very density)

Following the same procedure in simulation example 1, but here we consider the true vector of regression parameter is

$\beta^{true} = (0.85)^t$ , which is the very density case, we simulate $\boldsymbol{y_i}$ as follows

$$y_i = \sum_{i=1}^{7} 0.85X_i + \epsilon_i \quad (23)$$

We generate $X \sim N(0, \Sigma)$, with n=100 under different types of error terms densities. The following table shows the values of MAE and SD criterions for the very density case.

Table 2. The MAE and SD values for very density under different scenarios.

| | Methods | | $e_i \sim N(0,1))$ | $e_i \sim N(2,2) + N(3,3)$ | $e_i \sim t_{(5)}$ | $e_i \sim Lap(0,1)$ | $e_i \sim Lap(1,1) + Lap(2,2)$ |
|---|---|---|---|---|---|---|---|
| Sim2 | n=15 | BLBQReg | (0.735)(0.46) | (0.654)(0.403) | (0.672)(0.642) | (0.738)(0.673) | (1.364)(1.227) |
| | | BBReg | (0.503)(0.742) | (0.563)(0.458) | (0.623)(0.646) | (0.582)(0.634) | (1.365)(1.267) |
| | | BALBRReg | (0.637)(0.472) | (0.465)(0.534) | (0.729)(0.638) | (0.740)(0.668) | (1.364)(1.243) |
| | n=75 | BLBQReg | (0.436)(0.516) | (0.734)(0.450) | (0.694)(0.605) | (0.539)(0.662) | (1.221)(1.470) |
| | | BBReg | (0.653)(0.536) | (0.431)(0.546) | (0.693)(0.730) | (0.638)(0.473) | (1.340)(1.442) |
| | | BLBQReg | (0.483)(0.538) | (0.745)(0.342) | (0.634)(0.545) | (0.404)(0.432) | (1.365)(1.403) |
| | n=150 | BALBRReg | (0.386)(0.347) | (0.322)(0.301) | (0.430)(0.443) | (0.263)(0.373) | (0.648)(0.734) |
| | | BBReg | (0.423)(0.175) | (0.201)(0.243) | (0.322)(0.240) | (0.293)(0.306) | (0.553)(0.610) |
| | | BALBRReg | (0.123)(0.432) | (0.152)(0.111) | (0.278)(0.135) | (0.225)(0.205) | (0.543)(0.453) |

By looking at the values of MAE and SD for the different sample sizes and the different types of errors densities, we can see that our performance of our proposed model BALBRReg comparing with BLBQReg and BBReg, where the less values of MAE and SD are in our proposed model, which indicates the more estimation accuracy and the best variable selection in the view of the true vectors of parameters estimates.

3- Simulation case 3 (sparse case)

The true vector of parameters estimates is

$\beta^{true} = (5,0,0,0,5,0,5)^t$ and the regression model is

$$y_i = 5X_{i1} + 5X_{i5} + 5X_{i7} + e_i, \quad (24)$$

The data generated from $X \sim N(0, \Sigma)$, with 100 observations with five different types of errors as describes in the previous examples. The following table shows the values of MAE and SD criterions.

Table 3. MAE and SD values for sparse model under different sample size and different error models

| | Methods | | $e_i \sim N(0,1))$ | $e_i \sim N(2,2) + N(3,3)$ | $e_i \sim t_{(5)}$ | $e_i \sim Lap(0,1)$ | $e_i \sim Lap(1,1) + Lap(2,2)$ |
|---|---|---|---|---|---|---|---|
| Sim3 | n=15 | BLBQReg | (0.845)(0.534) | (0.753)(0.693) | (0.725)(0.649) | (0.631)(0.773) | (1.648)(1.548) |
| | | BBReg | (0.634)(0.873) | (0.682)(0.563) | (0.660)(0.742) | (0.623)(0.730) | (1.403)(1.341) |
| | | BALBRReg | (0.734)(0.504) | (0.502)(0.528) | (0.845)(0.720) | (0.602)(0.521) | (1.463)(1.372) |
| | n=75 | BLBQReg | (0.734)(0.435) | (0.532)(0.536) | (0.644)(0.705) | (0.682)(0.724) | (1.734)(1.636) |
| | | BBReg | (0.564)(0.534) | (0.463)(0.605) | (0.745)(0.632) | (0.535)(0.401) | (1.530)(1.356) |
| | | BALBRReg | (0.435)(0.354) | (0.620)(0.546) | (0.565)(0.627) | (0.496)(0.567) | (1.648)(1.592) |
| | n=150 | BLBQReg | (0.212)(0.241) | (0.367)(0.401) | (0.511)(0.500) | (0.395)(0.400) | (0.730)(0.846) |
| | | BBReg | (0.203)(0.116) | (0.206)(0.374) | (0.293)(0.362) | (0.362)(0.356) | (0.486)(0.536) |
| | | BALBRReg | (0.211)(0.231) | (0.200)(0.220) | (0.201)(0.200) | (0.377)(0.536) | (0.374)(0.342) |

From the table 3, the values of MAE and SD for the proposed model are less than the other different models, which indicates the out performance of the BALBRReg over the BLBQRReg and BBReg models.

Consequently, the estimation accuracy and variable selection with the proposed model is more reliable in comparing with other models.

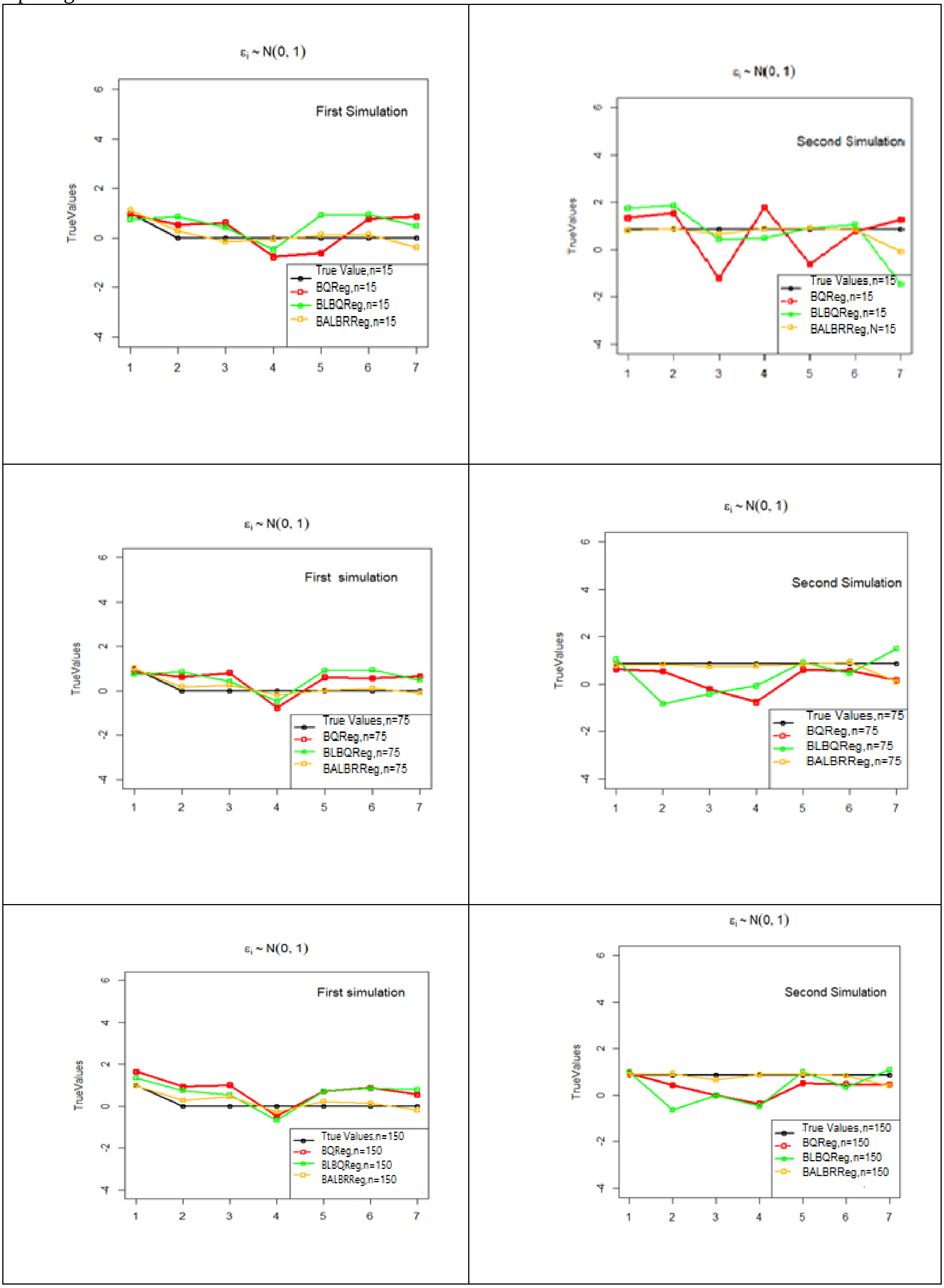


Figure 1. The closed of estimation values with true value

The following figure summarized the simulation examples results, where in the first and second examples when the error term follows the standard normal distribution, our proposed model shows the closed parameters estimates for the true vector which indicates the outperformance of the proposed model.

## 4. Real data analysis

For the application proposes, we select the general hospital in Hala, where data have collected for 200 patients from the historical record. the response variables were blood sugar level and the predictors are as follows: age ($X_1$), weight ($X_2$), genetic factors ($X_3$), number of meals during the day ($X_4$), Dose the patient use diabetic? ($X_5$), is the patient suffering from a disorder of the pituitary glands? ($X_6$), number of exercise hours ($X_7$), internet hours ($X_8$), marital status ($X_9$), residence place ($X_{10}$), monthly income ($X_{11}$), is the patient exposed to psychological pressure? ($X_{12}$), is the patient undergoing mental stress? ( $X_{13}$), number of soft and energy drinks during the day ($X_{14}$), injury or surgery history (yes or no) ($X_{15}$).

After collecting the data for 200 patients, we employed the proposed method and compared the results with the same method that used in simulation analysis, the mean square error (MSE) has used to judge the performance of the different methods. The following table explain the estimates of the MSE as well the standard error

Table 4. The values of MSE as S.E

| Methods | BLBQReg | BBReg | BALBRReg |
|---|---|---|---|
| MSE | 243 | 174 | 102 |
| S.E | 0.885 | 0.632 | 0.379 |

Table 4 shows that the proposed model (BALBRReg) have the minimum estimates of the MSE and S.D. comparing with (BLBQReg) and (BBReg), that is the outperformance of the proposed model in point view of estimation accuracy.

## 5. Conclusions

In this paper we have presented new regularization technique in point view of Bayesian for the adaptive lasso Binary regression model based on the scale mixture of normal mixing with Rayleigh distribution as representation of the Laplace prior distribution of the parameter $\boldsymbol{\beta}$. New hierarchical model and prior distributions, as well as the full conditional posterior distributions have been developed for implementing an efficient and easy Gibbs sampler algorithm. The proposed model, Bayesian Adaptive Lasso Binary Response Regression (BALBRReg) displayed the outperformance in the point view of variable selection and the prediction accuracy comparing with Bayesian binary regression (BBReg), and Bayesian lasso Binary quantile regression (BLBQReg) based on the simulation examples and the real data application.

## References


[1] A.N. Flaih, and T. Alshaybawee, A. F. Hadi, "Bayesian extensions on Lasso and adaptive Lasso Tobit regressions”, *Periodicals of Engineering and Natural Sciences*, Vol.8, No.2, pp. 1131-1140,2020.

[2] C. David. "*Modelling binary data*", CRC press, 2002.

[3] A. Katerina, J.He, and K. Yu.,"Binary quantile regression and variable selection: A new approach", *Econometric Reviews* ,Vol.38,No.6,pp. 679-694,2019.

[4] S. Aparecida, and H. S. Migon ,"Bayesian binary regression model: an application to in-hospital death after AMI prediction", *Pesquisa Operacional* ,Vol.24, No.2,pp. 253-267 ,2004.

[5] H. K. Abbas, and R. Alhamzawi, "Bayesian Adaptive Lasso binary regression with ridge parameter.", In *Journal of Physics: Conference Series*, Vol.1294, No.3, p.032036.IOP Publishing, 2019.

[6] B. Dries and D. V. Poel, "Binary quantile regression: a Bayesian approach based on the asymmetric Laplace distribution", *Journal of Applied Econometrics*, Vol.27, No.7, pp .1174-1188,2012.

[7] J. S. long and J. Freese, "Predicted probabilities for count models",*The Stata Journal* ,Vol.1,No 1, pp. 51–57,2001.

[8] T. Robert, "Regression shrinkage and selection via the lasso", *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol.58, No.1, pp. 267-288, 1996.
[9] Park T and Casella G, The Bayesian Lasso. Journal of the American Statistical Association, 103, pp. 681–686, 2008.
[10] M. Himel, and N. Yi, "A new Bayesian lasso", *Statistics and its interface*, Vol.7, No.4, pp. 571-582 ,2014.
[11] Z. Hui, "The adaptive lasso and its oracle properties", *Journal of the American statistical association*, Vol.101, No. 476, pp.1418-1429,2006.
[12] L. C. Minh-Ngoc Tran, and D. Nott,"Bayesian adaptive lasso", *Annals of the Institute of Statistical Mathematics*, Vol.66, No.2, pp. 221-244,2014.