A New Bayesian Group Bridge to Solve the Tobit Model

Saja Aljanabi*, Rahim Alhamzawi

Statistics Department, College of Administration and Economics, University of Al-Qadisiyah, Al Diwaniyah, Iraq

Abstract

In this paper, we propose a new approach of regularization for the left censored data (Tobit). Specifically, we propose a new Bayesian group Bridge for left-censored regression (BGBRLC). We developed a new Bayesian hierarchical model and we suggest a new Gibbs sampler for posterior sampling. The results show that the new approach performs very well compared to some existing approaches.

Keywords: left- censored regression, variable selection, Bayesian group Bridge left-censored regression (BGBRLC).

<div dir="rtl">

طريقة بيزية جديدة لمجموعة البرج لحل نموذج التوبت

سجى الجنابي*،رحيم الحمزاوي

قسم الإحصاء ، كلية الإدارة والاقتصاد ، جامعة القادسية ، الديوانية ، العراق.

الخلاصه

في هذه الورقة ، نقترح طريقة جديدة لتنظيم البيانات الرقابة اليسار (التوبت). على وجه التحديد ، نقترح طريقة بيزية جديدة لمجموعة البرج للانحدار الخاضع للرقابة اليسار (BGBRLC) . لقد طورنا نموذجًا هرميًا جديدًا بايزيًا ونقترح أخذ عينات جديدة من Gibbs لأخذ العينات الخلفية. أظهرت النتائج أن النهج الجديد يؤدي أداءً جيدًا مقارنةً ببعض النهج الحالية.

</div>

1-Introduction

Left censored regression is a statistical method in which the observed response variable is censored from below. Examples of such data are various and cover many different areas, such as agriculture, genetics, environment and medicine, etc. Left censored regression is considered as one of the good methods for assessing the correlation between a set of explanatory variables and a dependent variable. One of the important left censored regression problem is when explanatory variables are very large. Therefore, it is difficult to identify important variables. Most research focuses on variable selection to obtain the appropriate model. Traditional methods of variable selection include Mallow's C_k, suggested by Mallows [1]:

$$C_k = \frac{RSS(k)}{S^2} - n + 2k,$$

where S^2 is the mean square error of the model, $RSS(k)$ is a residuals sum of squares, k is a number of co-variates in the model, and n is the sample size of data. The small values of C_k mean that the model is relatively accurate. Woodroofe [2] showed that C_k selects the conservative model. Nishii [3] showed that C_k is inconsistent in selecting the right model, and often selecting a larger model when $n \to \infty$.

---

*Email: sajaljanabi@yahoo.com

Akaike [4] proposed Akaike Information Criterion (AIC), which is defined by

$$AIC = -2ln\,L + 2k,$$

where L is a maximum likelihood function (MLE). Javed and Mantalos [5] showed that the selected model using AIC is inconsistent when the sample size is large. For the sake of eliminating this issue, Schwarz [6] presented the Bayes information criterion (BIC)

$$BIC = -2ln\,L + k\,ln(n),$$

This method overcomes the problem of (AIC) and selects a model with good properties. However, when $k > n$, this method cannot handle the problem of variable selection.

George and Mc-Culloch [7] presented a stochastic search variable selection (SSVS) as an attractive way to select a subset of covariates using a mixture of prior distributions that allows some coefficients equal to zero

One of the disadvantages of this approach is the long time to select the correct model (i.e., it is time consuming). In addition, in high dimensional data, the algorithm cannot visit the correct model.

Recently, regularization methods became more popular because they simultaneously select and estimate the important coefficients; see for example Hans [8], Li *et al*. [9] Alhamzawi and Yu [10], Tibshirani [11] Liu *et al*. [12], Alhamzawi [13]. Mallick and Yi [14], Xu and Ghosh [15], and Alhamzawi and Ali [16]. The general formula of the regularization methods is as follows

$$\hat{\beta} = argmin\ (y - X\beta)'(y - X\beta) + \mathcal{G}_\lambda(\beta), \tag{1}$$

where $\mathcal{G}_\lambda(\beta)$ is a function of the model coefficients which controls the degree of penalty in terms of tuning parameter $\lambda > 0$. Hoerl and Kinnard [17] proposed the Ridge regression which has a better predictive performance than (OLS) estimates, with a lower variance. However, the Ridge regression cannot produce an optimal model, because it always retains all predictors in the model.

Frank and Friedman [18] suggested that the Bridge regression has attractive features such as Oracle, unbiasedness, as well as the variable selection and parameter estimation of the model, but the approximate covariance matrix and bootstrap calculated standard errors are unsteady.

Tibshirani [19] proposed the Lasso regression which automatically selects the important variable by shrinking some unimportant coefficients to zero.

In recent years, researchers focused on selecting influential groups of variables. Yuan and Lin [20] proposed a group Lasso, which was expanded by Kim et al. [21] to general loss functions. The group lasso regression cannot select a binary variable.

Huang et al. [22] proposed a group bridge regression, which is capable of selecting a bi-level variable with oracle property and sparsity [23,24].

Aljanabi.S and Alhamzawi [25] (Accepted paper) proposed a new Bayesian group lasso in left-censored regression models for the simultaneous variable selection and parameter estimation , where the results of data analysis and simulation showed that the proposed method performed better than the other approaches.

In this research, we propose a new Bayes group bridge for left-censored data. Then, a new Gibbs sampler algorithm for variable selection is implemented. Simulation researches and real data analysis show that the new approach's performance is very well in comparison to the existing methods.

In Section 2, we provide an overview of the left-censored model. In section 3, we describe the Bayesian group bridge regression for left-censored data and present a new Bayesian hierarchical model, In Section 4, we carry out Monte Carlo simulations to demonstrate the performance of the proposed method. In section 5, we analyze the Real data and in section 6 we draw the conclusions.

2. Methods

Consider the left-censored model

$$y_i = \begin{cases} y_i^\star & if & y_i^\star > c \\ c & if & y_i^\star \le c \end{cases}$$
$$\boldsymbol{y}^\star = \boldsymbol{X\beta} + \boldsymbol{u} \tag{2}$$

where c is a left censored point,
$\boldsymbol{y}^\star = (y_1^\star, \dots, y_n^\star)', \boldsymbol{X} = (\boldsymbol{X}_1, \dots, \boldsymbol{X}_G)', with\ each\ \boldsymbol{X}_g, g = 1, \dots, G$,
$\boldsymbol{\beta} = (\beta_1, \dots, \beta_G)', \boldsymbol{u} = (u_1, \dots, u_n)', and\ u_i \sim N(0, \sigma^2).$

3. Bayesian Group Bridge For Left-censored model

Huang et al. [22] suggested that the bridge group is able to select the important groups and select within each group,

$$\hat{\beta} = argmin(y - X\beta)'(y - X\beta) + \sum_{g=1}^{G} \lambda_g \left\| \beta_g \right\|_1^{\varsigma},$$

where $y = (y_1, \dots y_n)'$ and ς denote the parameter concavity.

Entering multiple parameters will aggregate the information between variables within the group and accommodate the shrinkage through specific parameters. Despite the good and desirable characteristics of this technique, it does not provide correct or reliable standard errors [26]. The Bayesian approach overcomes these disadvantages and can provide standard errors. Following Huang *et al*. [22], the Bayesian group bridge for censored data can be written as:

$$\hat{\beta} = argmin \, (y^\star - X\beta)'(y^\star - X\beta) + \sum_{g=1}^{G} \lambda_g \left\| \beta_g \right\|_1^{\varsigma},$$

where $\beta = (\beta_1, \dots, \beta_G)'$, G is the number of the groups and $\|.\|$ is the l_1 norm of β_g. We will use scale mixture of uniforms (SMU) for representing the generalized Gaussian (GG) prior, making the Markov Chain Monte Carlo (MCMC) algorithm work with good computational efficiency. The conditional GG prior distribution of $\boldsymbol{\beta}$ is given by Mallick and Yi [26], as follows

$$\pi(\boldsymbol{\beta}|\varsigma, \lambda_1, \dots, \lambda_g) \propto \prod_{g=1}^{G} exp\left(-\lambda_g \|\beta_g\|_1^{\varsigma}\right).$$

The most important step in the Bayesian approach is to determine the prior distribution of parameters. It is also of great importance that the selection must be accurate because the opposite will lead to many problems, as previously shown by Kenny and Donson [27], Alhamzawi and Yu [28], and Alhamzawi and Ali [29]. Following Mallick and Yi [26], to perform the Bayesian analysis, we set the next prior distribution of $\boldsymbol{\beta}$ as follows;

$$\pi(\boldsymbol{\beta}) = C(\lambda, \varsigma) exp\left(-\lambda \|\boldsymbol{\beta}\|_1^{\varsigma}\right),$$

where $C(\lambda, \varsigma) = \frac{\lambda^{\frac{k}{\varsigma}} \Gamma(k+1)}{2^k \Gamma(\frac{k}{\varsigma}+1)}$ is the normalizing constant.

$$\pi(\boldsymbol{\beta}) = \frac{\lambda^{\frac{k}{\varsigma}} \Gamma(k+1)}{2^q \Gamma\left(\frac{k}{\varsigma}+1\right)} exp\left(-\lambda \|\beta_g\|_1^{\varsigma}\right), \tag{3}$$

$$= \frac{\lambda^{\frac{k}{\varsigma}+1} \Gamma(k+1)}{2^k \Gamma\left(\frac{k}{\varsigma}+1\right)} \int_{\|\boldsymbol{\beta}\|_1 < w^{\frac{1}{\varsigma}}} exp\{-\lambda w\} dw$$

Let $V_q(Q) = \frac{2^k w^{\frac{k}{\varsigma}}}{\Gamma(k+1)}, w > 0$

$$= \int_{\|\boldsymbol{\beta}\|_1 < w^{\frac{1}{\varsigma}}} \frac{\lambda^{\frac{k}{\varsigma}+1}}{\Gamma\left(\frac{k}{\varsigma}+1\right)} w^{\frac{k}{\varsigma}} exp\{-\lambda w\} dw$$

In the present study, we convert the above formula in the following manner:
Let $v = \lambda w \implies dv = \lambda \, dw$ then

$$\frac{\lambda^k}{2^k} e^{-\lambda \|\beta\|_1^{\varsigma}} = \int_{\|\lambda \boldsymbol{\beta}\|_1 < v^{\frac{1}{\varsigma}}} \frac{\lambda^{\frac{k}{\varsigma}+1}}{\Gamma\left(\frac{k}{\varsigma}+1\right)} \left(\frac{v}{\lambda}\right)^{\frac{k}{\varsigma}} exp(-v) \frac{dv}{\lambda}$$

$$= \int_{\|\lambda \boldsymbol{\beta}\|_1 < v^{\frac{1}{\varsigma}}} \frac{1}{\Gamma\left(\frac{k}{\varsigma}+1\right)} v^{\frac{k}{\varsigma}} exp(-v) \, dv \tag{4}$$

3.1 Hierarchical Representation

We construct our Bayesian hierarchical model following the hierarchical model of Mallick and Y, [26], as follows:

$$y_i = max\{y_i^\star, c\}, \quad i = 1, \ldots, n$$
$$\boldsymbol{y}^\star | \boldsymbol{X}, \boldsymbol{\beta}, \sigma^2 \sim N_n(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 I_n)$$
$$\beta_g | v, \varsigma \sim Multivariate\ Uniform, g = 1, \ldots \ldots, G$$
$$v | \lambda_1, \ldots, \lambda_G, \varsigma \sim \prod_{g=1}^{G} Gamma\left(\frac{k}{\varsigma}, 1\right)$$
$$\sigma^2 \sim \frac{b^r}{\Gamma(a)} \sigma^{2-r-1} \exp\left(-\frac{b}{\sigma^2}\right), \quad r, b > 0$$
$$\lambda_g \sim \lambda_g exp^{-\theta\lambda_g}$$

where $v = (v_1, \ldots, v_G)$.

3.2 Full Conditional Distributions

1- The full conditional distribution of $y_i^\star$ is

$$y_i^\star | y_i, \boldsymbol{\beta} \sim \begin{cases} \gamma(y_i) & if \quad y_i > c, \\ N(\boldsymbol{x_i'}\boldsymbol{\beta}, \sigma^2) & otherwise, \end{cases}$$

where γ is a degenerated distribution.

2- The full conditional distribution of $(\boldsymbol{\beta})$ is

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

$$\boldsymbol{\beta}|\boldsymbol{y}^\star, \boldsymbol{X}, v, \sigma^2, \lambda_1, \ldots, \lambda_G, \varsigma \sim N_k\left(\widehat{\boldsymbol{\beta}}, \sigma^2(\boldsymbol{X'X})^{-1}\right) \prod_{g=1}^{G}\left\{-\frac{v_g^{\frac{1}{\varsigma}}}{\lambda_g} < \beta_g < \frac{v_g^{\frac{1}{\varsigma}}}{\lambda_g}\right\}, \tag{5}$$

here $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'y}^\star$.

3- The full conditional distribution of (v) is

$$\pi(\boldsymbol{v}|\boldsymbol{y}^\star, \boldsymbol{X}, \boldsymbol{\beta}, \lambda) \propto \pi(\boldsymbol{v})I\left\{v_g^{\frac{1}{\varsigma}} > \|\lambda_g\beta_g\|\right\}$$

$$\boldsymbol{v}|\boldsymbol{y}^\star, \boldsymbol{X}, \boldsymbol{\beta}, \lambda_1, \ldots, \lambda_G, \varsigma \sim \prod_{g=1}^{G} Exponential(1)I\left\{v_g^{\frac{1}{\varsigma}} > \|\lambda_g\beta_g\|\right\}, \tag{6}$$

4- The full conditional posterior distribution of $\pi(\sigma^2)$ is

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

$$\sigma^2|\boldsymbol{y}^\star, \boldsymbol{X}, \boldsymbol{\beta} \sim Inverse\ Gamma\left(\frac{n}{2} + r, \frac{1}{2}(\boldsymbol{y}^\star - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y}^\star - \boldsymbol{X}\boldsymbol{\beta}) + b\right), \tag{7}$$

5- The full conditional posterior distribution of $\pi(\lambda)$ is

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

$$\propto \lambda_g^{\frac{m_g}{\varsigma}} \lambda_g exp^{-\theta\lambda_g} \prod_{g=1}^{G}\left\{\lambda < \frac{v_g^{\frac{1}{\varsigma}}}{\|\beta_g\|}\right\}$$

$$\lambda_g \propto \lambda_g^{\frac{m_g}{\varsigma}+1} exp^{-\theta\lambda_g} \prod_{g=1}^{G}\left\{\lambda < \frac{v_g^{\frac{1}{\varsigma}}}{\|\beta_g\|}\right\}$$

$$\lambda_g \sim gamma\left(\left(\frac{m_g}{\varsigma} + 1\right), \theta\right), \tag{8}$$

where I(.) is an indicator function.

3.3 Posterior Computation

In the section, following Mallick and Yi [26], we develop a Gibbs sampling algorithm to update the latent variables and the other parameters, according to the following steps:

i.   Generate $y_i^\star$

$$y_i^\star|y_i,\boldsymbol{\beta}\sim\begin{cases}\gamma(y_i) & if \quad y_i>c\,,\\N(\boldsymbol{x_i'\beta},\sigma^2) & otherwise,\end{cases}$$

ii.   Generate v_g from the full conditional distribution,

$$Generate\ v_g^\star\sim Exponential(1).$$
$$v_g=v_g^\star+\left\|\lambda_g\beta_g\right\|_1^\varsigma,\quad g=1,\dots,G\,.$$

iii.   Generate $\boldsymbol{\beta}$ from the multivariate normal distribution with mean $(\boldsymbol{X'X})^{-1}\boldsymbol{X'}y^\star$ and the variance is $\left(\sigma^2(\boldsymbol{X'X})^{-1}\right)$.

iv.   Generate σ^2 from the Inverse Gamma distribution with the shape parameter $\frac{n}{2}+r$ and the rate parameter $\frac{(y^\star-\boldsymbol{X\beta})'(y^\star-\boldsymbol{X\beta})}{2}+b$.

v.   Generate λ_g from Gamma distribution, with the shape $\left(\frac{m_g}{\varsigma}+1\right)$ and rate (θ).

4. Simulation Study

Here, we carry out Monte Carlo simulations to demonstrate the performance of the proposed method for Bayesian group bridge regression for left censored data (BGBRLC) . The BGBRLC is compared with the frequentist left-censored regression (FLCR), Bayesian regression for the left censored data (BRLC), Bayesian Lasso regression for the left censored data (BLRLCR) and Bayesian group Lasso regression for left censored data (BGLRLC) . These methods are evaluated based on the median of mean absolute deviations (MMAD) over 1000 simulations. The convergence of the BGBRLC algorithm is checked by trace plots and the histograms of the posterior samples for the regression parameters. The data in the simulations are simulated by

$y_i=\max\{y_i^\star,0\},\quad i=1,2,\dots,n,$

where $y_i^\star=\boldsymbol{x_i^t\beta}+\boldsymbol{\varepsilon_i}$ and $\varepsilon_i\sim N(0,4)$. We generate 50 observations from the above mode, where x_i represents the ith row vector of 8 predictors in the matrix *X*. The rows of X are simulated independently from $N(0,\Sigma)$ where the (i,j)th element of Σ is 0.95. The true regression coefficients, including the intercept term, are $\beta=\big((5,-5,\ 0),\ (0,\ 0,\ 0),(0,\ 0,\ 0)\big)$ which are divided into three groups; $(5,-5,\ 0),\ (0,\ 0,\ 0)$ and $(0,\ 0,\ 0)$. The results of MMAD and SD are summarized in Table- 1, which shows that the proposed method out-performed the other approaches. We also notice from Table-2 that the proposed approach produces results that are much closer to the true regression coefficients as compared to those produced by the other methods.

Table 1-MMADs for Simulation 1

Method	MMAD	SD
BGBRLC	0.6838	0.1256
BGLRLC	0.6906	0.1302
FLCR	0.6981	0.1408
BRLC	0.7002	0.1448
BLRLCR	1.1034	0.7502

The bold numbers correspond to the smallest MMAD.

Table 2-Parameter estimations for simulation 1

Method	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8
β (True)	5	-5	0	0	0	0	0	0	0
BGBRLC	4.828	-5.117	0.359	0.244	-0.748	1.176	-0.661	0.552	-0.811
BGLRLC	4.827	-5.742	0.545	0.817	-0.781	1.162	-0.621	0.596	-0.783
FLCR	4.868	-5.627	0.546	0.861	-0.725	1.18	-0.689	0.49	-0.759
BRLC	4.814	-5.763	0.527	0.886	-0.753	1.24	-0.755	0.528	-0.739
BLRLCR	4.788	-4.465	0	0	0	0	0	0	0

We summarize The trace plots for the simulation study in Figure-1, which shows that the samples of the BGBRLC method very readily traverse the posterior space very.
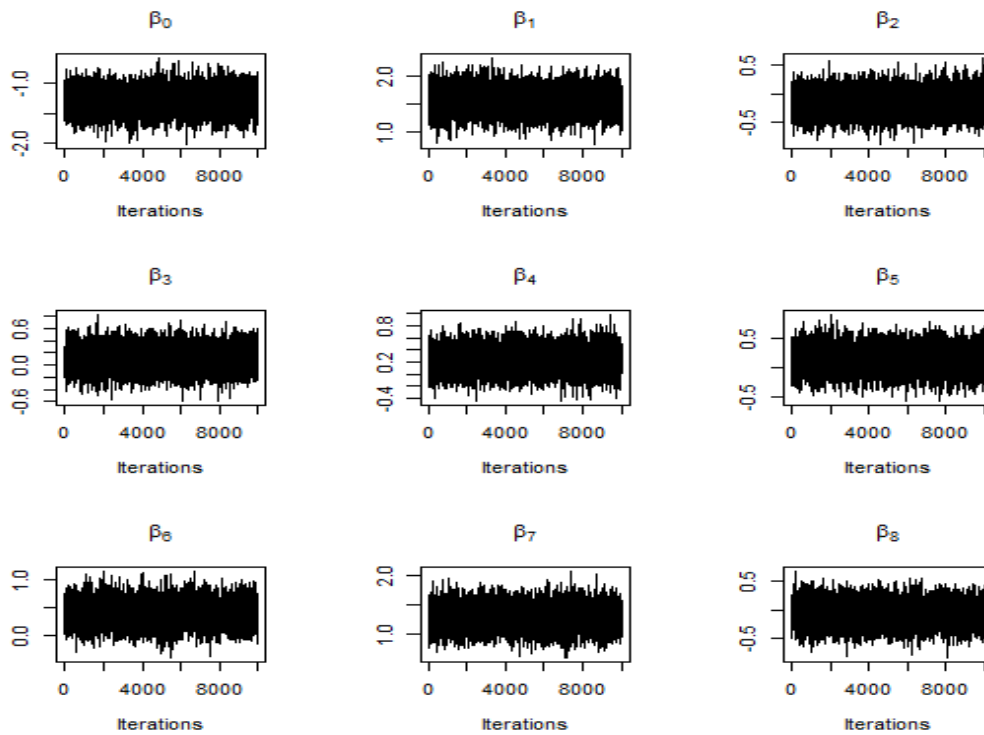

Figure 1-Trace plots for the variables in Simulation study.

In Figure-2, the posterior histograms of our proposed approach show that the conditional posterior for each parameter is the desired stationary truncated univariate normal distributions.
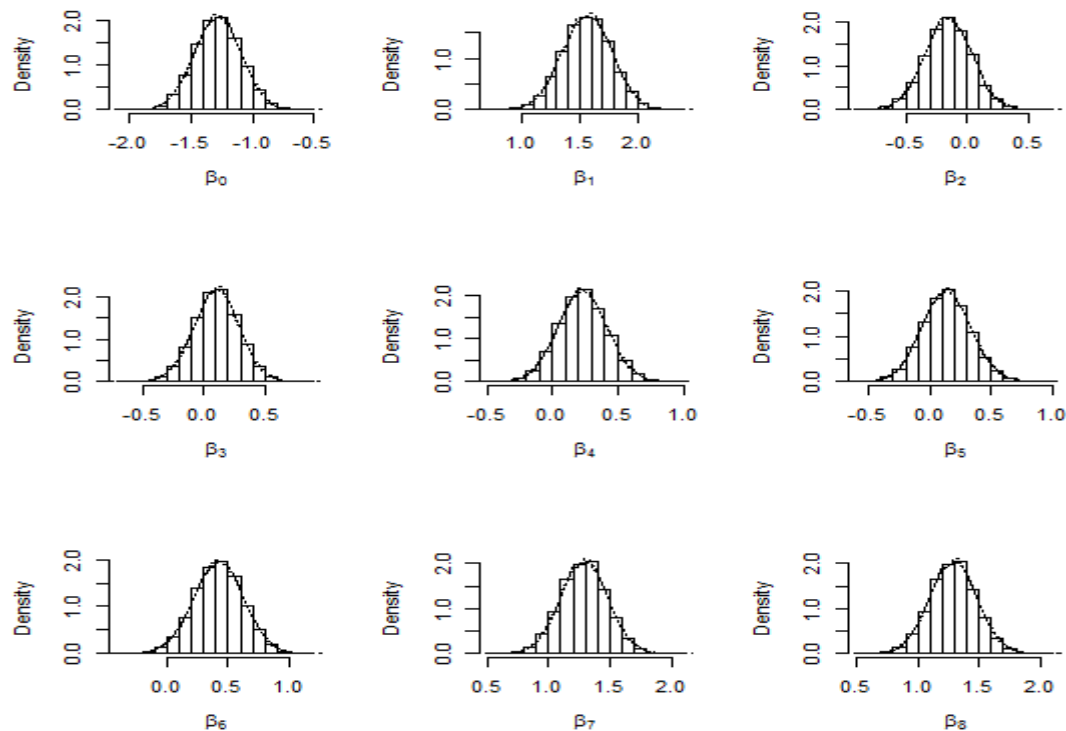

Figure 2-Posterior histograms for the simulation study.

5. Real Data

Here, the proposed approach is illustrated with the data of active sperms. This dataset has 200 observations on 8 variables. The response variable is the count of active sperms, while the other seven variables are covariates, as shown below.

y is the count of active sperm, the normal sperm count is 60-150 *1000000.

x_1 (Testosterone) the normal level of testosterone in blood 8.2-34 n.mol/l.

x_2 (Prolactin) the normal level of prolactin it is less than 1.5-19 ng/ml.

x_3 (Semen pH) the normal value of semen pH is 7.1-8.

x_4 (Semen viscosity) the normal value of semen viscosity in 20-30 minutes of gonorrhea.

x_5 (Sperm antibodies) if a person's blood contains sperm antibodies, then a value of 0 is given and, if doesn't, then the value is 1.

x_6 (varicocele) if a person suffers from varicocele, then a value of 0 is given and, if doesn't, the value is 1.

x_7 (Smoking) if a person smokes, then a value of 0 is given and, if doesn't, the value is 1.

Table 3-Posterior mean for parameter estimates of real data example.

Variables	BLRLCR	FRLCR	FLCR	BGLRLC	BGBRLC
Interceptس	3.385	2.9034	20.9419	22.0685	**22.3513**
X_1	1.0344	3.4665	1.3333	1.3663	**1.3594**
X_2	-0.9896	0.0731	-1.405	-1.4472	**-1.4624**
X_3	1.244	-0.0642	-1.9077	-2.1363	**-2.1561**
X_4	1.4815	0.5876	1.2577	1.2593	**1.2516**
X_5	-0.1198	-4.6287	-9.6294	-9.5889	**-9.5894**
X_6	5.954	0.2231	20.1326	20.1301	**20.4669**
X_7	1.9526	-0.8898	5.9632	6.1189	**6.2577**

In Table-3, we listed the results of the real data example. To evaluate the methods, the DIC was computed for the five approaches (BGBRLC, BGLRLC, FLCR, FRLCR, BLRLCR) and the values were 1710.291, 1713.448, 1822.805 , 1817.453 and 1836.229, respectively. The DIC results show that the BGBRLC performs better than the other approaches.

6. Conclusions

In this paper. We have analyzed the real data using the R software and applied the simulation examples. We compared our proposed method with other methods. The results showed that the new method performs better than some existing methods. The new method can be easily extended to other approaches such as Bayesian group bridge for binary data and Bayesian group bridge for right and interval censored data.

References

1. Mallows, C. L. **1973**. Some comments on C p. *Technometrics*, **15**(4): 661-675.
2. Woodroofe, M. **1982**. On model selection and the arc sine laws. *The Annals of Statistics*, **10**(4): 1182-1194.
3. Nishii, R. **1984**. Asymptotic properties of criteria for selection of variables in multiple regression. *The Annals of Statistics*, 758-765.
4. Akaike, H. **1974**. A new look at the statistical model identification. *IEEE transactions on automatic control*, **19**(6): 716-723.
5. Javed, F., & Mantalos, P. **2013**. GARCH-type models and performance of information criteria. *Communications in Statistics-Simulation and Computation*, **42**(8): 1917-1933.

6. Schwarz, G. **1978**. Estimating the dimension of a model Ann Stat 6: 461–464. *Find this article online*.
7. George, E. I., & McCulloch, R. E. **1993**. Variable selection via Gibbs sampling. *Journal of the American Statistical Association*, **88**(423): 881-889.
8. Hans, C. **2009**. Bayesian lasso regression. *Biometrika*, **96**(4): 835-845.
9. Li, Q. and Lin, N. **2010**. The Bayesian elastic net. *Bayesian Analysis* **5**(1): 151–170.
10. Alhamzawi, R., and Yu, K. **2014**. Bayesian Lasso-mixed quantile regression. *Journal of Statistical Computation and Simulation*, **84**(4): 868-880.
11. Tibshirani, R. **2011**. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **73**(3): 273-282.
12. Liu, X., Wang, Z., and Wu, Y. **2013**. Group variable selection and estimation in the tobit censored response model. *Computational Statistics & Data Analysis*, **60**: 80-89.
13. Alhamzawi, R., Yu, K. and Benoit, D. F. **2012**. Bayesian adaptive Lasso quantile regression. *Statistical Modelling*, **12**(3): 279-297.
14. Mallick, H., and Yi, N. **2014**. A new Bayesian lasso. *Statistics and its interface*, **7**(4): 571.
15. Xu, X. and Ghosh, M. **2015**. Bayesian variable selection and estimation for group lasso. *Bayesian Analysis*, **10**(4): 909-936.
16. Alhamzawi, R. and Ali, H. T. M. **2018**. Bayesian tobit quantile regression with penalty. *Communications in Statistics-Simulation and Computation*, **47**(6): 1739-1750.
17. Hoerl, A. E. and Kennard, R. W. **1970**. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, **12**(1): 55–67.
18. Frank, I. and Friedman, J. H. **1993**. A statistical view of some chemometrics regression tools (with discussion). *Technometrics*, **35**(2): 109–135.
19. Tibshirani, R. **1996**. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*: Series B (Statistical Methodology), **58**(1): 267–288.
20. Yuan, M. and Lin, Y. **2006**. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society:Series B (Statistical Methodology)*, **68**(1): 49-67.
21. Kim, Y., Kim, J., and Kim, Y. **2006**. Blockwise sparse regression. *Statistica Sinica*, **16**(2): 375.
22. Huang, J., Ma, S., Xie, H. and Zhang, C. **2009**. A group bridge approach for variable selection. *Biometrika*, **96**(2): 339–355.
23. Xu, Z., Zhang, H., Wang, Y., Chang, X. and Liang, Y. **2010**. L1/2 regularization. *Science China Information Sciences*, **53**(6): 1159–1169.
24. Polson, N. G., Scott, J. G., and Windle, J. **2014**. The bayesian bridge. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **76**(4): 713-733.
25. Aljanabi, S., and Alhamzawi, R. **2020.** Bayesian group Lasso regression for left-censored data. *Periodicals of Engineering and Natural Sciences*, Accepted paper.
26. Mallick, H., and Yi, N. **2017**. Bayesian group bridge for bi-level variable selection. *Computational statistics & data analysis*, **110**: 115-133.
27. Kinney, S. K., and Dunson, D. B. **2007**. Fixed and random effects selection in linear and logistic models. *Biometrics*, **63**(3): 690-698.
28. Alhamzawi, R., Yu, K., Vinciotti, V., and Tucker, A. **2011**. Prior elicitation for mixed quantile regression with an allometric model. *Environmetrics*, **22**(7), 911-920.
29. Alhamzawi, R., and Ali, H. T. M. **2018**. The Bayesian elastic net regression. *Communications in Statistics-Simulation and Computation*, **47**(4): 1168-1178.