Republic of Iraq
Ministry of Higher Education & scientific research
University of Al-Qadisiyah
College of Administration and Economics
Department of Statistics

Bayesian Lasso Left and Right Censored Regressions with an Application

A Thesis Submitted to the Council of the College of Administration and Economics at University of Al-Qadisiyah as a Partial Fulfillment of the Requirements for the Degree of Master in Statistics Sciences

By

Zainab Saadoon Oleiwi

Supervised by

Assist. Prof. Dr. Bahr Kadhim Mohammed

2023 *A.D* ……………………………..………..………………………… 1444 *A.H*

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

﴿وَلَوْلَا فَضْلُ اللَّهِ عَلَيْكَ وَرَحْمَتُهُ لَهَمَّت طَّائِفَةٌ مِّنْهُمْ أَن
يُضِلُّوكَ وَمَا يُضِلُّونَ إِلَّا أَنفُسَهُمْ ۖ وَمَا يَضُرُّونَكَ مِن شَيْءٍ ۚ
وَأَنزَلَ اللَّهُ عَلَيْكَ الْكِتَابَ وَالْحِكْمَةَ وَعَلَّمَكَ مَا لَمْ تَكُن تَعْلَمُ ۚ
وَكَانَ فَضْلُ اللَّهِ عَلَيْكَ عَظِيمًا﴾

صَدَقَ اللهُ العَلِيُّ العَظِيم

«سورة النساء: الآية 113»

Acknowledgments

First, I praise to be to Almighty Allah for helping me to achieve this work.

I would like to thank Dr. Bahr Kadhim Mohammed my supervisor who help me in completing this study through his great efforts and continuous guidance and assessment. Moreover, I would love to send my thanks and appreciation to the unyielding person Prof. Dr. Rahim Al-Hamzawi who supported me on the scientific and analytical side. I also extend my deepest thanks, gratitude, and appreciation to all the professors in the Statistics department.

I would also like to thank my committee members, for the time and effort they spent to read and comment on my work.

To Zahraa Kadhim, thank you for continuously pushing me forward when I was too tired. Thank you for always reminding me of how strong I am.

Finally, I truly owe my deep gratitude to all of my family and my parents in particular, for their unlimited support and encouragement during my studies period.

✍ *Zainab Al-Husseini*

Dedication

To My Family

Abstract

The regularization methods, and particularly Lasso method is widest and most methods in the issue of selecting variables in regression analysis. This thesis focused on studying the regulating called LASSO method through Bayes theory, where three models have been employed to represent the prior distribution of the regression parameter (Laplace distribution). The first model assumed using of scale mixture normal distribution mixing with the exponential distribution. The proposed model is second representation of scale mixture uniform distribution mixing with standard exponential distribution. The third scale mixture uniform distribution mixing with gamma distribution. The three models have applied in the censored regression especially, the left-censored regression and the right-censored regression. The three models were applied according to the Bayes approach through the implementation of the Gibbs sampler algorithm through the simulation method by assuming three trials for simulation and using the R programming language, where we performed simulation experiments of different sample sizes and different variances values for errors, and by using the median criterion of the mean squared error and then judging the performance of the different methods. In order to demonstrate the efficiency of the proposed method, this method was employed on real data that represents a sample taken from previous research. This sample is characterized as the response variable is a censored variable from the left. This data is available in the R language. The second practical application is for data in which the response variable is characterized as a censored variable from the right. Where the sample represents the level of urea in the blood with a group of explanatory variables. The results showed in each of the two applied examples that the proposed method has a high predictive ability of the model compared to other methods from the principle of prediction accuracy and selection of variables.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviations	**Meaning**
BLLCR	Bayesian LASSO Left Censored Regression
BLRCR	Bayesian LASSO Right Censored Regression
BLTR	Bayesian LASSO Tobit Regression
BTAL	Bayesian Tobit Adaptive LASSO
BTL	Bayesian Tobit LASSO
LASSO	Least Absolute Shrinkage and Selection Operator
MCMC	Markov Chain Monte Carlo
MCP	Minimax Concave Penalty
MLE	Maximum Likelihood Estimation
NBLLCR	New Bayesian LASSO Left Censored Regression
NBLRCR	New Bayesian LASSO Right Censored Regression
OLS	Ordinary Least Squares
RCR	Right Censored Regression
SCAD	Smoothly Clipped Absolute Deviation
SSE	Sum of Squares of Error
SSVS	Stochastic Search Variable Selection
VS	Variable Selection

Chapter One

Introduction

and

Literature Review

1.1 Introduction

Regression analysis is a statistical method for analyzing data that includes two or more variables when the goal is to discover the relationship. Regression analysis is the most common statistical method, as it is widely used in various sciences to describe the relationship between variables in the form of an equation. It enables the researcher to analyze relationships and predict the values of the dependent variable by knowing the values of the independent variables. Of course, the prediction of using the independent variables together is better than predicting the use of any of them separately, provided that the correlation between these independent variables is low and the correlation of each of them with the dependent variable is high.

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

The use of classical methods such as the method of least squares, which is widely used to estimate the multiple linear regression model, to obtain the estimations of the unknown parameters is sometimes defective and gives weak estimates, especially when the sample size is small $(n < p)$. Therefore, regularization methods have been proposed such as the LASSO method and others to obtain the best estimate for the unknown parameter from between all possible estimates.

The Bayesian analysis became more popular because of the development of computer approximations to integrals and the appearance of easy-to-use programs to implement these arithmetic operations. And the use of Bayesian statistics was not limited to the development of wide research in Bayesian methodology, but also in the use of Bayesian methods to process many problems in applied domains.

Censored regression models, often called (limited regression models), are characterized by the fact that the dependent variable is determined in some way that depends on the nature of the phenomenon. They differ from the truncated regression model. The regression model is called truncated when observations outside a specific range are missing for the dependent variable and the independent variables, while the regression model is called limited when the independent variables are observed over an open range and the dependent variable is within a specific range.

In this thesis, 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 and right sides.

This thesis is divided into four chapters, the first chapter includes an introduction, the research problem, the objective of the research, and the literature review, while the second chapter contains some critical basic concepts in this thesis and the proposed method. The third chapter includes the experimental and application sides, while the fourth chapter included a review of the most important conclusions and recommendations.

1.2 Thesis Problem

When the data set is restricted at specific value the upper or lower of data set is censored. Censored regression is used for similar data with same reason. The classical methods such as OLS, MLE is considered bias and inconsistent when the data is censored, there for alternative methods proposed in literature. Bayesian approach are widely used due its estimation accuracy. Unfortunately, when the independent variables exceed the number of observations or large enough, Bayesian methods are not reliable. Variable selection method is preferred with such case. It is observed the previous studies were not considered the left and right side. Particularly, with mixture methods.

1.3 The Objective of the thesis

This thesis seeks to satisfy two objectives as follows:

1. Study the Bayesian LASSO variable selection procedure based on three exact models proposed (normal and exponential, uniform and gamma, uniform and standard exponential).
2. Employed the three Bayesian LASSO models in the previous step the left-censored regression and right-censored regression.

1.4 Literature review

Due to a censoring mechanism in some data sets, we do not see numbers upper or lower than a specific size, when the censoring mechanism is used, the observed data consists of a mix of measurements of some latent variable and observations that occur as a result of the censoring process, y_i is censored.

- → y_i is left-censored or censored from lower if $y_i \geq y_L$ for every y_i.
- → y_i is right-censored or censored from upper if $y_i \leq y_U$ for every y_i.

Variable selection process refers to a collection of tasks in which the goal is finding the optimal subset of relevant variables that can be utilized to make precise modifications to the outcomes of a given dependent variable. Identifying essential and influential factors on the dependent variable might be challenging when the number of variables is too great. As a result, in the data analysis, the (VS) has deemed important. For this reason, many researchers concentrate on classical approaches to find the best model.

Efroymson (1960) provided a method of deleting non-significant variables step by step. In an automated approach for selecting independent variables, each stage considers a variable for addition to or deletion from the collection of independent variables based on some predetermined criterion. It is a mixture of forward selection (FS) and backward elimination (BE).

Mallows (1973) put a statistic known as Mallows C_p statistic and its basis is to evaluate the appropriateness of least squares to models with normal errors and constant variance, it is defined as follows:

$$C_p = \frac{RSS_p}{S^2} + 2p - n,$$

where RSS_p is the residual sum of squares from a model, and S^2 is variance estimate of σ^2, n is the sample size, and p is the total of parameters. The model in which the C_p value is small is the best model or the most exact model.

Akaike (1974) proposed the Akaike Information Criterion method. It's one of the most typical ways to choose variables. An AIC procedure's value can be used to compare various models. The best model is the one with the lowest value, the formula for Akaike's process is:

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

where L represents the parameter's maximum likelihood estimation value (MLE), and p represents the number of estimated parameters.

Hocking (1976) proposed all possible regressions method, which depends on finding all possible models (2^p) that contain one independent variable and even (p) of independent variables. Models contain the same number of independent variables are placed in totals on it, the number of totals is (p), to choose the required model, this method is characterized by reliance on experience and the use of relevant analytical results, the normal of the data, and some measures of statistical differentiation, such as the mean of squares of error (MSE) and the Mallows C_p statistic.

Schwarz (1978) suggested the Bayesian information criterion (BIC), is the development of AIC, and the BIC procedure formula is:

$$BIC = -2 \ln L + p \ln n,$$

where n is the sample size.

It is a criterion for selecting a model from a specific set of models, the preferred model is the one that has the least BIC. This method solves the (AIC) problem by selecting a model with good properties, but it has some drawbacks:

1. The BIC criterion has an approximation problem, the BIC is only valid if ($n > p$) of model parameters (p).
2. This criterion is incapable of dealing with large sets of models.

Recently, a lot of effort has been exerted to develop 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.

George and McCulloch (1993) introduced one of the most popular Bayesian variable selection approaches, and for is Stochastic Search Variable Selection (SSVS). The goal of this procedure is to develop a probabilistic method for selecting promising subsets. This procedure involves embedding the model into a hierarchical normal mixture model. This algorithm is the first to introduce some of the basic principles of modern Bayesian variable selection methods. usually takes a long time to select the significant variables due to the large number of variables involved.

Donoho and Johnstone (1994) for the first-time regularization techniques were used (VS), where they proposed the soft-threshold estimator to obtain a smooth estimation of a function in the wavelet approximation. The results proved the perfect estimation of the function. Tibshirani developed it after that a year 1996 to obtain an estimate of the coefficients themselves and not just the function.

Tibshirani (1996) proposed a new technique for estimating linear models called LASSO, which is an abbreviation for "Least Absolute Shrinkage and Selection Operator", that reduces the sum of squares of residuals subject to the sum of the absolute value of coefficients that are less than a fixed value, and because of the nature of this constraint, it tends to produce some coefficients that are exactly equal to zero and thus give interpretable models. Using a simulation experiment, showed LASSO has a properties combination of the subset selection method and the ridge regression method.

Fan and Li (2001) suggested the SCAD penalty. SCAD estimator has oracle properties, and it has many desirable properties including unbiasedness, sparsity, and continuity. Through simulation, it was shown that the SCAD method works positively compared to other regularization methods.

Zou and Hastie (2005) proposed the Elastic-Net as the development of the LASSO technique by adding the ridge penalty parameter (λ_2) with the LASSO penalty parameter (λ_1). This technique has the performance the best in estimating coefficients, selecting variables, and dealing with the problem of high correlations between independent variables. As well as, this technique is essential in dealing with the process of choosing variables when the number of variables is much greater than the sample size ($p \gg n$). On the other hand, this method preserved the sparsity property of the LASSO method.

Zou (2006) introduced an update to the LASSO technique, and it is Adaptive LASSO, the concept of this technique is to assign various adaptive weights to a variety of parameters in the penalty function, which results in a penalty reduction for parameters that are close to zero. As a result, its estimates are consistent and unbiased. So, it has found that the adaptive LASSO has the characteristic of the oracle, unlike LASSO and Elastic-Net, it does not have this feature.

Zou and Zhang (2009) introduced the adaptive Elastic-Net technique to discuss the issue of estimating and selecting variables in the case presence of high-dimensional data and the problem of multicollinearity, where the ridge penalty parameter (λ_2) is combined with the weighted LASSO penalty parameter (λ_1), assuming that the LASSO penalty parameter is different for each parameter of the model, they proved that this technique possesses the oracle property. Simulation results showed the efficiency of this method in the case of multicollinearity problems compared with other methods of regularization.

Zhang (2010) suggested the MCP penalty, a method of penalty variable selection in high-dimensional linear regression. MCP penalty that is fast, continuous, almost unbiased, and accurate in penalized variable selection. Subset selection in it is unbiased. However, it is computationally costly. Without assuming the strong irrepresentable requirement required by the LASSO, the MCP has a high likelihood of matching the signs of the unknowns and hence proper selection. The results proved the high accuracy of variable selection and the computational efficiency of this method.

The following are some previous studies related to the thesis topic.

Tobin (1958) suggested the (MLE) was used to estimate the parameters of the regression model using the classical methods, and that is by studying the household spending on durable goods, where he noticed that the data of the dependent variable contains an important characteristic that makes most of the observations have zero expenses and have income, not to mention that this is a violation of one of the assumptions of the (OLS) method.

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. The simulation results showed that the results of the Bayesian LASSO are similar to the results of the ordinary LASSO.

Hans (2009) compared the standard LASSO regression and the Bayesian LASSO regression, when independent double-exponential prior initial distributions are applied to the regression parameters. He found the standard LASSO method is not necessarily in agreement with the predictions of the Bayesian method.

Mallick and Yi (2014) introduced a new Bayesian LASSO method to solve the LASSO problem in representing the density function of the Laplace distribution, where he proposed the use of a uniform distribution of the scale mixture with a specific gamma $(2, \lambda)$. The simulation results proved the high predictive ability and variable selection in the models.

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. The real data results showed the superiority of the new Bayesian method.

Alhusseini et al. (2020) proposed to formulate a new hierarchical model for (BTL) and (BTAL) regression taking into account Laplace distribution taken the proposal by (Flaih et al., 2020).

In this study, we introduced a mixed representation of the Laplace distribution by performing transformations and mathematical operations which was obtained through the uniform continuous distribution $(\frac{-\sigma^2}{\lambda}, \frac{\sigma^2}{\lambda})$ multiplied by the standard exponential distribution (z), has employed by Bayesian LASSO regression for left and right-censored data. The results show that the proposed method performs very well compared with the classical methods for left and right-censored data. We have proposed a Bayesian regularization method for left and right-censored responses based on the Bayesian regularized method of Park & Casella (2008). Also, we have proposed a Bayesian regularization method for left and right-censored responses based on the Bayesian regularized method of Mallick & Yi (2014). In practice, the results show that the proposed methods perform very well in terms of convergence.

Chapter Two

Theoretical Side

2.1 Preface

In this chapter, consisting of six parts, some basic concepts of the theoretical aspects of the methods applied in this thesis will be clarified. The first part deal with a simple review of the linear regression model, which includes two types, simple and multiple, and the (OLS) method. In the second part, we will explain the concept of limited dependent variables, which include censored data from the left side, and censored data from the right side. Additionally, the interval-censored regression model. The third part will present a brief overview of the regularization methods, and an accurate presentation of the LASSO method in selecting variables, their advantages, and disadvantages. The fourth part will explain the Bayesian LASSO estimation method. The chapter also includes the suggested methods.

2.2 Linear Regression Models

Many studies and researches depend on advanced methods in order to obtain highly accurate results, Statistics and its related branches have a great impact on building best models and analyzing data through them to reach proper decisions.

Regression analysis is the most important branch of statistics, which is concerned with building the mathematical relationship between the dependent variable and the independent variables, this relationship is represented as a linear formula called the regression equation, as its accuracy depends on the correctness of estimating its parameters, which requires the availability of the analysis hypotheses.

Regression analysis is divided into two main parts: linear and nonlinear regression. The general form of a linear regression model is:

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

where $\boldsymbol{y}$ is a vector of $(n \times 1)$ of responses, $\boldsymbol{X}$ is a matrix with dimension $(n \times p)$ of predictors, $\boldsymbol{\beta}$ is a vector $(p \times 1)$ of unknown parameters, and $\boldsymbol{\epsilon}$ is a vector $(n \times 1)$ of random errors.

2.2.1 Ordinary Least Squares method

The ordinary least squares method is one of the most important and common methods for estimating the parameters of the linear regression model. This method is characterized by good characteristics that made it one of the best and most widely used methods, including unbiased, consistency, efficiency, and sufficiency, but when the assumptions of the analysis are not available, their estimates become biased and inconsistent. This method is based on the principle of minimizing the sum of the squares of errors to the least possible. (Balestra, 1970). The mathematical formula for obtaining the (OLS) estimator for the parameters of the regression model using matrices is as follows:

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

The (OLS) method gives the best linear unbiased estimate (BLUE) with the least variance of the model parameters.

2.3 Limited Dependent Variables

The limited dependent variables in regression models mean that there is a limit to the dependent variable and some independent variables reach that limit where the dependent variable is observed within a specified range, while the independent variables are observed within an open range. Limited dependent variable models address two issues important censored and truncation. A limited dependent variable y_i^* is a continuous variable with a lot of independent variables repeated at the lower or upper bound. (Maddala, 1987) James Tobin is one of the first to write about censored data and censored models, and he called them "limited dependent variables". (Tobin, 1958)

The censored regression model is one of the most common statistical models used in many studies and research especially when the data is restricted (determined) in one part and free (undetermined) in another, which makes it difficult to use and apply classical regression models to that data because it will be biased towards zero and inconsistent. This data is called censored data. The dependent variable might be censored left-censored or right-censored, or both (interval).

2.3.1 Left-Censored Data

A data point is lower than a particular value but is unknown. If the latent variable y_i^* is upper the limit and the limit for the censored observations, the real value for the dependent variable y_i is observed. The dependent variable y_i is a continuous variable with no zero. If the dependent variable's y_i actual values are greater than the lower limit, they are observed.

$$y_i = \begin{cases} y_i^* & if \quad y_i^* > y_L, \\ y_L & if \quad y_i^* \leq y_L, \end{cases} \tag{2.1}$$

where y_L is the restriction point.

or the dependent variable y_i can also write it like this:

$$y_i = max(y_i^*, y_L),$$

The censored regression model is the Tobit model when ($y_L = 0$). y_L is usually zero, but not always.

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

we can express the Tobit model as:

$$y_i = \begin{cases} y_i^* & if \quad y_i^* > 0, \\ 0 & if \quad y_i^* \leq 0, \end{cases}$$

or can be expressed:

$$y_i = max(y_i^*, 0).$$

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). The Tobit model is a mixture of the Probit model and the regression model. (Fernando, 2011)

For example, a person's age is considered one of the censored data from the left side, so no person's age is less than zero.

2.3.2 Right-Censored Data

A data point is upper than a particular value but is an unknown. If the latent variable y_i^* is lower than the limit and the limit for the censored observations, the real value for the dependent variable y_i is observed. If the dependent variable's actual values are less than the upper limit, they are observed. (Koul et al., 1981; Kohler et al., 2002)

$$y_i = \begin{cases} y_i^* & if \quad y_i^* < y_U, \\ y_U & if \quad y_i^* \geq y_U, \end{cases} \tag{2.2}$$

where y_U is the restriction point.

or can also write it like this:

$$y_i = min(y_i^*, y_U).$$

2.3.3 Interval-Censored Data

The data observed in this type of censoring time interval is located somewhere in a time interval between two points. That is, there are unknown values of the dependent variable y_i that exceed the lower and upper limit of the scale and are monitored at these limits.

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

where y_i^* is a latent dependent variable (unobserved). However, the dependent variable y_i (observed) is written as:

$$y_i = \begin{cases} y_L & if \quad y_i^* \leq y_L, \\ y_i^* & if \quad y_L < y_i^* < y_U, \\ y_U & if \quad y_i^* \geq y_U, \end{cases}$$

Here y_L is the lower limit and y_U is the upper limit of the dependent variable. The dependent variable is not left-censored or right-censored when $y_L = -\infty$ or $y_U = \infty$, and to estimate the model's upper limit, lower limit, or interval, values must be provided. Censored data consists of a large number of observations in which the dependent variable y_i takes one value or a limited number of values. (Henningvsen, 2010; Alan et al., 2014; Chay & Powell, 2001; Amemiya, 1973)

The standard Tobit model is one of the important and commonly used models in dealing with this type of data when the dependent variable is limited. This model describes the relationship between the dependent variable (positive) and the independent variables. The Tobit model deals with the observed data by dividing it into two parts and the model function is mixed so that the dependent variable has a specific distribution according to its range, and this model is of great importance in applied research such as medical and economic research. The Tobit model is the oldest censored regression model. Tobit model is a special case when the dependent variable is blocked from lower at zero ($y_L = 0$). (Kohler et al., 2002)

$$y_i = \begin{cases} 0 & if & y_i^* \leq 0, \\ y_i^* & if & 0 < y_i^* < y_U, \\ y_U & if & y_i^* \geq y_U, \end{cases}$$

y_i is observed if it exceeds zero.

An example of censored data from both sides, if we have a study on the monthly income of families and a limit has been set so that the income is not less than \$500 and not more than \$2000, here cases appear that have a monthly income that is less than the lower or greater than the upper limit, and they are recorded without specifying them exactly. These values are considered unknown and outside the range.

2.4 Regularization Methods

Regularization methods have received great interest in the statistical literature, for their high efficiency in selecting important variables and excluding unimportant variables in the regression model.

The classical estimation methods, such as the (OLS) method, perform poorly in the case where the number of independent variables is large. Although the (OLS) estimators give a small variance and zero bias. But here it will give a significant variance and a little bias, and will be difficult to interpret the regression models. Therefore, the researchers resorted to using regularizing methods to address this problem by adding a penalty function to the sum of the error squares. That is, some bias is introduced in order to reduce prediction variance and obtain accurate results.

Regularization methods were called this name because it regularizes between bias and variance, and it is also called the shrinkage methods because it works to shrink some parameters and make the others equal to zero.

2.4.1 The LASSO (Least Absolute Shrinkage and Selection Operator)

A method proposed by Tibshirani in 1996 works on selecting the variables and estimating parameters of the regression model at the same time. In contrast to (OLS), the LASSO estimation method is biased but more accurate according to the following formula:

$$\hat{\boldsymbol{\beta}}_{LASSO} = \underbrace{arg\ min}_{\boldsymbol{\beta}}\ SSE \quad subject\ to \quad \sum_{j=1}^{p} |\boldsymbol{\beta}_j| \leq t,$$

This method reduces the sum of the squares of the errors and is subject to the constraint ("the sum of the absolute values of the parameters being less than a certain constant, let it be t"), where a penalty function λ is added equal to the absolute value of the parameters as follows:

$$\hat{\boldsymbol{\beta}}_{LASSO} = \underbrace{arg\ min}_{\boldsymbol{\beta}}\ \|\boldsymbol{y} - \boldsymbol{X\beta}\|_2^2 + \lambda \sum_{j=1}^{p} |\boldsymbol{\beta}_j|.$$

λ is the tuning parameter or penalty parameter, and its ranges between $[0, \infty)$. If it is ($\lambda = 0$), then no parameter is deleted and we get the (OLS) estimators. Either if it is ($\lambda = \infty$) then all parameters are deleted, and thus the bias amount increases, but when the value of λ decreases the variance increases. Because of the normal of this constraint, the LASSO method reduces some estimated parameters and makes other parameters equal to zero, thus reducing the variance of errors, and it becomes easy to interpret the regression model (Tibshirani, 1996; Savin, 2013).

The LASSO penalty function is also called L_1 – norm.

2.4.1.1 Advantages of LASSO technique:

1. The LASSO technique can be applied in various statistical models.
2. 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.
3. 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.

2.4.1.2 Disadvantages of LASSO technique:

1. In the case of the multicollinearity problem, LASSO does not have the optimal performance, where the ridge technique is the best.
2. When the number of independent variables (p) is greater than the sample size (n), the process of selecting variables will be restricted so that it does not exceed the sample size (n).
3. It doesn't have oracle properties.

2.5 Bayesian LASSO Estimation

The Bayesian method is one of the important methods used in estimating the parameters of the linear model because of its importance in finding accurate estimates of the parameters and in overcoming the problems facing the estimation process using classical methods. This large importance of the Bayesian approach in estimating regression models has made its use common in recent research and studies, where the Bayesian approach provides an efficient method in the case of small samples, also in overcoming some of the difficulties that accompany the process of estimating model parameters using the classical approach (Rencher & Schaalje, 2008).

The Bayesian method in the LASSO technique has become of major interest in recent years because of its great importance in inference.

LASSO is estimated by the classical method and the Bayesian method, where the LASSO penalty parameter is estimated by the cross-validation in the classical method, which often results in the selection of variables in the model being inconsistent. (Chand, 2012). In this thesis, we will estimate it using the Bayesian method through three models for the left-censored and three for the right-censored.

2.5.1 Scale Mixture of Normal

In 1974, Andrews and Mallows introduced a useful representation of calculations of MCMC, this representation is the scale mixture of normal mixing with exponential distribution. In 2008, Park and Casella employed this scale mixture as the Laplace distribution in the Bayesian LASSO linear regression,

$$\frac{a}{2}e^{-a|z|} = \int_0^\infty \frac{1}{\sqrt{2\pi s}} e^{-\frac{z^2}{2s}} \frac{a^2}{2} e^{-\frac{a^2 s}{2}} ds, \quad a > 0 \tag{2.3}$$

In this study, we will employ the scale mixture in (2.3) to develop the Gibbs sampler algorithm by proposing new hierarchical prior model for the left-censored data. If we substitute $a = \frac{\lambda}{\sigma^2}$, $z = \beta$, $s = \tau^2\sigma^2$, $ds = \sigma^2 d\tau^2$. Then the formulation (2.3) can be rewritten as follows:

$$\frac{\lambda}{2\sigma^2} e^{-\frac{\lambda}{\sigma^2}|\beta|} = \int_0^\infty \frac{1}{\sqrt{2\pi\tau^2\sigma^2}} e^{-\frac{\beta^2}{2\tau^2\sigma^2}} \frac{\lambda}{2\sigma^2} e^{-\frac{\lambda^2\tau^2}{2\sigma^2}} d\tau^2 \tag{2.4}$$

the formula (2.4) is using the normal distribution mixing with exponential distribution.

2.5.2 The Hierarchical Priors Model of Left-Censored Data

The following hierarchical priors model suggested based on the linear left-censored structural equals (2.1), and based on the scale mixture (2.3):

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

$$\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\beta}, \sigma^2 \sim N(\boldsymbol{X\beta}, \sigma^2 \boldsymbol{I}_n),$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{y}^* - \boldsymbol{X\beta})'(\boldsymbol{y}^* - \boldsymbol{X\beta})\right],$$

$$\boldsymbol{\beta}|\boldsymbol{\tau^2} \sim \frac{1}{\sqrt{2\pi\sigma^2\boldsymbol{\tau^2}}} e^{-\frac{(\boldsymbol{\beta}-0)^2}{2\sigma^2\boldsymbol{\tau^2}}} \sim N\,(\boldsymbol{\beta}; 0, \sigma^2\boldsymbol{\tau^2}), \tag{2.5}$$

$$\boldsymbol{\tau^2} \sim \frac{\lambda^2}{2\sigma^2} e^{-\frac{\lambda^2\boldsymbol{\tau^2}}{2\sigma^2}},$$

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

$$\lambda^2 \sim Gamma(q\,, \theta).$$

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

2.5.3 The Gibbs Sampling Algorithm

In this study, we use the Gibbs sampler algorithm to implement the hierarchical prior model (2.5). The Gibbs sampler algorithm can sample the interested parameter values from the conditional distribution of any parameter given all the other parameters. Also, we use the inverse gamma distribution as prior density for σ^2 :

$$\pi(\sigma^2) = \frac{\alpha^h}{\Gamma(h)} (\sigma^2)^{-h-1} e^{-\alpha/\sigma^2} \;\; ; \;\; \sigma^2 > 0 \; (h > 0\,, \alpha > 0),$$

Now, we can write down the full joint density of the interested parameters

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

$$=\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}(\boldsymbol{y}^*-\boldsymbol{X\beta})'(\boldsymbol{y}^*-\boldsymbol{X\beta})}\frac{\alpha^h}{\Gamma(h)}(\sigma^2)^{-h-1}e^{-\frac{\alpha}{\sigma^2}}\frac{\theta^q}{\Gamma(q)}(\lambda^2)^{q-1} \tag{2.6}$$

$$e^{-\theta\lambda^2}\prod_{j=1}^{k}\frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}}e^{-\frac{(\boldsymbol{\beta}_j-0)^2}{2\sigma^2\tau_j^2}}\frac{\lambda^2}{2\sigma^2}e^{-\frac{\lambda^2\tau_j^2}{2\sigma^2}}.$$

Then full conditional posterior distributions are defined as follows:

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

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

2. The full conditional posterior distributions of $\boldsymbol{\beta}$ is defined by:

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

$$\propto\frac{1}{\sqrt{2\pi\sigma^2}}\;e^{-\frac{1}{2\sigma^2}(\boldsymbol{y}^*-\boldsymbol{X\beta})'(\boldsymbol{y}^*-\boldsymbol{X\beta})}\prod_{j=1}^{k}\frac{1}{\sqrt{2\pi\tau_j^2\sigma^2}}e^{-\frac{\beta_j^2}{2\tau_j^2\sigma^2}}$$

$$\propto e^{-\frac{1}{2\sigma^2}(\boldsymbol{y}^*-\boldsymbol{X\beta})'(\boldsymbol{y}^*-\boldsymbol{X\beta})-\sum_{j=1}^{k}\frac{\beta_j^2}{2\tau_j^2\sigma^2}}$$

$$=exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{y}^*-\boldsymbol{X\beta})'(\boldsymbol{y}^*-\boldsymbol{X\beta})-\frac{1}{2\sigma^2}\frac{\boldsymbol{\beta}'\boldsymbol{\beta}}{\boldsymbol{D_\tau}}\right]$$

$$=exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{y}^*-\boldsymbol{X\beta})'(\boldsymbol{y}^*-\boldsymbol{X\beta})-\frac{1}{2\sigma^2}\boldsymbol{\beta}'\boldsymbol{D_\tau^{-1}\beta}\right]$$

where $\boldsymbol{D_\tau}=diag\;(\tau_1^2,\dots,\tau_k^2)$

$$=exp\left[-\frac{1}{2\sigma}\{(\boldsymbol{\beta}'(\boldsymbol{X}'\boldsymbol{X})\boldsymbol{\beta}\;-2\boldsymbol{y}^*\boldsymbol{X\beta}+\boldsymbol{y}^{*\prime}\boldsymbol{y}^*)+\boldsymbol{\beta}'\boldsymbol{D_\tau^{-1}\beta}\}\right]$$

$$=exp\left[-\frac{1}{2\sigma^2}\{(\boldsymbol{\beta}'(\boldsymbol{X}'\boldsymbol{X}+\boldsymbol{D_\tau^{-1}})\boldsymbol{\beta}\;-2\boldsymbol{y}^*\boldsymbol{X\beta}+\boldsymbol{y}^{*\prime}\boldsymbol{y}^*)\}\right]$$

if we let $\boldsymbol{H} = \boldsymbol{X}'\boldsymbol{X} + \boldsymbol{D}_{\boldsymbol{\tau}}^{-\mathbf{1}}$, then

$$= exp\left[-\frac{1}{2\sigma^2}\{\boldsymbol{\beta}'\boldsymbol{H}\boldsymbol{\beta} - 2\boldsymbol{y}^*\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{y}^{*\prime}\boldsymbol{y}^*\}\right]$$

Now, suppose that

$$(\boldsymbol{\beta} - \boldsymbol{H}^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{y}^*)'\boldsymbol{H}(\boldsymbol{\beta} - \boldsymbol{H}^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{y}^*) = \boldsymbol{\beta}\boldsymbol{H}'\boldsymbol{\beta} - 2\boldsymbol{y}^*\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{y}^{*\prime}(\boldsymbol{X}\boldsymbol{H}^{-\mathbf{1}}\boldsymbol{X}')\boldsymbol{y}^*.$$

Hence,

$$= exp\ [(\boldsymbol{\beta} - \boldsymbol{H}^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{y}^*)'\boldsymbol{H}(\boldsymbol{\beta} - \boldsymbol{H}^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{y}^*) + \boldsymbol{y}^{*\prime}(\boldsymbol{I}_n - \boldsymbol{X}\boldsymbol{H}^{-\mathbf{1}}\boldsymbol{X}')\boldsymbol{y}^*)],$$

Then,

$$\pi(\boldsymbol{\beta}|\,.) \propto exp\left[-\frac{1}{2\sigma^2}\{(\boldsymbol{\beta} - \boldsymbol{H}^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{y}^*)'\boldsymbol{H}(\boldsymbol{\beta} - \boldsymbol{H}^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{y}^*)\}\right],$$

which means that the $\boldsymbol{\beta}$ has normal posterior distribution with mean = $\boldsymbol{H}^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{y}^*$ and variance = $\sigma^2\boldsymbol{H}^{-\mathbf{1}}$.

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

From the full joint density (2.6) the parts that includes σ^2 are as follows,

$$\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}(\boldsymbol{y}^*-\boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y}^*-\boldsymbol{X}\boldsymbol{\beta})}\ \frac{\alpha^h}{\Gamma(h)}(\sigma^2)^{-h-1}\ e^{-\frac{\alpha}{\sigma^2}}$$

$$\prod_{j=1}^{k}\frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}}e^{-\frac{\beta_j^2}{2\tau_2^2\sigma^2}}\ \frac{\lambda^2}{2\sigma^2}e^{-\frac{\lambda^2\tau_j^2}{2\sigma^2}},$$

then, the full conditional posterior distribution is:

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

$$\propto \left(\frac{1}{\sqrt{\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}(\boldsymbol{y}^*-\boldsymbol{X\beta})'(\boldsymbol{y}^*-\boldsymbol{X\beta})} \left(\frac{1}{\sqrt{\sigma^2}}\right)^k e^{-\frac{\boldsymbol{\beta}'\boldsymbol{D}_{\boldsymbol{\tau}}^{-1}\boldsymbol{\beta}}{2\sigma^2}}$$

$$(\sigma^2)^{-h-1}\, e^{-\frac{\alpha}{\sigma^2}}\, \frac{\lambda^2}{2\sigma^2} e^{-\frac{\lambda^2\tau^2}{2\sigma^2}}$$

$$\propto (\sigma^2)^{-\frac{n}{2}-\frac{k}{2}-h-1-1}\, exp\left[-\left\{\frac{1}{2\sigma^2}(\boldsymbol{y}^*-\boldsymbol{X\beta})'(\boldsymbol{y}^*-\boldsymbol{X\beta}) + \frac{\boldsymbol{\beta}'\boldsymbol{D}_{\boldsymbol{\tau}}^{-1}\boldsymbol{\beta}}{2\sigma^2} + \frac{\alpha}{\sigma^2} + \frac{\lambda^2\tau^2}{2\sigma^2}\right\}\right],$$

the last expression can be viewed as inverse gamma distribution with shape parameter $\left(\frac{n}{2}+\frac{k}{2}+h+1\right)$ and scale parameter $(\boldsymbol{y}^*-\boldsymbol{X\beta})'(\boldsymbol{y}^*-\boldsymbol{X\beta}) + \frac{\boldsymbol{\beta}'\boldsymbol{D}_{\boldsymbol{\tau}}^{-1}\boldsymbol{\beta}}{2} + \alpha + \lambda^2\tau^2$.

4. The full conditional posterior distribution of τ_j^2 is:

The parts that included τ_j^2 in the joint distribution (2.6) are as follows:

$$\prod_{j=1}^{k} \frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}} e^{-\frac{\beta_j^2}{2\tau_j^2\sigma^2}}\ \frac{\lambda^2}{2\sigma^2} e^{-\frac{\lambda^2\tau_j^2}{2\sigma^2}} \tag{2.7}$$

Then, the posterior distribution of $\frac{1}{\tau_j^2}$ can be defined as the inverse Gaussian distribution. From (2.7), we have

$$\left(\tau_j^2\right)^{-\frac{1}{2}} exp\left[-\frac{1}{2}\left(\frac{\beta^2}{\sigma^2\tau_j^2} + \frac{\lambda^2\tau_j^2}{\sigma^2}\right)\right], \tag{2.8}$$

Based on (2.7) and (2.8), we can write the posterior distribution of $\frac{1}{\tau_j^2}$ as follows:

$$\propto \prod_{j=1}^{k}\left(\frac{1}{\tau_j^2}\right)^{-\frac{3}{2}} exp\left\{-\frac{\beta_j^2\left({}^{1}\!/_{\tau_j^2}-\sqrt{\frac{\lambda^2}{\beta^2}}\right)}{2\sigma^2\left({}^{1}\!/_{\tau_j^2}\right)}\right\} \tag{2.9}$$

as result, we can name the distribution in (2.9) as the inverse Gaussian with mean (location) $\sqrt{\lambda^2/\beta^2}$.

5. The full conditional posterior distribution of λ^2 is:

Based on the full joint distribution (2.6), we can write the posterior distribution of λ^2 as follows:

$\pi(\lambda^2|\theta, q, \sigma^2) \propto \pi(\lambda^2|\sigma^2, \tau^2)\ \pi(\lambda^2)$

$$\propto \prod_{j=1}^{k}\left(\frac{\lambda^2}{2\sigma^2} e^{-\frac{\lambda^2\tau_j^2}{2\sigma^2}}\right)(\lambda^2)^{q-1} e^{-\theta\lambda^2}$$

$$\propto (\lambda^2)^{(q+k)-1}\, exp\left\{-\lambda^2\left(\theta+\frac{1}{2}\sum_{j=1}^{k}\tau_j^2\right)\right\}. \tag{2.10}$$

From (2.10), we can conclude that λ^2 follows gamma distribution with shape parameter $(q+k+1)$ and rate parameter $\sum_{j=1}^{k}\tau_j^2\,/2$.

2.6 The Proposed Scale Mixture

Based on the following mathematically fact,

$$\int_{w>\frac{|x|}{\sigma^2}} \lambda e^{-\lambda w}\, dw = e^{-\frac{\lambda|x|}{\sigma^2}} \tag{2.11}$$

we can propose the following scale mixture formula. In (2.11), let $x = \beta,\ \lambda w = z$, and by multiply both sides by $\frac{\lambda}{2\sigma^2}$, we get

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

$$\frac{\lambda}{2\sigma^2}e^{-\frac{\lambda|\beta|}{\sigma^2}} = \int_{z > \frac{\lambda|\beta|}{\sigma^2}} \frac{\lambda}{2\sigma^2}e^{-z}\,dz \tag{2.12}$$

so, the formulation (2.12) is the scale mixture of standard exponential mixing with uniform ($\frac{-\sigma^2}{\lambda},\frac{\sigma^2}{\lambda}$).

2.6.1 The Hierarchical Prior Model of Left-Censored Data

Based on the proposed scale mixture (2.12), and (2.1). The hierarchical prior model is formulated as follows:

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

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

$$\boldsymbol{\beta}|\ \sigma^2,\lambda \sim Uniform\left(-\frac{\sigma^2}{\lambda},\frac{\sigma^2}{\lambda}\right), \tag{2.13}$$

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

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

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

2.6.2 The Gibbs Sampling Algorithms

suppose that the full joint density as follows:

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

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

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

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:

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

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

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

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

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

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

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

Therefor,

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

4. The full conditional posterior distribution of z is:

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

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

Therefor,

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

5. The full conditional posterior distribution of λ is:

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

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

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

Therefor,

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

2.7 The Scale Mixture of Uniform Distribution

In 2014, Mallick and Yi proposed the following scale mixture to represent the prior distribution of Laplace,

$$\frac{\lambda}{2\sqrt{\sigma^2}} e^{-\frac{\lambda|\beta|}{\sqrt{\sigma^2}}} = \int_{u > \frac{|\beta|}{\sqrt{\sigma^2}}} \frac{1}{2u\sqrt{\sigma^2}} \frac{\lambda^2}{\Gamma(2)} u^{2-1} e^{-\lambda u} du,$$

This scale mixture of uniform mixing with gamma $(2, \lambda)$ has used in the Bayesian linear regression model. In this thesis we use this scale mixture representation with slight change, by substitute σ^2 instead of $\sqrt{\sigma^2}$, then

$$\frac{\lambda}{2\sigma^2} e^{-\frac{\lambda|\beta|}{\sigma^2}} = \int_{u > \frac{|\beta|}{\sigma^2}} \frac{1}{2u\sigma^2} \frac{\lambda^2}{\Gamma(2)} u^{2-1} e^{-\lambda u} du, \tag{2.14}$$

we still ensure that scale mixture is unimodal by conditioning on σ^2.

2.7.1 The Hierarchical Prior Model of the Left-Censored Data

Based on the scale mixture (2.14) and the left-censored structure model. The hierarchical prior model defined as follows:

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

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

$$\boldsymbol{\beta} | \sigma^2, \boldsymbol{u} \sim \prod_{j=1}^{k} Uniform\left(-\sigma^2 u_j, \sigma^2 u_j\right),$$

$$\boldsymbol{u} | \lambda \sim \prod_{j=1}^{k} Gamma\,(2\,, \lambda),$$

$$\sigma^2 \sim Inverse - Gamma(a, r),$$

$$\lambda \sim Gamma(c, d). \tag{2.15}$$

2.7.2 The Gibbs sampling Algorithm

We suppose that the full joint density is:

$$f(\boldsymbol{y}|\boldsymbol{\beta}, \sigma^2)\, \pi(\sigma^2) \prod_{j=1}^{k} \pi\left(\beta_j | u_j, \sigma^2\right) \pi\left(u_j | \lambda\right)$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y}^* - \boldsymbol{X\beta})'(\boldsymbol{y}^* - \boldsymbol{X\beta})\right\} \frac{r^a}{\Gamma(a)} (\sigma^2)^{-a-1} e^{-r/_{\sigma^2}}$$

$$\prod_{j=1}^{k} \frac{1}{2\sigma^2 u_j} \frac{\lambda^2}{\Gamma(2)} {u_j}^{2-1} e^{-\lambda u_j}$$

No, we will write down the full conditional posterior distributions.

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

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

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

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

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

$$\propto exp\left\{-\frac{1}{2\sigma^2}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}_{OLS}\right)' \boldsymbol{X}'\boldsymbol{X}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}_{OLS}\right)\right\}$$

$$\prod_{j=1}^{k} I\left\{|\beta_j| < \sigma^2 u_j\right\}$$

Therefore,

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

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

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

$$\propto \left(\frac{1}{\sqrt{\sigma^2}}\right)^n exp\left\{-\frac{1}{2\sigma^2}(y^* - \boldsymbol{X\beta})'(y^* - \boldsymbol{X\beta})\right\}\frac{r^a}{\Gamma(a)}(\sigma^2)^{-a-1}$$

$$e^{-r/\sigma^2}\left(\frac{1}{2u_j\sigma^2}\right)^k \prod_{j=1}^{k}\left\{\sigma^2 > Max_j\left(\frac{|\beta_j|}{u_j}\right)\right\}$$

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

$$\prod_{j=1}^{k}\left\{\sigma^2 > Max_j\left(\frac{|\beta_j|}{u_j}\right)\right\}$$

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

$$\prod_{j=1}^{k}\left\{\sigma^2 > Max_j\left(\frac{|\beta_j|}{u_j}\right)\right\}$$

Therefor,

$$\sigma^2|\boldsymbol{y},\boldsymbol{X},\boldsymbol{\beta},u \sim Inverse - Gamma\left(\frac{n}{2} + a + k, (\boldsymbol{y}^* - \boldsymbol{X\beta})'(\boldsymbol{y}^* - \boldsymbol{X\beta}) + \frac{r}{2}\right)$$

$$\prod_{j=1}^{k}\left\{\sigma^2 > Max_j\left(\frac{|\beta_j|}{u_j}\right)\right\}$$

4. The full conditional posterior distribution of u_j is:

a. $u_j^* \sim Exp\ (\lambda)$

b. $u_j = u_j^* + \frac{|\beta_j|}{\sqrt{\sigma^2}}$

5. The full conditional posterior distribution of λ is:

$$\pi(\lambda|u) \propto \pi(u|\lambda)\,\pi(\lambda)$$

$$\propto \left[\prod_{j=1}^{k} \frac{\lambda^2}{\Gamma(2)} u_j^{2-1} e^{-\lambda u_j}\right] \frac{\lambda^{c-1}}{\Gamma(c)} d^c e^{-d\lambda}$$

$$\propto (\lambda^2)^k\, e^{-\lambda \sum_{j=1}^{k} u_j}\, \lambda^{c-1}\, e^{-d\lambda},$$

$$\propto (\lambda)^{2k+c-1}\, exp\left\{-\lambda\left(\sum_{j=1}^{k} u_j + d\right)\right\}.$$

This is gamma distribution with shape parameter $(2k + c)$ and rate parameter $(\sum_{j=1}^{k} u_j)$.

2.8 Extensions of the Right-Censored Data

After introducing the above three scale mixture representation for the Laplace distribution as the prior distribution of the interested parameter of left-censored regression, we will rework the above scale mixture but for the right-censored regression model, the structural formulation of the right-censored linear regression is defined by:

$$y_i = \begin{cases} y_i^* & if \;\; y_i^* < y_U, \\ y_U & if \;\; y_i^* \geq y_U, \end{cases} \tag{2.16}$$

or equivalently,

$$y_i = min(y_i^*, y_U), \qquad \text{where } \; y_i^* = \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_i,$$

y_i^* is the latent variable or the unobservable variable. Furthermore, the following write down represents the Gibbs sample algorithms:

2.8.1 Gibbs sampling Algorithm for Scale Mixture of Normal

Sampling parameters of the right-censored regression model (2.16), unobserved variables, the hierarchical prior model (2.5) and full posterior distributions of the full joint model (2.6) guide us to the exact Gibbs sampler with the following steps:

1. Sampling $\boldsymbol{y}^*$: we draw samples from

$$y_i^*|y_i\ ,\boldsymbol{\beta} \sim \begin{cases} y_i & if\ \ y_i^* < y_U, \\ N(\boldsymbol{X\beta}, \sigma^2 \boldsymbol{I}_n) & if\ \ y_i^* \geq y_U, \end{cases}$$

2. Sampling $\boldsymbol{\beta}$: we draw samples for $\boldsymbol{\beta}$ from normal distribution with mean = $(\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{D}_{\boldsymbol{\tau}}^{-\mathbf{1}})^{-1}\boldsymbol{X}'\boldsymbol{y}^*$ and variance = $\sigma^2(\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{D}_{\boldsymbol{\tau}}^{-\mathbf{1}})^{-1}$; $\boldsymbol{D}_{\boldsymbol{\tau}} =$ $diag\ (\tau_1^2, \dots, \tau_k^2)$.

3. Sampling σ^2: we draw samples from inverse gamma distribution with shape parameter $(\frac{n}{2} + \frac{k}{2} + h + 1)$ and scale parameter $(\boldsymbol{y}^* - \boldsymbol{X\beta})'(\boldsymbol{y}^* - \boldsymbol{X\beta}) + \frac{\boldsymbol{\beta}'\boldsymbol{D}_{\tau}^{-\mathbf{1}}\boldsymbol{\beta}}{2} + \alpha + \lambda^2\tau^2$.

4. Sampling τ_j^{-2}: we draw samples from inverse Gaussian with mean (location) $\sqrt{\lambda^2/\beta^2}$ and shape parameter λ^2.

5. Sampling λ^2: we draw samples from gamma distribution with shape parameter $(q + k + 1)$ and rate parameter $\sum_{j=1}^{k} \tau_j^2\ /2$.

2.8.2 Gibbs Sampling Algorithm for Proposed Scale Mixture

Sampling parameters of the right-censored regression model (2.16), unobserved Variables, the hierarchical prior model (2.13) guide us to the exact Gibbs sampler with the following steps:

1. Sampling $\boldsymbol{y}^*$: we draw samples from

$$y_i^* | y_i , \boldsymbol{\beta} \sim \begin{cases} y_i & if \;\; y_i^* < y_U, \\ N(\boldsymbol{X\beta}, \sigma^2 \boldsymbol{I}_n) & if \;\; y_i^* \geq y_U, \end{cases}$$

2. Sampling $\boldsymbol{\beta}$: we draw samples from

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

3. Sampling σ^2: we draw samples from

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

4. Sampling z: we draw samples from

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

5. Sampling λ: we draw samples from

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

2.8.3 Gibbs Sampling Algorithm for Scale Mixture of Uniform

Sampling parameters of the right-censored regression model (2.16), unobserved variables, the hierarchical prior model (2.15) and full posterior distributions of the full joint model guide us to the exact Gibbs sampler with the following steps:

1. Sampling $\boldsymbol{y}^*$: we draw samples from

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

2. Sampling $\boldsymbol{\beta}$: we draw samples from normal distribution

$$N\left(\widehat{\boldsymbol{\beta}}_{OLS}, \sigma^2 (\boldsymbol{X'X})^{-1}\right) \prod_{j=1}^{k} I\left\{|\beta_j| < \sigma^2 u_j\right\}$$

3. Sampling σ^2: we draw samples from inverse gamma distribution

$$Inverse - Gamma\left(\frac{n}{2} + q + k, (\boldsymbol{y}^* - \boldsymbol{X\beta})'(\boldsymbol{y}^* - \boldsymbol{X\beta}) + \frac{r}{2}\right)$$

$$\prod_{j=1}^{k} \left\{\sigma^2 > Max_j\left(\frac{|\beta_j|}{u_j}\right)\right\}$$

4. Sampling u_j: we draw samples from this is gamma distribution with rate parameter (λ)
5. Sampling λ: we draw samples from this is gamma distribution with shape parameter $(2k + c)$ and rate parameter $(\sum_{j=1}^{k} u_j)$.

Chapter Three

Practical Side

3.1 Simulation Study

For the purpose of applying what was mentioned on the theoretical side, the simulation method was used, which simulated a very large number of hypothetical cases that can appear in practice so that the results are more comprehensive and general.

3.1.1 Simulation Study for Left-Censored Data

In this section, we demonstrate the prediction accuracy of the methods: linear left-censored regression (Tobit), Bayesian LASSO left-censored regression (BLLCR), the proposed Bayesian LASSO left-censored regression (NBLLCR), 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 the prediction accuracy, we evaluate the median of mean squared errors (MMSE) for the simulation study based on 100 replications. We ran our algorithm for 12000 iterations discarding the first 2000 as burn in.

Example 1 (Left-censored with sparse case)

In this example, we generate data from the correct model

$$y_i = max(0, y_i^*), \qquad i = 1, \dots, n,$$

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

We set $\beta^{10\times1} = (6,1,0,0,3,0,0,0,0,0)'$ 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 proposed 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 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
	MMSE SD	MMSE SD	MMSE SD	MMSE SD
(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.

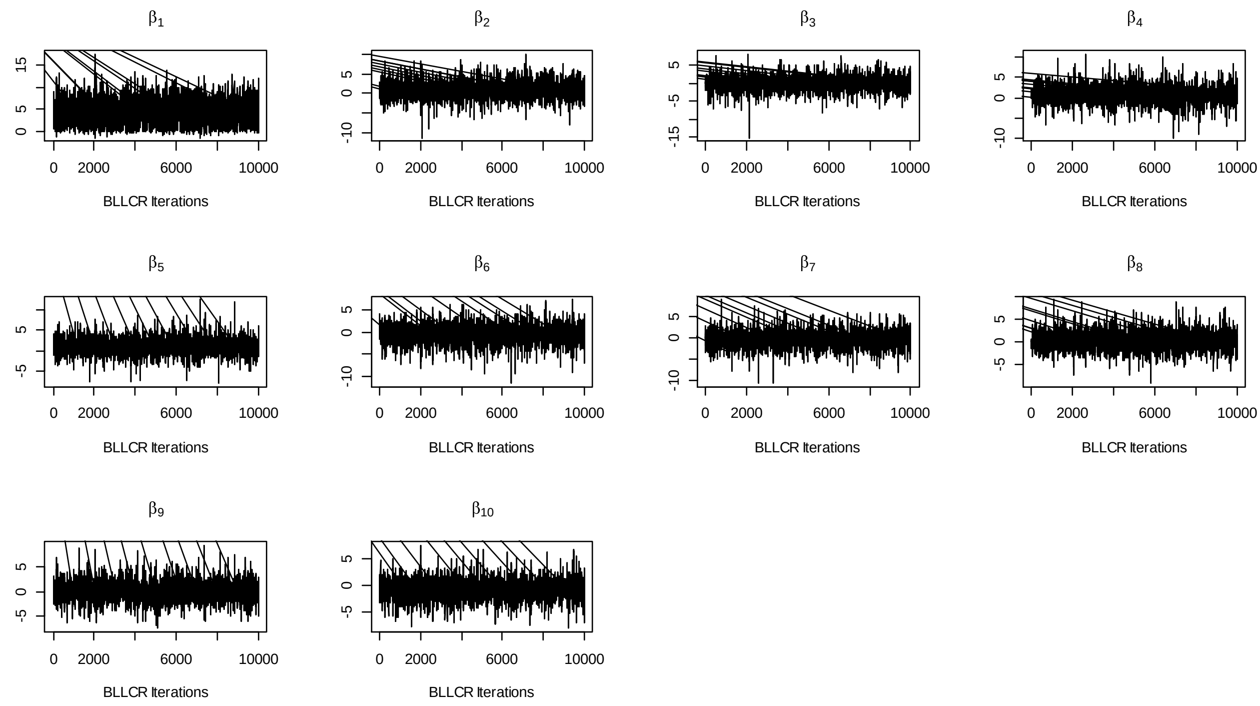

Figure 1: Trace plots of parameters in simulation 1 using BLLCR method.

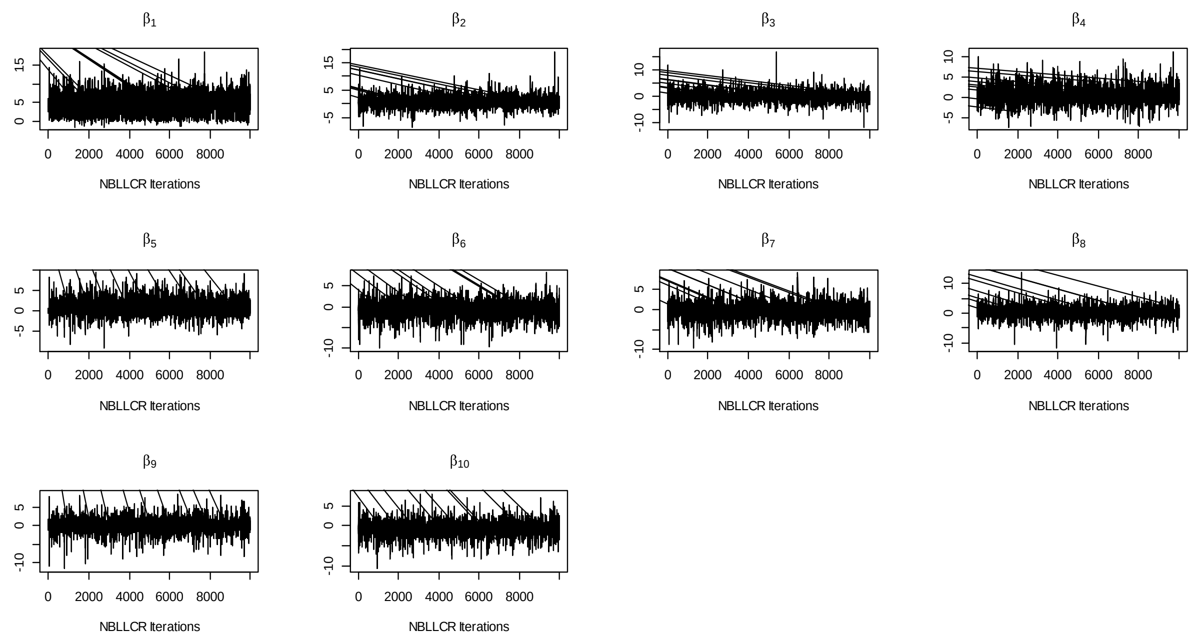

Figure 2: Trace plots of parameters in simulation 1 using NBLLCR method.

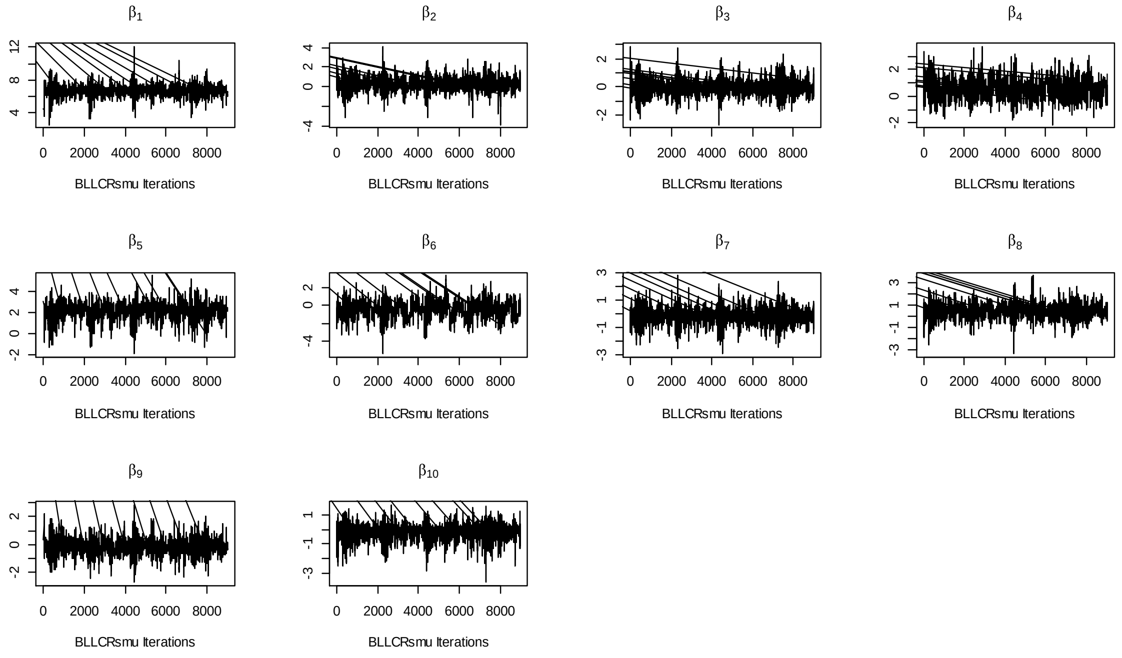

Figure 3: Trace plots of parameters in simulation 1 using BLLCRsmu method.

Example 2 (Left-censored with dense case)

Here we set $\beta^{10\times1} = (6,1,1,1,1,1,1,1,1,1)'$, leaving other setups exactly the same as in Example (1). The results are listed 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.

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
	MMSE SD	MMSE SD	MMSE SD	MMSE SD
(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)

Example 3 (Left-censored with very sparse case)

Here we set $\beta^{10\times1} = (6,0,0,0,0,0,0,0,0,0)'$, leaving other setups exactly the same as in Example (1). The results are listed in Table (3). The results show that the proposed 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.

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
	MMSE SD	MMSE SD	MMSE SD	MMSE SD
(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)

3.1.2 Simulation Study for Right-Censored Data

In this section, we demonstrate the prediction accuracy of the methods: linear right-censored regression (RCR), Bayesian LASSO right-censored regression (BLRCR), the proposed Bayesian LASSO right-censored regression (NBLRCR), Bayesian LASSO right-censored regression using scale mixture uniform (BLRCRsmu). Similar to Section (3.1.1), the outcome variable is centered and the covariates are standardized to have 0 means and unit variances before applying the above methods. For the prediction accuracy, we evaluate the median of mean squared errors (MMSE) for the simulated studies based on 100 replications.

Example 4 (Right-censored with sparse case)

This example is similar to Example (1) except that we set $\beta^{10\times1} = (6,1,0,0,3,0,0,0,0,0)'$, $\sigma^2 = \{1,1.5,2\}$ and we generate data from the correct model

$$y_i = min(5, y_i^*), \qquad i = 1, \dots, n,$$

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

The results are listed in Table (4). The results show that the proposed Bayesian LASSO right-censored regression (NBLRCR) performs very well compared to other methods in the comparison. It has the smallest MMSE in 7 out of 9 experimental results. The Bayesian LASSO right-censored regression (BLRCR) also performs well compared to other methods in the comparison. It has the smallest MMSE in 2 out of 9 experimental results.

Table 4: Median mean squared error (MMSE) and their associated standard deviations (SD) are listed in the parentheses for Example (4). All results are averaged over 100 replications.

(n_t, n_p, σ^2)	RCR	BLRCR	NBLRCR	BLRCRsmu
	MMSE SD	MMSE SD	MMSE SD	MMSE SD
(100,200,1)	0.1153 (0.0336)	0.1409 (0.1124)	**0.0878** (0.0410)	0.1047 (0.0333)
(150,200,1)	0.2331 (0.1111)	0.2961 (0.1212)	**0.1320** (0.0315)	0.1981 (0.0838)
(200,200,1)	0.4286 (0.2022)	0.5258 (0.3168)	**0.2784** (0.1684)	0.3683 (0.1904)
(100,200,1.5)	0.0651 (0.0134)	0.0617 (0.0329)	**0.0377** (0.0078)	0.0565 (0.0120)
(150,200,1.5)	0.1238 (0.1136)	0.1432 (0.0435)	**0.0737** (0.0415)	0.1078 (0.0908)
(200,200,1.5)	0.3585 (0.2697)	0.3597 (0.0740)	**0.2337** (0.1508)	0.3010 (0.2128)
(100,200,2)	0.0560 (0.0174)	0.0724 (0.0301)	**0.0520** (0.0157)	0.0537 (0.0174)
(150,200,2)	0.1886 (0.0796)	**0.1197** (0.0328)	0.1266 (0.0617)	0.1691 (0.0749)
(200,200,2)	0.1889 (0.0801)	**0.1201** (0.0331)	0.1288 (0.0612)	0.1696 (0.0751)

Convergence of the corresponding our Gibbs sampler methods was assessed by trace plots of the simulated draws. The trace plots Figures (4 – 6) shows that our methods converge very fast.

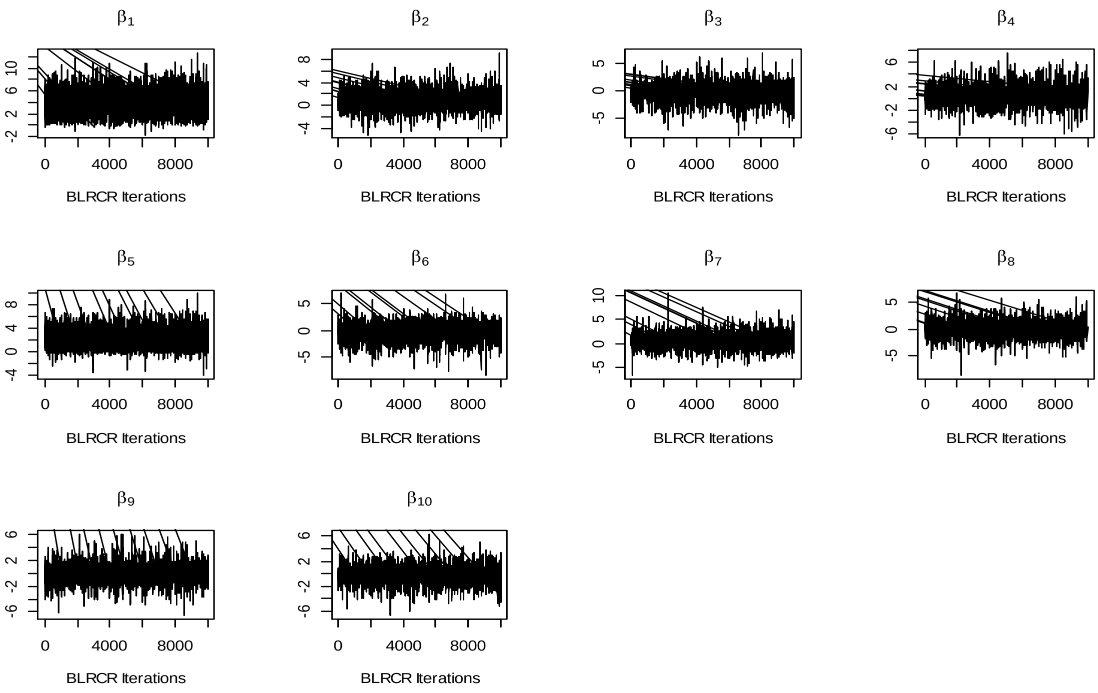

Figure 4: Trace plots of parameters in simulation 4 using BLRCR method.

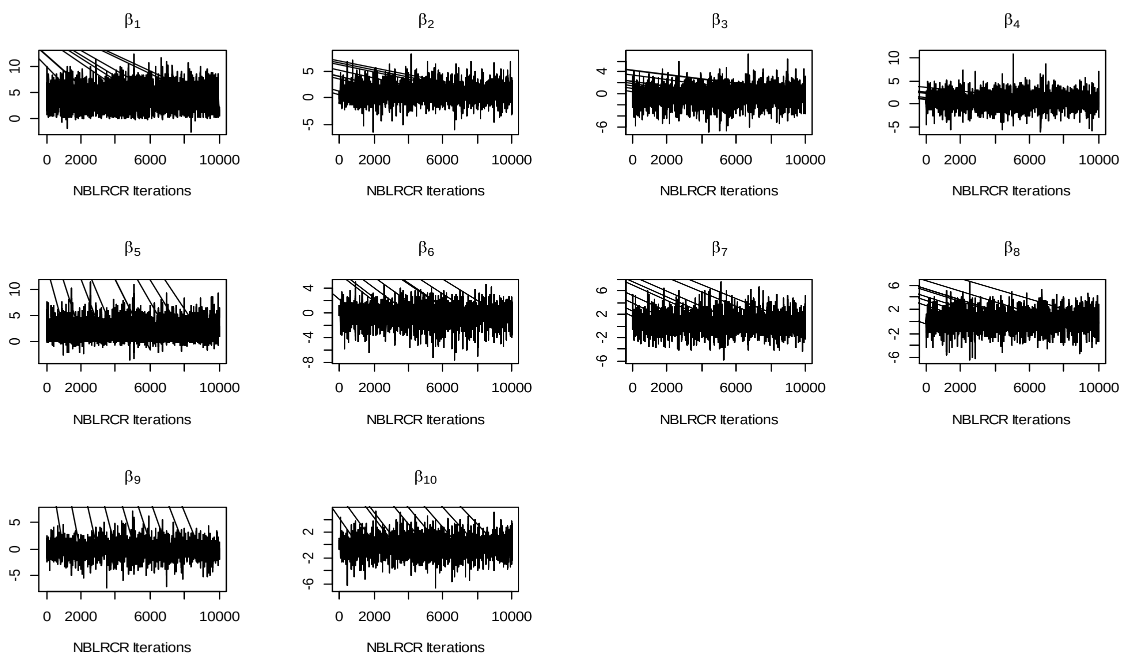

Figure 5: Trace plots of parameters in simulation 4 using NBLRCR method.

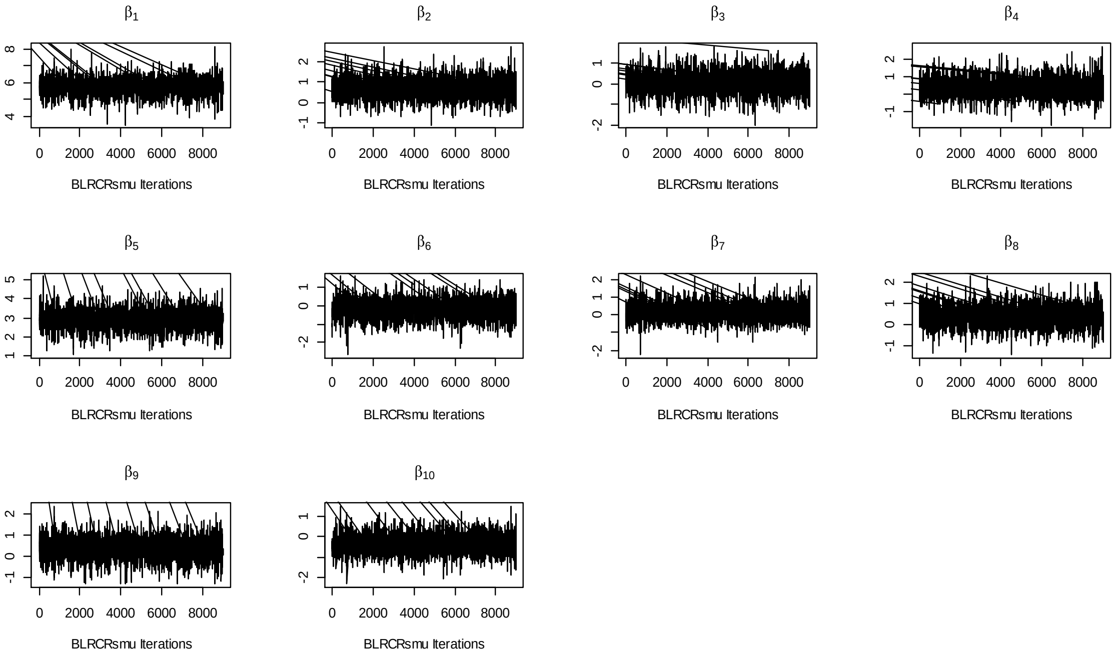

Figure 6: Trace plots of parameters in simulation 4 using BLRCRsmu method.

Example 5 (Right-censored with dense case)

This example is similar to Example (2) except that we set $\beta^{10\times 1} = (6,1,1,1,1,1,1,1,1,1)'$, $\sigma^{2} = \{1,1.5,2\}$ and we generate data from the correct model

$$y_i = min(5, y_i^*), \qquad i = 1, \dots, n,$$

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

The results are listed in Table (5). The results show that the proposed Bayesian LASSO right-censored regression (NBLRCR) performs very well compared to other methods in the comparison. It has the smallest MMSE in 7 out of 9 experimental results. The Bayesian LASSO right-censored regression using scale mixture uniform (BLRCRsmu) also performs well compared to other methods in the comparison. It has the smallest MMSE in 2 out of 9 experimental results.

Table 5: Median mean squared error (MMSE) and their associated standard deviations (SD) are listed in the parentheses for Example (5). All results are averaged over 100 replications.

(n_t, n_p, σ^2)	RCR	BLRCR	NBLRCR	BLRCRsmu
	MMSE SD	MMSE SD	MMSE SD	MMSE SD
(100,200,1)	0.2087 (0.0372)	0.5022 (0.1124)	**0.1485** (0.0490)	0.1506 (0.0376)
(100,200,1)	0.3448 (0.1619)	0.7973 (0.1993)	**0.3309** (0.1592)	0.3348 (0.1560)
(100,200,1)	0.6431 (0.2490)	1.4096 (0.6720)	**0.6179** (0.3090)	0.6233 (0.2321)
(150,200,1.5)	0.0832 (0.0436)	0.1600 (0.0620)	0.0890 (0.0120)	**0.0813** (0.0400)
(150,200,1.5)	0.2122 (0.0480)	0.3839 (0.0727)	**0.1677** (0.0262)	0.1708 (0.0483)
(150,200,1.5)	0.3085 (0.0758)	0.7735 (0.2405)	**0.2688** (0.0827)	0.2758 (0.0886)
(200,200,2)	0.0511 (0.0172)	0.0768 (0.0156)	0.0533 (0.0070)	**0.0496** (0.0157)
(200,200,2)	0.1106 (0.0420)	0.2071 (0.0742)	**0.1009** (0.0382)	0.1019 (0.0418)
(200,200,2)	0.2659 (0.1149)	0.4753 (0.0737)	**0.2492** (0.0718)	0.2539 (0.1031)

Example 6 (Right-censored with very sparse case)

This example is similar to Example (1) except that we generate data from the correct model

$$y_i = min(5, y_i^*), \qquad i = 1, \dots, n,$$

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

The results are listed in Table (6). The results show that the Bayesian LASSO right-censored regression (BLRCR) performs very well compared to other methods in the comparison. It has the smallest MMSE in 6 out of 9 experimental results. The proposed Bayesian LASSO right-censored regression (NBLRCR) also performs well compared to other methods in the comparison. It has the smallest MMSE in 3 out of 9 experimental results.

Table 6: Median mean squared error (MMSE) and their associated standard deviations (SD) are listed in the parentheses for Example (6). All results are averaged over 100 replications.

(n_t, n_p, σ^2)	RCR	BLRCR	NBLRCR	BLRCRsmu
	MMSE SD	MMSE SD	MMSE SD	MMSE SD
(100,200,1)	0.1075 (0.0516)	**0.0444** (0.0414)	0.0608 (0.0277)	0.0821 (0.0427)
(100,200,1)	0.2998 (0.1106)	**0.1414** (0.0549)	0.1509 (0.0547)	0.2212 (0.0794)
(100,200,1)	0.2587 (0.1850)	1.1212 (0.0852)	**0.1300** (0.0687)	0.1690 (0.1198)
(150,200,1.5)	0.1240 (0.0587)	0.0777 (0.0471)	**0.0743** (0.0414)	0.1069 (0.0549)
(150,200,1.5)	0.2465 (0.1547)	**0.0971** (0.0546)	0.1496 (0.0899)	0.2000 (0.1225)
(150,200,1.5)	0.2389 (0.1096)	**0.0954** (0.0717)	0.1461 (0.0670)	0.1685 (0.0858)
(200,200,2)	0.0587 (0.0234)	0.0406 (0.0201)	**0.0383** (0.0191)	0.0510 (0.0214)
(200,200,2)	0.1211 (0.0352)	**0.0438** (0.0104)	0.0836 (0.0333)	0.0957 (0.0317)
(200,200,2)	0.2687 (0.1187)	**0.0903** (0.0476)	0.1815 (0.0853)	0.2004 (0.0929)

3.2 Real Data

Based on what was described in the simulation using the R program and according to the requirements of the study, the data is analyzed in two ways:

3.2.1 Real Data (Left-Censored Data)

We demonstrate the performance of the methods with the 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 on 9 variables. We use a randomly subsample of this data set 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 the gender, age, years, children, religiousness, education, occupation, and rating.

The results are listed in Table (7) The results show that the proposed Bayesian LASSO left-censored regression (NBLLCR) performs very well compared to other methods in the comparison.

Table 7: Mean squared error (MSE) for the affairs data.

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

3.2.2 Real Data (Right-Censored Data)

Data for 62 patients were obtained. These data were collected from Al-Hashimiya General Hospital. The research variables consist of a dependent variable and 10 independent variables, which are;

Urea level in blood (y_i): (Uremia) is caused by extreme and usually irreversible damage to your kidneys. This is usually from chronic kidney disease. The kidneys are no longer able to filter the waste from your body and send it out through your urine. Instead, that waste gets into your bloodstream, causing a potentially life-threatening condition. The normal percentage of urea in blood around 6 to 24 mg/dL (2.1 to 8.5 mmol/L) is considered.

Age (x_1): In humans an age-related increase in plasma urea levels and no correlation between plasma creatinine and age. Fractional urea excretion decreases with age.

Urinary tract obstruction (x_2): A blockage (obstruction) where the ureter connects to the kidney or bladder. This prevents urine flow. A blockage of the ureter and kidney meet (ureteropelvic junction) may cause the kidney to swell and eventually stop working. This condition can be congenital or can develop with typical childhood growth, result from an injury or scarring, or in rare cases, develop from a tumor. A blockage where the ureter and bladder meet (ureterovesical junction) may cause urine to back up into the kidneys.

Congestive heart failure (x_3): With reduced ejection fraction is another risk factor for kidney disease. When the heart is unable to pump forcefully, the amount of blood it ejects with each contraction drops. This reduces the amount of blood that passes through the kidneys, causing urine and waste output to drop. Because salt isn't eliminated well, fluid may build up, causing heart failure to worsen.

Having a heart attack (x_4): The stress of a heart attack can result in hormonal changes within the body, and that can have a negative effect on how well the kidneys work. Changes in heart function may lead to kidney damage by decreasing the blood supply to the kidneys result increase in blood urea.

Gastrointestinal bleeding (x_5): When upper GI bleeding occurs, the blood is digested to protein. This protein is transported to the liver via the portal vein, and metabolized to BUN in the urea cycle. Higher BUN values are therefore associated with the digestion of blood.

Drought (x_6): Drought and lack of fluids in the body or increased protein, whether from food or muscle loss, which is one of the most important causes of deficiency in kidney functions.

Severe burns (x_7): Acute renal failure occurred immediately after burns is mostly due to reduced cardiac output, which is mainly caused by fluid loss. This is usually caused by delayed or inadequate fluid resuscitation but may also result from substantial muscle breakdown or hemolysis.

Pharmaceuticals (x_8): A biomarker that is released directly into the blood or urine by the kidney in response to injury may be a better early marker of drug-induced kidney toxicity than BUN and serum creatinine.

Sugar percentage (x_9): Increment of blood urea level with the increment of blood sugar level clearly indicates that the increase blood sugar level causes damage to the kidney. Research conducted by Anjaneyulu et al. (2004) had found that increase urea and serum creatinine in diabetic rats indicates progressive renal damage.

Blood fat levels (x_{10}): Being overweight can directly affect kidneys. Extra weight forces the kidneys to work harder and filter wastes above the normal level. Over time, this extra work increases the risk for kidney disease.

Table 8: showed the estimation of parameters.

Descriptive variables	Variables	NBLRCR	BLRCRsmu	BLRCR
Age	x_1	0.000	0.003	0.000
Urinary tract obstruction	x_2	0.654	0.793	0.543
Congestive heart failure	x_3	0.000	0.000	0.065
Having a heart attack	x_4	0.763	0.439	0.652
Gastrointestinal bleeding	x_5	0.000	0.000	0.005
Drought	x_6	0.732	0.874	0.609
Severe burns	x_7	0.000	0.000	0.070
Pharmaceuticals	x_8	2.210	3.095	0.054
Sugar percentage	x_9	0.854	0.986	0.549
Blood fat levels	x_{10}	0.000	0.000	0.005
MSE		**13.98**	16.43	17.86

We can see that the proposed model gave least the value for MSE, is 13.98 , and the from above table the estimation of parameters were taken from the subsequent distributions of the proposed model, by adding a threshold point to zeroing because Bayesian methods do not zero, and the proposed method has reduced many unimportant variables such as making a variable selection in the proposed model in the five variables (Age, Congestive heart failure, Gastrointestinal bleeding, Severe burns, and Blood fat levels) where the parameters were ($x_1 = 0,\ x_3 = 0,\ x_5 = 0,\ x_7 = 0, and\ x_{10} = 0$).

Chapter Four

Conclusions

and

Recommendations

4.1 Conclusions and Future Research

In this thesis, we have proposed several Bayesian methods for variable selection and parameter estimation in linear regression models with left and right-censored data. Some advantages over old approaches include fast convergence Gibbs sampler, efficient Gibbs sampler computation techniques, and the use of data augmentation to allow left and right-censored responses. Some contributions and future research topics are summarized below.

4.1.1 Main Contributions

Bayesian regularization methods for variable selection and parameter estimation for left and right-censored responses are proposed in this thesis.

These proposed methods are summarized as follows:

1. We have proposed a Bayesian regularization method for left and right-censored responses based on the Bayesian regularized method of Park & Casella (2008).
2. We have proposed a Bayesian regularization method for left and right-censored responses based on the Bayesian regularized method of Mallick & Yi (2014).
3. We have proposed a Bayesian regularization method for left and right-censored responses based on a proposed scale mixture formula for Laplace prior.
4. The results show that the proposed method performs very well compared with the classical methods for left and right-censored data.
5. In practice, the results show that the proposed methods perform very well in terms of convergence.

4.2 Recommendations

We recommend the usage of the proposed scale mixture of uniforms mixing the standard exponential distribution when:

1. Dealing with the presents of the multicollinearity problem under left and right-censored regression models.
2. Dealing with Bayesian group LASSO under left and right-censored regression models.
3. Dealing with Bayesian LASSO under an interval-censored regression model.
4. Dealing with Bayesian adaptive LASSO under left and right-censored regression models.

Bibliography

Bibliography

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, *19*(6), 716-723.

Alan, S., Honoré, B. E., Hu, L., & Leth-Petersen, S. (2014). Estimation of panel data regression models with two-sided censoring or truncation. *Journal of Econometric Methods*, *3*(1), 1-20.

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

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

Alhusseini, F. H. H., Flaih, A. N., & Alshaybawee, T. (2020). Bayesian extensions on LASSO and adaptive LASSO Tobit regressions. Periodicals of Engineering and Natural Sciences, 8(2), 1131-1140.

Amemiya, T. (1973). Regression analysis when the dependent variable is truncated normal. *Econometrica: Journal of the Econometric Society*, 997-1016.

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.

Andrews, D. F., & Mallows, C. L. (1974). Scale mixtures of normal distributions. Journal of the Royal Statistical Society: Series B (Methodological), 36(1), 99-102.

Balestra, P. (1970). On the efficiency of ordinary least-squares in regression models. Journal of the American Statistical Association, 65(331), 1330-1337.

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

Chand, S. (2012, January). On tuning parameter selection of LASSO-type methods-a monte carlo study. In Proceedings of 2012 9th international Bhurban conference on applied sciences & technology (IBCAST) (pp. 120-129). IEEE.

Chay, K. Y., & Powell, J. L. (2001). Semiparametric censored regression models. *Journal of Economic Perspectives, 15*(4), 29-42.

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.

Efroymson, M. A. (1960). Multiple regression analysis. *Mathematical methods for digital computers*, 191-203.

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.

Fernando, R. (2011). Logit, probit and tobit: Models for categorical and limited dependent variables. PLCS/RDC Statistics and Data Series at the West.

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.

George, E. I., & McCulloch, R. E. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association*, *88*(423), 881-889.

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

Henningsen, A. (2010). Estimating censored regression models in R using the censReg Package. *R package vignettes*, *5*, 12.

Hocking, R. R. (1976). A Biometrics invited paper. The analysis and selection of variables in linear regression. Biometrics, 1-49.

Kohler, M., Máthé, K., & Pintér, M. (2002). Prediction from randomly right-censored data. Journal of Multivariate Analysis, 80(1), 73-100.

Koul, H., Susarla, V., & Van Ryzin, J. (1981). Regression analysis with randomly right-censored data. The Annals of statistics, 1276-1288.

Maddala, G. S. (1987). Limited dependent variable models using panel data. Journal of Human resources, 307-338.

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

Mallows, C. L. (1973). Some comments on C_p, Technometries, 15, 661-675.

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.

Savin, I. (2013). A comparative study of the LASSO-type and heuristic model selection methods. Jahrbücher für Nationalökonomie und Statistik, 233(4), 526-549.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 461-464.

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.

المستخلص

تعتبر طرائق التنظيم وخاصاً طريقة LASSO من أوسع الطرائق انتشاراً في موضوع اختيار المتغيرات في تحليل الانحدار. هذه الرسالة ركزت على دراسة طريقة التنظيم المسماة بطريقة LASSO من خلال نظرية Bayes، حيث تم توظيف ثلاث نماذج لتمثيل التوزيع المسبق (توزيع لابلاس) لمعلمة الانحدار، النموذج الأول افترض استخدام تمثيل التوزيع الطبيعي مع التوزيع الاسي، والتمثيل الثاني كان من خلال اقتراح تمثيل للتوزيع المنتظم مع التوزيع الاسي القياسي، والتمثيل الثالث كان من خلال التوزيع المنتظم مع توزيع كاما. حيث تم تطبيق هذه النماذج الثلاثة فيما يسمى بانحدار البيانات المراقبة من اليسار وانحدار البيانات المراقبة من اليمين.

وقد تم تطبيق النماذج الثلاثة وفق نهج Bayes من خلال تنفيذ خوارزمية Gibbs sampler عن طريق أسلوب المحاكاة من خلال افتراض ثلاث تجارب للمحاكاة وباستخدام لغة البرمجة R، حيث قمنا بأجراء تجارب المحاكاة بأحجام عينات مختلفة وقيم تباينات مختلفة للأخطاء وباستخدام معيار الوسيط لمتوسط مربعات الخطأ ثم الحكم على أداء الطرائق المختلفة. ولبيان كفاءة الطريقة المقترحة تم توظيف هذه الطريقة على بيانات حقيقية تمثل عينة مأخوذة من بحث سابق تتصف هذه العينة بأن متغير الاستجابة هو متغير مراقب من اليسار، هذه البيانات متوفرة في لغة R. والتطبيق العملي الثاني هو لبيانات تتصف متغير الاستجابة بها بأنه متغير مراقب من اليمين حيث كانت العينة تمثل مستوى اليوريا في الدم مع مجموعة من المتغيرات التفسيرية. وأظهرت النتائج في كل من المثالين التطبيقين ان الطريقة المقترحة ذات قدرة تنبؤية عالية للنموذج مقارنة مع الطرائق الأخرى من مبدأ دقة التنبؤ واختيار المتغيرات.

جمهورية العراق

وزارة التعليم العالي والبحث العلمي

جامعة القادسية / كلية الإدارة والاقتصاد

قسم الإحصاء

انحدار لاسو البيزي للبيانات المراقبة من اليسار واليمين مع تطبيق عملي

رسالة مقدمة الى مجلس كلية الإدارة والاقتصاد في جامعة القادسية

وهي كجزء من متطلبات نيل درجة الماجستير في علوم الإحصاء

تقدمت بها

زينب سعدون عليوي

بإشراف

أ . م . د. بحر كاظم محمد

2023م 1444هـ