Variable Selection in Bayesian Penalized Lasso Left Censored Regression Model

Zainab Saadoon Oleiwi
Universirty of Al-Qadisiyah
admin.stat21.16@qu.edu.iq

Bahr kadhim mohammed
Universirty of Al-Qadisiyah
Bahr.mahemmed@qu.edu.iq

Abstract

A new Bayesian lasso left censored regression (NBLLCR) method is proposed. This proposed method is presented by continuous uniform distribution $(\frac{-\sigma^2}{\lambda}, \frac{\sigma^2}{\lambda})$ with standard exponential distribution for a mixed representation of the Laplace distribution. The proposed method is compared with several existing Bayesian and non-Bayesian method using simulation examples and real data analysis. The results of the simulation studies and real data analysis show that the proposed method perform very well compared with other approaches.

Keywords: variable selection, left censored data, Bayesian regression, Laplace distribution.

1. Introduction

The concept of regression is one of the parametric statistical methods that indicate the extent to which the dependent variable value is affected by the change in the values of the independent variables. The linear regression model is given by the following formula:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{U}, \qquad \boldsymbol{U} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$$

where $\boldsymbol{y}$ is a vector of dependent observations, $\boldsymbol{X}$ is a matrix of independent observations, $\boldsymbol{\beta}$ is a vector of regression coefficients, and $\boldsymbol{U}$ is a vector of random errors.

The ordinary least squares method is one of the most popular traditional methods because it gives good predictions under specific assumptions. The least squares estimator is:

$$\widehat{\boldsymbol{\beta}}_{OLS} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y},$$

It makes the sum of squared errors (SSE) as minimum as possible.

Recently, a lot of effort has gone into developing different methods of variable selection in high-dimensional models. Regularization methods have grown in popularity as a result of their ability to at the same time select and estimate important coefficients. As a result, the variable selection (VS) characteristic was considered very important in the data analysis, because determining the important variables in the model can be difficult when the number of covariates is large. Donoho and Johnstone (1994) used for the first time, regularization techniques (VS). Where they proposed the soft-threshold estimator to obtain a smooth estimation of a function in the wavelet approximation. Tibshirani developed it after that in 1996 to obtain an estimate of the coefficients. The lasso technique can be applied in various statistical models. High predictive accuracy by reducing some coefficients to zero and thus decreasing the value of the variance while sacrificing a little bias, especially when the sample size (n) is small and the number of predictors (p) is large. Increasing the interpretability of the model. We often want to find a smaller set of predictors that have the strongest effects when we have a large number of them. The (VS) is used in regularization methods as part of the parameter estimation process, and examples of regularization techniques are lasso (Tibshirani, 1996), SCAD (Fan & Li, 2001), elastic net (Zou & Hastie, 2005), adaptive lasso (Zou, 2006), adaptive elastic net (Zou & Zhang, 2009), and MCP (Zhang, 2010). The Bayesian method is one of the important methods used in estimating the parameters of the model because of its importance in finding accurate estimates of the parameters and in overcoming the problems facing the estimation process using classical methods (Rencher & Schaalje, 2008). The Bayesian method in the lasso technique has become of great interest in recent years because of its great importance in inference. Park and Casella (2008) suggested the Bayesian lasso regression for linear models. To mix the normal distribution with the exponential distribution in representing the density function of the Laplace distribution. Hans (2009) compared the standard lasso regression and the Bayesian lasso regression and found the standard lasso method is not necessarily in agreement with the predictions of the Bayesian method. Mallick and Yi (2014) proposed the use of a uniform distribution of the scale mixture with a specific gamma (2, λ) by introducing a new Bayesian lasso method to solve the lasso problem in representing the density function of the Laplace distribution. Alhamzawi (2016) proposed a new method for the evaluation of the Tobit quantile regression model using a Bayesian elastic net. The method is called sparsity. He also used the gamma priors to develop a hierarchical prior model and introduced a new Gibbs sampling algorithm for the MCMC algorithm. The results of the study revealed that the proposed model outperforms other regularization methods. Alhusseini (2017) introduced the proposed model for the Tobit regression based on the lasso method. The Laplace distribution is a scale mixture of definite gamma and uniform distribution. The new Gibbs sampling algorithm has also been proposed. A simulation study and real data results of the studies revealed that the proposed model outperforms other methods. Flaih et al. (2020) introduced a new hierarchical model with new Gibbs samples as Bayesian analysis. A mixture of the normal distribution with Rayleigh density was used to represent the density function of the previous Laplace distribution.

In this paper, we propose a mixed representation of the Laplace distribution by following mathematical procedures and transformations for the mixed representation of the Laplace distribution, and a representation was obtained expressed by the continuous uniform distribution $(\frac{-\sigma^2}{\lambda}, \frac{\sigma^2}{\lambda})$ multiplied by the standard exponential distribution. This proposal was mapped to both the Bayesian regression of the lasso method and to censored data from the left side.

---

2. Left Censored Data

The latent variable in many real-world applications has a high number of observations that are less than a specific value, which is known as left censored data. These facts have been found in numerous scientific disciplines such as; economy, medicine, chemistry, and physics. When the values of the dependent variable are unknown but the values of the independent variables are known, censored regression models are utilized, if the dependent variable's actual values are greater than the lower limit, they are observed.

$$y_i = \begin{cases} y_i^* & if \quad y_i^* > y_L, \\ y_L & if \quad y_i^* \le y_L, \end{cases} \qquad (1)$$

$$y_i^* = \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_i, \qquad \varepsilon_i \sim N(0, \sigma^2)$$

where y_i dependent variable, y_i^* laten variable (unobserved), y_L is the restriction point, $\boldsymbol{x}_i'$ is a vector of predictors, $\boldsymbol{\beta}$ is a vector of the regression coefficients, and ε_i is an error term. The censored regression model is the Tobit model when ($y_L = 0$).

$$y_i = \begin{cases} y_i^* & if \quad y_i^* > 0, \\ 0 & if \quad y_i^* \le 0, \end{cases} \qquad i = 1, \dots, n,$$

$$y_i^* = \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_i. \qquad \varepsilon_i \sim N(0, \sigma^2)$$

If no data are censored, the Tobit model is the same as an OLS regression. If the actual value is less than a cutoff point y_L, the left censored value is unobserved. (Carson & Sun, 2007; Amemiya, 1984; Anastasopoulos et al., 2008; Chib, 1992;

Tobin, 1958; Alshaybawee et al., 2017; Alhamzawi et al., 2011; Alhamzawi & Ali, 2018; Alhamzawi, 2021).

3. The Proposed Scale Mixture

Based on the following mathematically fact,

$$\int_{w>\frac{|x|}{\sigma^2}} \lambda e^{-\lambda w}\, dw = e^{-\frac{\lambda|x|}{\sigma^2}} \tag{2}$$

we can propose the following scale mixture formula. In (2), let $x = \beta,\ \lambda w = z$, and by multiply both sides by $\frac{\lambda}{2\sigma^2}$, we get

$$\frac{\lambda}{2\sigma^2}\int_{\frac{z}{\lambda}>\frac{|\beta|}{\sigma^2}} \lambda e^{-z}\frac{1}{\lambda}dz = \frac{\lambda}{2\sigma^2}e^{-\frac{\lambda|\beta|}{\sigma^2}}$$

$$\frac{\lambda}{2\sigma^2}e^{-\frac{\lambda|\beta|}{\sigma^2}} = \int_{z>\frac{\lambda|\beta|}{\sigma^2}} \frac{\lambda}{2\sigma^2}e^{-z}\, dz \tag{3}$$

so, the formulation (3) is the scale mixture of standard exponential mixing with uniform $(\frac{-\sigma^2}{\lambda}, \frac{\sigma^2}{\lambda})$.

3.1 The Hierarchical Prior Model of Left-Censored Data

Based on the proposed scale mixture (3), and (1). The hierarchical prior model is formulated as follows:

$$y_i = \begin{cases} y_i^* & if\ \ y_i^* > y_L, \\ y_L & if\ \ y_i^* \le y_L, \end{cases}$$

$$y_i^*|\boldsymbol{X}, \boldsymbol{\beta}, \sigma^2 \sim N(\boldsymbol{X\beta}, \sigma^2\boldsymbol{I}_n), \tag{4}$$

$$\boldsymbol{\beta}|\,\sigma^2,\lambda \sim Uniform\left(-\frac{\sigma^2}{\lambda},\frac{\sigma^2}{\lambda}\right),$$

$$\sigma^2 \sim \pi(\sigma^2)\; d\sigma^2,$$

$$\lambda \sim Gamma(c\,,d),$$

$$z \sim Exp(1).$$

Where $\boldsymbol{X}$ is the standardized covariate matrix, and $\boldsymbol{y}^*$ are the centered unobserved response variable values.

3.2 The Gibbs Sampling Algorithms

We suppose that the full joint density as follows:

$$f(\boldsymbol{y}^*|\boldsymbol{\beta},\sigma^2)\;\pi(\sigma^2)\;\pi(\lambda)\;\prod_{j=1}^{k}\pi\,(\boldsymbol{\beta}|\sigma^2,\lambda)\;\pi(z_j)\;I\left\{z_j > \frac{\lambda|\beta_j|}{\sigma^2}\right\}$$

$$= exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y}^* - \boldsymbol{X\beta})'(\boldsymbol{y}^* - \boldsymbol{X\beta})\right\}\frac{(\sigma^2)^{-q-1}}{\Gamma(q)}\;\theta^q\;e^{-\theta/\sigma^2}$$

$$\frac{(\lambda)^{c-1}}{\Gamma(c)}\;d^c\;e^{-d\lambda}\;e^{-\sum_{j=1}^{k}z_j}\;\prod_{j=1}^{k}\frac{\lambda}{2\sigma^2}I\left\{z_j > \frac{\lambda|\beta_j|}{\sigma^2}\right\}$$

Now, the full conditional posterior distributions are defined by:

1. The full conditional posterior distribution of $\boldsymbol{y}^*$ is:

$$y_i^*|y_i\,,\boldsymbol{\beta} \sim \begin{cases} y_i & if\quad y_i^* > y_L, \\ N(\boldsymbol{X\beta},\sigma^2\boldsymbol{I}_n) & if\quad y_i^* \le y_L, \end{cases}$$

2. The full conditional posterior distribution of $\boldsymbol{\beta}$ is:

$$\pi(\boldsymbol{\beta}|\boldsymbol{y}^*,\boldsymbol{X},\sigma^2,z) \propto \pi(\boldsymbol{y}^*|\boldsymbol{X},\boldsymbol{\beta},\sigma^2)\;\pi(\boldsymbol{\beta}|z,\sigma^2,\lambda)$$

$$\propto exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y}^* - \boldsymbol{X\beta})'(\boldsymbol{y}^* - \boldsymbol{X\beta})\right\}\prod_{j=1}^{k}I\left\{|\beta_j| < \frac{z_j\sigma^2}{\lambda}\right\}$$

$$\propto exp\left\{-\frac{1}{2\sigma^2}\left(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}_{OLS}\right)'\boldsymbol{X}'\boldsymbol{X}\left(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}_{OLS}\right)\right\}\prod_{j=1}^{k} I\left\{\frac{-z_j\sigma^2}{\lambda}<\beta_j<\frac{z_j\sigma^2}{\lambda}\right\}.$$

Hence,

$$\boldsymbol{\beta}|\boldsymbol{y},\boldsymbol{X},z,\lambda,\sigma^2 \sim N_k(\widehat{\boldsymbol{\beta}}_{OLS},\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1})\prod_{j=1}^{k} I\left\{\frac{-z_j\sigma^2}{\lambda}<\beta_j<\frac{z_j\sigma^2}{\lambda}\right\}$$

3. The full conditional posterior distribution of σ^2 is:

$$\pi(\sigma^2|\boldsymbol{y}^*,\boldsymbol{X},\boldsymbol{\beta}) \propto \pi(\boldsymbol{y}^*|\boldsymbol{X},\boldsymbol{\beta},\sigma^2)\,\pi(\sigma^2)\,\pi(\boldsymbol{\beta}|\sigma^2,\lambda,z)$$

$$\propto \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y}^*-\boldsymbol{X\beta})'(\boldsymbol{y}^*-\boldsymbol{X\beta})\right\}(\sigma^2)^{-q-1}e^{-\theta/\sigma^2}$$

$$(\sigma^2)^{-k}\prod_{j=1}^{k} I\left\{z_j>\frac{\lambda|\beta_j|}{\sigma^2}\right\}$$

$$\propto (\sigma^2)^{-\frac{n}{2}-q-1-k} exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y}^*-\boldsymbol{X\beta})'(\boldsymbol{y}^*-\boldsymbol{X\beta})\right\} e^{-\theta/\sigma^2}$$

$$I\left\{\sigma^2>Max_j\left(\frac{\lambda|\beta_j|}{z_j}\right)\right\}$$

$$\propto (\sigma^2)^{-\frac{n}{2}-q-k-1} exp\left\{-\left[\frac{1}{2\sigma^2}(\boldsymbol{y}^*-\boldsymbol{X\beta})'(\boldsymbol{y}^*-\boldsymbol{X\beta})+\frac{\theta}{\sigma^2}\right]\right\}$$

$$I\left\{\sigma^2>Max_j\left(\frac{\lambda|\beta_j|}{\sigma_j^2}\right)\right\}.$$

Therefore,

$$\sigma^2|\boldsymbol{y},\boldsymbol{X},\boldsymbol{\beta},z,\lambda \sim Inverse-Gamma\left(\frac{n}{2}+q+k,\frac{(\boldsymbol{y}^*-\boldsymbol{X\beta})'(\boldsymbol{y}^*-\boldsymbol{X\beta})}{2+\theta}\right)$$

4. The full conditional posterior distribution of z is:

$$\pi(z|\boldsymbol{\beta},\lambda,\sigma^2) \propto \pi(z)\,\pi(\boldsymbol{\beta}|z,\lambda,\sigma^2)$$

$$\propto \prod_{j=1}^{k} e^{-z_j} I\left\{z_j > \frac{\lambda|\beta_j|}{\sigma^2}\right\}.$$

Therefore,

$$z \sim \prod_{j=1}^{k} standard\ exponential\ \ I\left\{z_j > \frac{\lambda|\beta_j|}{\sigma^2}\right\}$$

5. The full conditional posterior distribution of λ is:

$\pi(\lambda|\boldsymbol{\beta}) \propto \pi(\boldsymbol{\beta}|\lambda)\ \pi(\lambda)$

$$\propto \left(\frac{\lambda}{2\sigma^2}\right)^k\ \lambda^{c-1}\ e^{-d\lambda}\ \prod_{j=1}^{k} I\left\{z_j > \frac{\lambda|\beta_j|}{\sigma^2}\right\}$$

$$\propto \lambda^{k+c-1}\ e^{-d\lambda}\ \prod_{j=1}^{k} I\left\{\lambda < \frac{z_j\sigma^2}{|\beta_j|}\right\}.$$

Therefore,

$$\lambda \sim Gamma\ (k+c, d) \prod_{j=1}^{k} I\left\{\lambda < \frac{z_j\sigma^2}{|\beta_j|}\right\}$$

---

4. Simulation Study and Real Data

4.1 Simulation Study

In this section, we demonstrate the prediction accuracy of the methods; linear left censored regression (Tobit), Bayesian lasso left censored regression (BLLCR), the new Bayesian lasso left censored regression (NBLLCR), and Bayesian lasso left censored regression using scale mixture uniform (BLLCRsmu). The outcome variable is centered and the covariates are standardized to have 0 means and unit variances before applying the above methods. For prediction accuracy, we evaluate the median of mean squared errors (MMSE) for the simulated studies based on 100 replications.

4.1.1 Example 1 (Left censored with sparse case)

In this example, we generate data from the correct model (Alhamzawi and Ali, 2020)

$$y_i = max\{0, y_i^*\}, \qquad i = 1, \dots, n,$$

$$y_i^* = \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_i, \qquad \varepsilon_i \sim N(0, \sigma^2)$$

We set $\beta^{10\times 1} = (6,1,0,0,3,0,0,0,0,0)'$ and $\sigma = \{1,3,5\}$. For each simulation study, we generate a training set (n_t) with $n_t = \{100,150,200\}$ observations and a testing set with 200 observations. The covariates are simulated from the multivariate normal distribution with mean zero, variance 1, and pairwise correlations between x_i and x_j equal to $0.5^{|i-j|}$ $\forall\ i \neq j$.

The results are listed in Table 1. The results show that the new Bayesian lasso left censored regression (NBLLCR) performs very well compared to other methods in the comparison. It has the smallest MMSE in 5 out of 9 experimental results. The Bayesian lasso left censored regression using scale mixture uniform (BLLCRsmu) also performs well compared to other methods in the comparison. It has the smallest MMSE in 3 out of 9 experimental results.

Table 1: Median mean squared error (MMSE) and their associated standard deviations (SD) are listed in the parentheses for Example 1. All results are averaged over 100 replications.

(n_t, n_p, σ^2)	Tobit	BLLCR	NBLLCR	BLLCRsmu
(100, 200, 1)	0.2102 (0.0879)	0.4866 (0.2924)	**0.1684** (0.0944)	0.1770 (0.0746)
(100, 200, 9)	1.4887 (0.7463)	1.6287 (0.6630)	**0.8467** (0.4599)	1.0394 (0.5372)
(100, 200, 25)	3.9987 (2.2121)	4.5781 (2.3237)	**2.8790** (1.7789)	2.9307 (1.6958)
(150, 200, 1)	0.1349 (0.0696)	0.2422 (0.1171)	0.1247 (0.0754)	**0.1221** (0.0642)
(150, 200, 9)	0.8980 (0.3243)	1.2659 (0.5821)	**0.6485** (0.2042)	0.7319 (0.2369)
(150, 200, 25)	2.6600 (1.2711)	2.5127 (1.5803)	2.0214 (1.1362)	**1.9633** (1.1111)
(200, 200, 1)	0.1111 (0.0661)	0.1411 (0.0883)	**0.0857** (0.0624)	0.1009 (0.0631)
(200, 200, 9)	0.7527 (0.3544)	0.7273 (0.3896)	**0.5378** (0.2738)	0.6142 (0.3105)
(200, 200, 25)	1.7358 (0.8284)	1.9657 (0.9514)	1.3485 (0.6781)	**1.3257** (0.6370)

Convergence of the corresponding our Gibbs sampler methods was assessed by trace plots of the simulated draws. The trace plots Figures $1-3$ shows that our methods converge very fast.

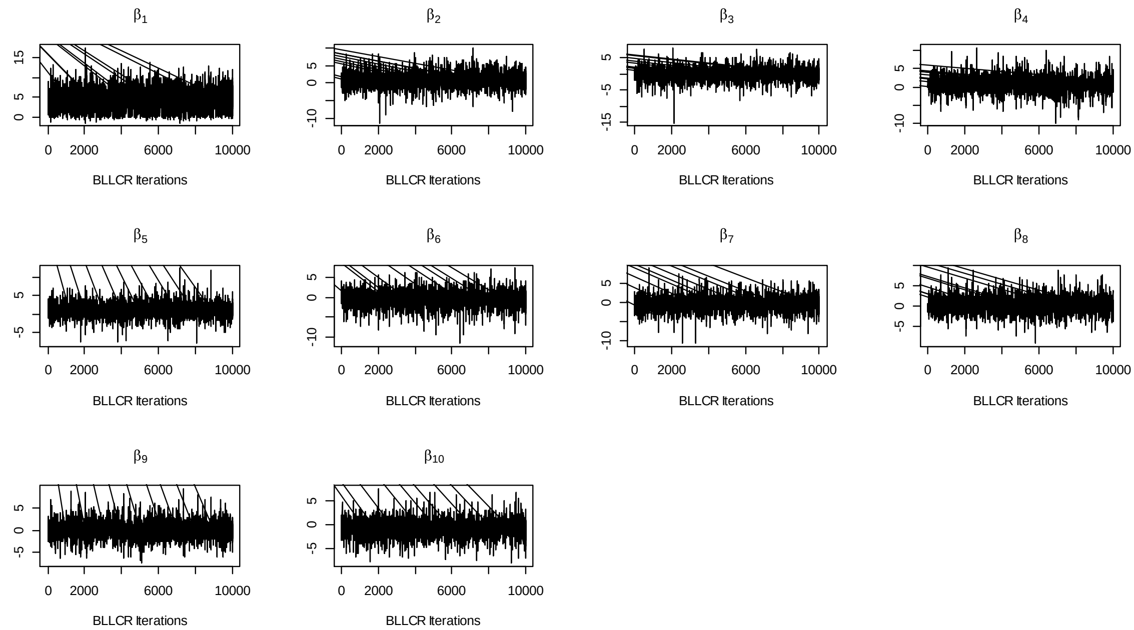

Figure 1: Trace plots of parameters in Simulation 1 using BLLCR method.

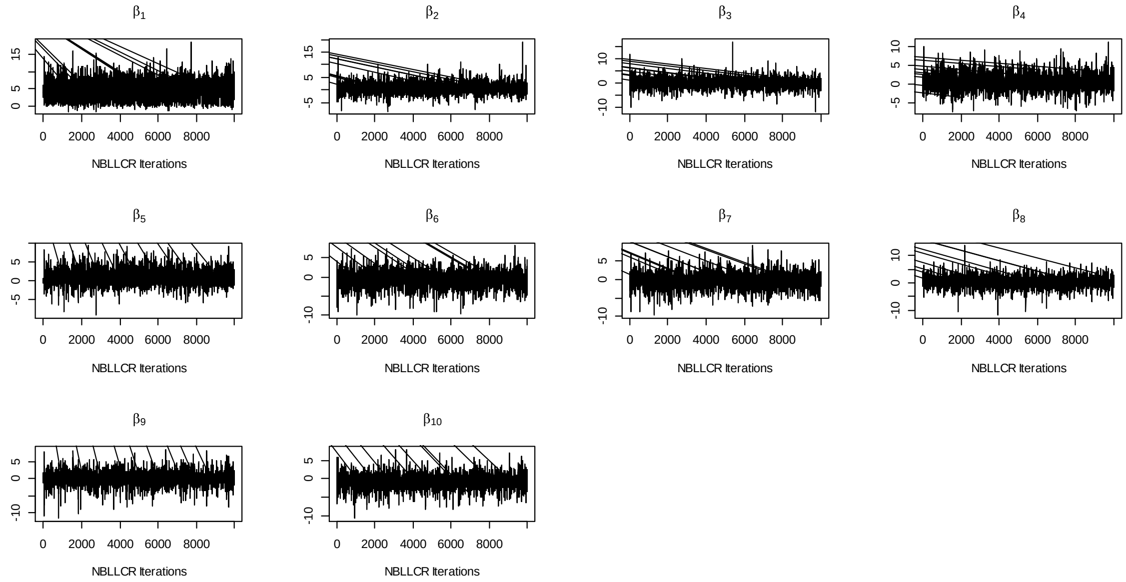

Figure 2: Trace plots of parameters in Simulation 1 using NBLLCR method.

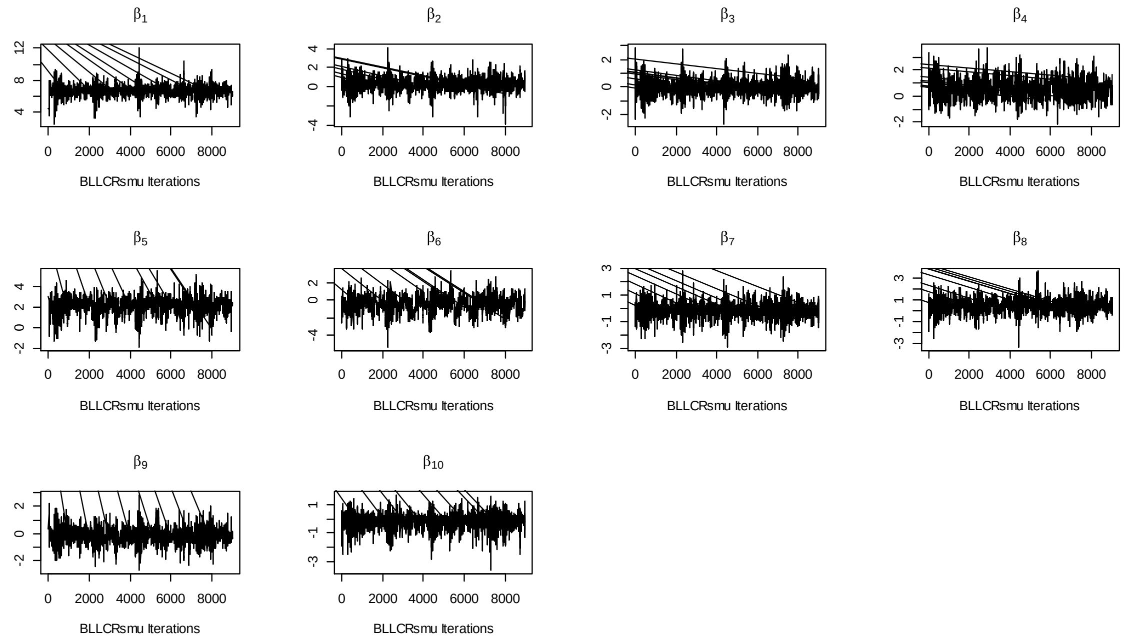

Figure 3: Trace plots of parameters in Simulation 1 using BLLCR method.

4.1.2 Example 2 (Left censored with dense case)

Here we set $\beta^{10\times 1} = (6,1,1,1,1,1,1,1,1,1)'$, leaving other setups exactly the same as in Example 1.

Table 2: Median mean squared error (MMSE) and their associated standard deviations (SD) are listed in the parentheses for Example 2. All results are averaged over 100 replications.

(n_t, n_p, σ^2)	Tobit	BLLCR	NBLLCR	BLLCRsmu
(100, 200, 1)	0.1771 (0.0762)	1.1996 (0.2868)	0.3049 (0.1845)	**0.1740** (0.0798)
(100, 200, 9)	1.4155 (0.6258)	3.7530 (1.1582)	1.3331 (0.8354)	**1.2620** (0.6742)
(100, 200, 25)	5.3689 (2.9968)	6.0693 (2.3489)	4.4063 (2.0186)	**3.9529** (1.9710)
(150, 200, 1)	**0.1340** (0.0405)	0.6167 (0.1878)	0.2550 (0.1122)	0.1352 (0.0417)
(150, 200, 9)	1.0950 (0.6451)	2.4081 (0.5277)	0.9955 (0.4794)	**0.9890** (0.5311)
(150, 200, 25)	2.7741 (1.0659)	5.1234 (2.5299)	2.6020 (0.9744)	**2.4885** (0.9346)
(200, 200, 1)	**0.0963** (0.0451)	0.2967 (0.1102)	0.1552 (0.0724)	0.0976 (0.0469)
(200, 200, 9)	**0.6958** (0.3017)	1.9084 (0.4891)	0.7644 (0.3170)	0.7009 (0.3045)
(200, 200, 25)	2.0036 (0.8296)	3.3814 (1.4898)	1.9139 (0.7927)	**1.7670** (0.7850)

The results are lists in Table 2. The results show that The Bayesian lasso left censored regression using scale mixture uniform (BLLCRsmu) performs very well compared to other methods in the comparison. It has the smallest MMSE in 6 out of 9 experimental results.

4.1.3 Example 3 (Left censored with very sparse case)

Here we set $\beta^{10\times 1} = (6,0,0,0,0,0,0,0,0,0)'$, leaving other setups exactly the same as in Example 1. The results are lists in Table 3. The results show that the new

Table 3: Median mean squared error (MMSE) and their associated standard deviations (SD) are listed in the parentheses for Example 3. All results are averaged over 100 replications.

(n_t, n_p, σ^2)	Tobit	BLLCR	NBLLCR	BLLCRsmu
(100, 200, 1)	0.2151 (0.0782)	0.1979 (0.1410)	**0.0923** (0.0482)	0.1652 (0.0560)
(150, 200, 9)	1.4375 (0.8445)	1.7693 (0.4738)	**0.7129** (0.4676)	0.8499 (0.5963)
(200, 200, 25)	4.5167 (2.0976)	**2.2472** (1.8837)	3.0784 (1.4806)	2.4561 (1.3518)
(100, 200, 1)	0.1026 (0.0564)	0.0919 (0.0750)	**0.0435** (0.0324)	0.0813 (0.0483)
(150, 200, 9)	0.8673 (0.3041)	0.5420 (0.2819)	**0.5054** (0.2150)	0.5792 (0.1988)
(200, 200, 25)	1.8641 (0.8668)	**0.8514** (0.6257)	1.1804 (0.6008)	1.9366 (0.4496)
(100, 200, 1)	0.0812 (0.0318)	0.0749 (0.0259)	**0.0410** (0.0185)	0.0661 (0.0269)
(150, 200, 9)	0.5502 (0.1860)	0.4886 (0.2758)	**0.3573** (0.1474)	0.3891 (0.1495)
(200, 200, 25)	1.5103 (0.5687)	**0.8622** (0.4930)	1.0590 (0.4349)	1.8843 (0.3484)

New Bayesian lasso left censored regression (NBLLCR) performs very well compared to other methods in the comparison. It has the smallest MMSE in 6 out of 9 experimental results. The Bayesian lasso left censored regression (BLLCR) also performs well compared to other methods in the comparison. It has the smallest MMSE in 3 out of 9 experimental results.

4.2 Real Data

In this section, we demonstrate the performance of the methods with extramarital affairs data. A detailed discussion of this data set can be found in Chernozhukov and Hong (2011), and this data set is available in the R package AER. The original data has 601 observations and on 9 variables. We use a random subsample of this dataset which has 100 observations The dependent variable is affairs (the number of extramarital sexual intercourse during the past year). The other eight independent variables include: gender (x_1), age (x_2), years (x_3), children (x_4), religiousness (x_5), education (x_6), occupation (x_7), and rating (x_8).

The results are listed in Table 4, The results show that the new Bayesian lasso left censored regression (NBLLCR) performs very well compared to other methods in the comparison.

Table 4: Mean squared error (MSE) for the affairs data.

Tobit	BLLCR	NBLLCR	BLLCRsmu
118.42581	146.17364	**31.93521**	123.16909

5. Conclusions

The median of mean squared errors (MMSE) was used to know prediction accuracy, and by applying the Markov Chain Monte Carlo (MCMC) method in the simulation and real data, the study found that the estimator of the new Bayesian lasso left censored regression (NBLLCR) method is the best compared to other methods, based on the value of mean squared error (MSE).

References

Alhamzawi, R. (2016). Bayesian elastic net Tobit quantile regression. Communications in Statistics-Simulation and Computation, 45(7), 2409-2427.

Alhamzawi, R. (2021). The reciprocal Bayesian bridge for left-censored data. Communications in Statistics-Simulation and Computation, 1-9.

Alhamzawi, R., & Ali, H. T. M. (2018). Bayesian tobit quantile regression with penalty. Communications in Statistics-Simulation and Computation, 47(6), 1739-1750.

Alhamzawi, R., & Taha Mohammad Ali, H. (2020). A new Gibbs sampler for Bayesian lasso. Communications in Statistics-Simulation and Computation, 49(7), 1855-1871.

Alhamzawi, R., Yu, K., Vinciotti, V., & Tucker, A. (2011). Prior elicitation for mixed quantile regression with an allometric model. Environmetrics, 22(7), 911-920.

Alhusseini, F. H. H. (2017). New Bayesian Lasso in Tobit Quantile Regression. Romanian Statistical Review Supplement, 65(6), 213-229.

Alshaybawee, T., Midi, H., & Alhamzawi, R. (2017). Bayesian elastic net single index quantile regression. Journal of Applied Statistics, 44(5), 853-871.

Amemiya, T. (1984). Tobit models: A survey. Journal of econometrics, 24(1-2), 3-61.

Anastasopoulos, P. C., Tarko, A. P., & Mannering, F. L. (2008). Tobit analysis of vehicle accident rates on interstate highways. Accident Analysis & Prevention, 40(2), 768-775.

Carson, R. T., & Sun, Y. (2007). The Tobit model with a non-zero threshold. *The Econometrics Journal*, *10*(3), 488-502.

Chib, S. (1992). Bayes inference in the Tobit censored regression model. Journal of Econometrics, 51(1-2), 79-99.

Donoho, D. L., & Johnstone, J. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *biometrika*, *81*(3), 425-455.

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, *96*(456), 1348-1360.

Flaih, A. N., Alshaybawee, T., & Alhusseini, F. H. H. (2020). Sparsity via new Bayesian Lasso. Periodicals of Engineering and Natural Sciences, 8(1), 345-359.

Hans, C. (2009). Bayesian Lasso regression. *Biometrika*, *96*(4), 835-845.

Mallick, H., & Yi, N. (2014). A new Bayesian lasso. Statistics and its interface, 7(4), 571-582.

Park, T., & Casella, G. (2008). The bayesian Lasso. *Journal of the American Statistical Association*, *103*(482), 681-686.

Rencher, A. C., & Schaalje, G. B. (2008). Linear models in statistics. John Wiley & Sons.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267-288.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica: journal of the Econometric Society*, 24-36.

Zhang, C. H. (2010). Nearly unbiased variable selection under minimax concave penalty. *The Annals of statistics*, *38*(2), 894-942.

Zou, H. (2006). The adaptive lasso and its oracle properties. Journal of the American statistical association, 101(476), 1418-1429.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, *67*(2), 301-320.

Zou, H., & Zhang, H. H. (2009). On the adaptive elastic-net with a diverging number of parameters. Annals of statistics, 37(4), 1733.