Symbolic Data Analysis

Statistical Inference on Interval-valued Data Regression

by

Yaotong Cai

(Under the Direction of Lynne Billard)

Abstract

Interval-valued data are one of the most common forms of symbolic data. Previous studies have provided a number of approaches to conduct linear regression models for interval data, while few have involved issues surrounding inference on the regression coefficient estimates. In this dissertation, we propose a method of statistical inference on coefficient estimates for interval data regression by means of the maximum likelihood principle. Under some assumptions, this method not only enables us to provide point estimators of the parameters in linear regression models, but also gives the distributions of the point estimators, as well as the confidence intervals. Performances of the proposed method are evaluated by simulations as well as real data analyses.

Index words: Symbolic data analysis, Interval-valued data, Linear regression, Statistical inference, Maximum likelihood

Symbolic Data Analysis

Statistical Inference on Interval-valued Data Regression

by

Yaotong Cai

B.S., RENMIN University of China, 2013

A Dissertation Submitted to the Graduate Faculty

of The University of Georgia in Statistics

of the

Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2018

©2018

Yaotong Cai

All Rights Reserved

Symbolic Data Analysis

Statistical Inference on Interval-valued Data Regression

by

Yaotong Cai

Approved:

Major Professor: Lynne Billard

Committee: T.N. Sriram
William P. McCormick
Ping Ma
Wenxuan Zhong

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2018

Symbolic Data Analysis:
Statistical Inference on Interval-valued Data
Regression

Yaotong Cai

May 24th, 2018

Acknowledgments

I would like to express my deepest appreciation and gratitude to Professor Lynne Billard, my major advisor, for her delicated guidance, tremendous help and cares through my research. Her inspiration, enthusiasm, and intelligence has set a good example for me as a successful scientific researcher. Without her patient guidance and help, I would never have completed this dissertation.

I would also like to extend my sincere thanks to Dr. T.N. Sriram, Dr. William McCormick, Dr. Ping Ma, and Dr. Wenxuan Zhong to serve as my advisory committee. I appreciate their invaluable aids, thoughtful comments, and precious time they have spent to review my dissertation. I would also like to thank Dr. Dan Hall for his help and guidance during the SAS shootout. I appreciate all the faculty, staff, and my friends in the department for their tremendous instruction, help, and guidance throughout my years study in the program.

Lastly, special thanks to my parents, who give me endless love and consistent support and encouragement throughout my life and study for years. Their love is always the greatest motivation for me to complete the study.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

Symbolic data is a new type of data. Considering data of p random variables, unlike classical data which are represented by a single point in p-dimensional space $\mathcal{R}^p$, values of symbolic data are of p-dimensional hypercube forms in $\mathcal{R}^p$, or Cartesian products of p distributions. Because of the different structure, one of the most distinctive features of symbolic data is the internal variations they have within the observation. To take account of the internal variations, conventional theories and methodologies are not sufficient to properly study and analyse symbolic data. Therefore, novel approaches for analysis on the new type of data are needed to be proposed and studied.

There are different types of symbolic data. Interval, multi-valued, histogram as well as distributions are the most common. Details of the first three types of symbolic data can be found in Chapter 2. Symbolic data mainly arise from two sources: the data are inherently symbolic, e.g., daily temperature of an area is [lowest temperature, highest temperature]; or data aggregation, e.g., the amount of monthly premiums for drivers aged within a certain range can be represented by a histogram.

Billard and Diday (2006) [1] provides a great number of examples as well as applications of symbolic data. It can be observed that two major challenges trigger the study and analysis

of symbolic data. The first comes from larger and larger data sets we face to deal with in the future. As the size and dimensions of a data set become huge, capacities of conventional analysis approaches are restricted with bad and inefficient performances due to limitations of computer power; the other is for solving some problems by data analysis. We need to focus more at group levels instead of individual levels. To conquer these challenges, by symbolic data analysis, we aggregate observations of classical data into groups with characteristics that are of great interest, and as a result, the data set is reorganized to be of reduced size and dimensions while retaining as much information as possible. One example to consider is a data set of individual medical records from a health insurance company. With demand-based aggregations, the original data set is converted into one with symbolic data, representing features such as marital status by gender groups, not those of individuals themselves. More details of this example can be found in Chapter 2.

Among different types of symbolic data, interval-valued data are one of the most common formats to be studied. For linear regression models for interval-valued data, several approaches have been proposed. The Center method, introduced by Billard and Diday (2000) [2] fits a linear regression model on the center points of the intervals; then predictions of the response were obtained by applying the fitted model to the lower and upper points of independent variables, respectively. The Center and Range method by de Carvalho et al. (2004) [3], Neto et al. (2005) [4] and Neto and de Carvalho (2008) [5], utilizes both centers and ranges of intervals to fit regression models. The Constrained method, proposed by Lima Neto et al. (2005, 2010) [6][7], sets constraints on the coefficient estimates to guarantee the upper bound of the predicted response is always not smaller than the lower bound of the predicted response; the Symbolic Covariance method, introduced by Xu (2010) [8], computes the coefficient estimates by means of the symbolic sample covariance. Among these approaches, most concentrate on the estimation of coefficients, while few have involved issues surrounding inference on the regression coefficient estimates.

This dissertation aims to propose a maximum likelihood estimator method on the coefficient estimates for linear regression of interval-valued data. The proposed method enables to study internal variations within an interval-valued observation by the principle of maximum likelihood, from where we obtain the point estimators of regression coefficients and their distributions. As a result, this approach can also give confidence intervals for the parameters in linear regression models.

The dissertation is organized as follows: Chapter 2 reviews the concept of symbolic data, the main types, recent studies on symbolic data analysis, and regression methods on interval-valued data. In Chapter 3, we propose the statistical inference approach by maximum likelihood principle for the coefficient estimates in interval data regression. In Chapter 4, simulations and results by the proposed method are studied; Chapter 5 contains examples with real data sets to evaluate performances of the proposed method. In Chapter 6, future research ideas are discussed.

Chapter 2

LITERATURE REVIEW

For the purpose of building the foundation for our work, a review of the literature is presented in this chapter. In Section 2.1, the concepts of symbolic data as well as sources where symbolic data arise are introduced. Section 2.2 discusses the main types of symbolic data and their descriptive statistics. Current studies surrounding symbolic data analysis are described in Section 2.3. Section 2.4 reviews several linear regression methods on interval valued data proposed in the literature, along with their advantages and disadvantages.

2.1 Symbolic Data

Unlike classical data on p random variables, which are represented by single points in p-dimensional space $\mathcal{R}^p$, realizations of symbolic data are represented by p-dimensional hypercubes in $\mathcal{R}^p$, or Cartesian products of p distributions. For example, in symbolic data, observations can take multi-values for a variable, e.g., the colors of a given species of birds can be {white, black}, etc. Symbolic data can be intervals, lists, histograms or distributions.

There are several sources where symbolic data arise. The first is when observations are inherently symbolic, such as the "colors of bird species" example in the above paragraph; the

second is when symbolic data arise by data aggregation. This occurs when we are interested in studying classes or groups. One example is about a data set comprising the medical records of individuals retained by a health insurance corporation, taken from Billard and Diday (2006) [1]. In the data set, values of several variables on geographic location information, such as region (north, south, etc.) and city (Boston, Atlanta, etc.) are recorded for each individual. It also contains some demographic variables, such as gender, marital status, age, health provider, etc. Another kind of variable included is about health: incidences of ailments and diseases, for instance. Table 2.1 (extracted from Billard and Diday, 2006) is a simplified table with entries to be classical data values as described above.

Table 2.1 - Classical data

ID	City/Town	Age	Gender	Marital Status	Weight	Pulse Rate	...
1	Boston	24	Male	Single	165	68	...
2	Boston	56	Male	Married	186	84	...
3	Chicago	48	Male	Married	175	73	...
4	El Paso	47	Female	Married	141	78	...

It should be noted that when the size of a data set becomes considerably large (e.g., n =100 million, $p > 100$), using conventional approaches to handle the data set may cause some problems. Firstly, the huge size of data challenges machines' capabilities to save and compute to conduct analysis; secondly, the primary importance of an analysis may not lie at the individual level but at groups with certain characteristics defined by some variables. Both of the above issues can be properly tackled by reorganizing the data set in the view of symbolic data. For instance, suppose the ages of married women is the list {29, 31, 33, 34, 42, 44, 47, 54, 61, 63, 64, 69, 71, 75, 82, 88}. These values can instead be represented as realizations within the interval [29, 88]; or the weight of this group of people can be represented as a histogram (also, from Billard and Diday, 2006): {[78, 110), 3/14; [110, 160),

7/14; [160, 170], 4/14}. Now, the variables "Age" and "Weight" for corresponding groups have become lists, or an interval, or a histogram, respectively.

Table 2.2 below shows the reorganized dataset from Table 2.1, composed of symbolic list and interval-valued data:

Table 2.2 - Symbolic data

ID	Marital Status × Gender	City/Town	Age	Weight	Pulse Rate	...
1	Single × Male	{Akron, Boston, Concord, Chicago, Marion, Quincy}	[6, 64]	[35, 268]	[57, 81]	...
2	Single × Female	{Amherst, Boston, Chicago, Ila, Medford, Yuma}	[11, 66]	[73, 166]	[62, 75]	...
3	Married × Male	{Albany, Atlanta, Barry, Bangor, Boston, Chicago, ...}	[24, 86]	[128, 239]	[59, 88]	...
4	Married × Female	{Atlanta, Boston, Buffalo, Byron, Detroit, El Paso, ...}	[21, 87]	[113, 178]	[58, 88]	...

Additionally, symbolic data also come from government census data, as well as confidentiality, such as selecting options of income ranges in a survey. Assume we want to add another column to Table 2.2 showing information on respondents' incomes. Because of its sensitive feature, people would only provide ranges covering their exact incomes in the survey. Hence, the new column named will contain intervals instead of single numbers to reflect their income levels, such as {[40, 50], [90, 100], [20, 30], [50, 70], [70, 90]...}.

2.2 Main Types of Symbolic Data

The main types of symbolic data include: interval-valued, multi- valued, histogram-valued, and distributions. Definitions as well as some examples to be introduced below are taken

from Billard and Diday (2006) [1].

Interval-valued Data

Among all the types of symbolic data, interval-valued data have been studied the most. One reason is it is the most common form of symbolic data; and it has been observed that methods to analyse interval-valued data can to be generalized to other types of symbolic data.

Note that all of the following definitions are based on the assumption that values across each interval are distributed uniformly. Denote $X_{(j)}$ to be the jth variable of a random sample $\mathbf{X}_i, i = 1, ..., n$, with the ith realization as the interval $[a_{ij}, b_{ij}] \subset \mathcal{R}$ where $a_{ij} \leq b_{ij}$, $j = 1, ..., p$. By the uniform assumption for a point in $X_{(j)}$ denoted by W, we have

$$P(W \leq \zeta) = \begin{cases} 0, \ \zeta \leq a_{ij}, \\ \frac{\zeta - a_{ij}}{b_{ij} - a_{ij}}, \ a_{ij} \leq \zeta \leq b_{ij}, \\ 1, \ b_{ij} \leq \zeta. \end{cases} \tag{2.1}$$

Further, Bertrand and Goupil (2000) [9] define the sample mean and sample variance of W, respectively, as

$$\overline{W} = \frac{1}{2n} \sum_{i=1}^{n} (a_{ij} + b_{ij}), \tag{2.2}$$

$$S^2 = \frac{1}{3n} \sum_{i=1}^{n} (a_{ij}^2 + a_{ij} b_{ij} + b_{ij}^2)^2 - \frac{1}{4n^2} [\sum_{i=1}^{n} (a_{ij} + b_{ij})]^2. \tag{2.3}$$

Billard (2007, 2008) [10] [11] show that the sample variance in (2.3) is a function of the total sum of squares (TSS) that can be divided into two terms: within sum of squares (WSS) to

represent the internal varition and between sum of squares (BSS) to represent the external variation:

$$nS^2 = TSS = WSS + BSS \tag{2.4}$$

where

$$WSS = \frac{1}{3}\sum_{i=1}^{n}[(a_{ij} - \overline{W}_{ij})^2 + (a_{ij} - \overline{W}_{ij})(b_{ij} - \overline{W}_{ij}) + (b_{ij} - \overline{W}_{ij})^2] \tag{2.5}$$

with

$$\overline{W}_{ij} = \frac{1}{2}(a_{ij} + b_{ij}), i = 1, ..., n, j = 1, ..., p,$$

and

$$BSS = \sum_{i=1}^{n}(\overline{W}_{ij} - \overline{W}_{(j)})^2 \tag{2.6}$$

where

$$\overline{W}_{(j)} = \frac{1}{2n}\sum_{i=1}^{n}(a_{ij} + b_{ij}), i = 1, ..., n, j = 1, ..., p.$$

From (2.5), we know that $a_{ij} - \overline{W}_{ij} = \overline{W}_{ij} - b_{ij} = \frac{a_{ij} - b_{ij}}{2}$. Thus, we have

$$WSS = \frac{1}{12}\sum_{i=1}^{n}(b_{ij} - a_{ij})^2. \tag{2.7}$$

Next, the sample covariance between two interval-valued variables can be shown similarly. Denote X_1 and X_2 to be two interval-valued random variables and assume $X_{i1} = [a_i, b_i]$, $X_{i2} = [c_i, d_i]$, $i = 1, ..., n$. From Billard (2008) [11], we have

$$\begin{aligned} Cov(X_1, X_2) &= \frac{1}{6n}\sum_{i=1}^{n}[2(a_i - \overline{X}_1)(c_i - \overline{X}_2) + (a_i - \overline{X}_1)(d_i - \overline{X}_2) \\ &+ (b_i - \overline{X}_1)(c_i - \overline{X}_2) + 2(b_i - \overline{X}_1)(d_i - \overline{X}_2)] \end{aligned} \tag{2.8}$$

where

$$\overline{X}_1 = \frac{1}{2n}\sum_{i=1}^{n}(a_i + b_i), \overline{X}_2 = \frac{1}{2n}\sum_{i=1}^{n}(c_i + d_i).$$

The correlation coefficient between X_1 and X_2 is then defined as

$$Corr(X_1, X_2) = \frac{Cov(X_1, X_2)}{S_1 S_2} \tag{2.9}$$

where S_1 and S_2 are the square roots of the sample variances of X_1 and X_2, respectively.

Multi-valued Data

Assume X is a multi-valued symbolic random variable, with each possible value taking one or more values from the list of values in its domain $\mathcal{X}$. We assume the complete list of possible values in $\mathcal{X}$ to be finite, and it may include well-defined quantitative or categorical values.

Taking the case of medical records retained by a health insurance corporation illustrated in Section 2.1 as an example, we can see in Table 2.2 that after aggregation of the original data, the "City/Town" variable is multi-valued with values to be lists of a number of cities in which individual of different groups live. Descriptive statistics for multi-valued data can be found in Bertrand and Goupil (2000) [9].

Histogram-valued Data

Suppose X to be a quantitative randome variable that takes values on a finite number of non-overlapping intervals $[a_i, b_i)$, $i = 1, 2, ...,$ with $a_i \leq b_i$. Then the realization of an observation for this variable has the form

$$\xi_u = \{[a_i, b_i), p_{ui}; i = 1, ..., s_u\} \tag{2.10}$$

where $s_u < \infty$ is the number of intervals forming the support for the realization for observation ξ_u, and p_{ui} is the weight for the subinterval $[a_{ui}, b_{ui})$, $i = 1, ..., s_u$, with $\sum_{k=1}^{s_u} p_{ui} = 1$.

An example of histogram-valued data was given in Billard (2011) [12]. To describe claims (in \$1000's) from 35-year-old females after data aggregation from an original automobile insurance data set, it is more appropriate to use a histogram-valued realization, or called "histogram data":

$$Y = \{[0,2), 0.05; [2,4), 0.25; [4,6), 0.45; [6,8), 0.20; [8,10], 0.05\}. \tag{2.11}$$

This cannot be interval-valued data, since the assumption that values within the interval are uniformly distributed across the interval cannot be satisfied in this context; now, the frequencies' values in different subintervals vary.

We can generalize (2.10) to a p-dimensional scenario easily. Suppose $\boldsymbol{X} = (X_1, \cdots, X_p)$ is a vector of histogram-valued random variables. Then for each observation w_u, the variable $X_j(u)$ takes values

$$X_j(w_u) = \{[a_{ujk}, b_{ujk}), p_{ujk}, k = 1, ..., s_{uj}\} \tag{2.12}$$

where the non-overlapping intervals $\xi_{ujk} = [a_{ujk}, b_{ujk})$ have relative frequencies p_{ujk}, $k = 1, ..., s_{uj}$, with $\sum_{k=1}^{s_{uj}} p_{ujk} = 1$ and s_{uj} is the number of subintervals in the histogram.

Based on the assumption that all values within each subinterval $[a_{ijk}, b_{ijk})$, $k = 1, ..., s_{ij}$, $i = 1, ..., n$, $j = 1, ..., p$, are uniformly distributed, Billard and Diday (2003) [13] defined the empirical distribution function of a histogram-valued variable as well as the sample variance for histogram-valued data, respectively. The empirical density function of X_j is expressed as

$$f_{X_j}(\xi) = \frac{1}{n} \sum_{u=1}^{n} \sum_{k:\xi \in \xi_{ujk}} p_{ujk} \frac{1}{b_{ujk} - a_{ujk}}. \tag{2.13}$$

The symbolic sample mean for X_j is

$$\bar{X_j} = \frac{1}{2n} \sum_{u=1}^{n} \sum_{k=1}^{s_{uj}} p_{ujk}(a_{ujk} + b_{ujk}) \tag{2.14}$$

and the sample variance S_j^2 is

$$S_j^2 = \frac{1}{3n}\sum_{u=1}^{n}\sum_{k=1}^{s_{uj}} p_{ujk}[a_{ujk}^2 + a_{ujk}b_{ujk} + b_{ujk}^2] - \frac{1}{4n^2}[\sum_{u=1}^{n}\sum_{k=1}^{s_{uj}} p_{ujk}(a_{ujk} + b_{ujk})]^2. \quad (2.15)$$

2.3 Studies on Symbolic Data Analysis

Current studies surrounding symbolic data analysis focus on the following aspects: (1) Principle component analysis methods for interval-valued data have been studied and developed by Cazes et al. (1997) [14], Lauro and Palumbo (2000) [15], Billard et al. (2008) [16] and Le-Rademacher and Billard(2012, 2013, 2016) [17] [18] [19]; (2) Factorial discriminant analysis methods for interval-valued data were proposed by Palumbo and Verde (2000) [20] and Lauro et al. (2000) [21] and were generalized to deal with face recognition by Hiremath and Prabhakar (2008) [22]; Silva and Brito (2006) [23] studied linear discriminant analysis for interval data and recently Silva and Brito (2015) [24] further proposed parametric and distance-based approaches for discriminant analysis of interval data; (3) Multidimensional scaling methods to deal with interval-valued and fuzzy dissimilarity data were proposed by Denoeux and Masson (2000) [25] and Masson and Denoeux (2002) [26], respectively. It was further introduced by Groenen et al. (2006) [27] and developed by Huang et al. (2006) [28] using a rough set concept. Terada and Yadohisa (2011) [29] proposed a multidimensional scaling method with the nested hypersphere model for percentile dissimilarities in a different direction; (4) In terms of classification methods, Ichino et al. (1996) [30] introduced a symbolic classifier as a region-oriented approach and Rasson and Lissoir (2000) [31] proposed a symbolic kernel classifier based on dissimilarity functions for interval-valued data; a tree-growing algorithm for classification was introduced by Perinel and Lechevallier (2000) [32]; Dinesh, Gowda and Nagabhushan (2005) [33] proposed a new generalized similarity symbolic distance measure for classification and Maia et al. (2008) [34] studied approaches

to interval-valued time series forecasting; (5) For clustering methods on symbolic data, a number of methods considering different types of symbolic data as well as clustering criteria were proposed in the following major articles: Gowda and Diday (1991, 1992) [35] [36] illustrated an agglomerative approach that forms composite symbolic objects by a joint operator based on minimum dissimilarity or maximum similarity in hierarchical clustering methods; and Ichino and Yaguchi (1994) [37] defined generalized Minkowski metrics for mixed feature variables and displayed dendrograms by standard linkage methods. Chavent (1998) [38] introduced a divisive clustering method which simultaneously performs hierarchy of objects and a monothetic characterization of each cluster; Guru et al. (2004) [39] and Guru and Kiranagi (2005) [40] proposed agglomerative clustering algorithms based on similarity and dissimilarity, respectively; and Kiranagi and Guru (2010) [41] introduced a new statistical measure for estimating the degree of dissimilarity between two symbolic objects with features of multivalued type and proposed interval type and magnitude type as two new simple representation techniques for dissimilarity computation;
(6) Regarding partitioning clustering algorithms for interval-valued data, Bock (2002) [42] proposed several clustering algorithms and Kohonen maps for symbolic data, Chavent and Lechevallier (2002) [43] introduced a dynamic clustering algorithm for interval-valued data, Souza and Carvalho (2004) [44] proposed partitioning clustering methods for interval-valued data based on city-block distances. De Carvalho et al. (2006) [45] presented a partitional dynamic clustering method for interval data based on adaptive Hausdorff distances, and De Carvalho (2007) [46] introduced adaptive and non-adaptive fuzzy clustering c-means methods. De Souza et al. (2006) [47] proposed a partitioning method for mixed feature-type symbolic data and Pimentel et al. (2011) [48] proposed a K-means clustering method based on kernelized squared L_2 distance for interval valued data.

In the framework of symbolic data analysis, a number of approaches on fitting linear regression models to interval-valued data have been proposed, which will be introduced in

detail in the next section.

2.4 Regression Methods on Symbolic Data

So far, studies of linear regression methods on symbolic data mainly surround interval-valued data and several approaches have been proposed.

The initial method was introduced by Billard and Diday (2000) [2] which is to use the centers of intervals to fit the regression model; Lima Neto et al. (2004) [4], de Carvalho et al. (2004) [3] and Lima Neto and de Carvalho (2008) [5] then developed methods utilizing both centers and ranges of intervals to fit regression models. Billard and Diday (2006) [1] proposed an approach as an improvement of previous methods, which is to use centers and ranges simultaneously in model fitting. Later on, a constrained method was introduced by Lima Neto et al. (2005, 2010) [6] [7]. Recently, another method called "Symbolic Covariance Method" was proposed by Xu (2010) [8]. This method utilizes the symbolic sample covariance of (2.8) to compute the coefficient estimates in the regression equation. In this section, these methods are reviewed briefly based on different categories.

The Center Method

Billard and Diday (2000) [2] proposed the first approach to fit a linear regression model to interval-valued data. They fitted regression models using the centers of intervals by classical methods and made predictions on the response variable by means of applying the model to both lower and upper bounds of a new interval-valued observation. This approach is the so-called center method (CM).

Denote X_j to be the jth variable among p independent interval-valued variables, and Y is the response variable. Let $X_{ij} = [a_{ij}, b_{ij}]$ and $Y_i = [c_i, d_i]$ be the ith observation of the variable X_j and the ith observation of the response, respectively, where $i = 1, ..., n, j = 1, ..., p$.

The center points of X_{ij} and Y_i can be expressed as

$$X_{ij}^c = \frac{a_{ij} + b_{ij}}{2}, Y_i^c = \frac{c_i + d_i}{2}, i = 1, ..., n, j = 1, ..., p. \tag{2.16}$$

The fitted regression model is

$$\mathbf{Y}^c = \mathbf{X}^c \boldsymbol{\beta}^c + \boldsymbol{\epsilon}^c \tag{2.17}$$

where $\mathbf{Y}^c = (Y_1^c, ..., Y_n^c)'$, $\mathbf{X}^c = (X_1^c, ..., X_n^c)', \boldsymbol{\beta}^c = (\beta_0, \beta_1, ..., \beta_p)$, $\boldsymbol{\epsilon} = (\epsilon_1, ..., \epsilon_n)'$ and $X_i^c = (1, X_{i1}^c, ..., X_{ip}^c)'$ for $i = 1, ..., n$.

As for the linear regression model for classical data, the least squares estimator of $\boldsymbol{\beta}^c$ is

$$\hat{\boldsymbol{\beta}}^c = ((X^c)' X^c)^{-1} (X^c)' Y^c \tag{2.18}$$

with the condition that $\mathbf{X}^c$ has full rank $(p + 1) \leq n$.

For a given observation $\mathbf{X}^* = (1, X_1^*, ..., X_p^*)$, where $X_j^* = [X_{jL}^*, X_{jU}^*], j = 1, ..., p$, the predicted value of the response is

$$\hat{Y}^* = [\hat{Y}_L, \hat{Y}_U] = [\mathbf{X}_L^* \hat{\boldsymbol{\beta}}^c, \mathbf{X}_U^* \hat{\boldsymbol{\beta}}^c] \tag{2.19}$$

where $\mathbf{X}_L^* = (1, X_{1L}^*, ..., X_{nL}^*)$ and $\mathbf{X}_U^* = (1, X_{1U}^*, ..., X_{nU}^*)$.

This approach only takes the center points of intervals into consideration when calculating the parameters, while it ignores other important information such as the internal variation within each observation.

Center and Range Method

The center and range method (CRM method) was introduced by Neto et al. (2004, 2008) [4] [5] and de Carvalho et al. (2004) [3] to estimate the parameter β using not only center

points but also ranges of intervals. In this method, the regression model on center points is the same as in the CM method above, which has the form as (2.11), and the parameter estimate can be expressed by (2.12).

In addition to the model built on center points, the CRM method considers another model on the ranges of the intervals. Let X_{ij}^r and Y_i^r, $i = 1, ..., n, j = 1, ..., p$, be the ranges of the interval-valued data, with $X_{ij}^r = (b_{ij} - a_{ij}), Y_i^r = (d_i - c_i)$. Then, the regression model on the range is

$$\mathbf{Y}^r = \mathbf{X}^r \boldsymbol{\beta}^r + \boldsymbol{\epsilon}^r \tag{2.20}$$

where $\mathbf{Y}^r = (Y_1^r, ..., Y_n^r)'$, $\mathbf{X}^r = (1, X_1^r, ..., X_n^r)', \boldsymbol{\beta}^r = (\beta_0^r, \beta_1^r, ..., \beta_p^r)$, $\boldsymbol{\epsilon} = (\epsilon_1^r, ..., \epsilon_n^r)'$ and $X_i^r = (1, X_{i1}^r, ..., X_{ip}^r)'$ for $i = 1, ..., n$.
As in (2.12), the least squares estimate of $\boldsymbol{\beta}^r$ is

$$\hat{\boldsymbol{\beta}}^r = ((X^r)' X^r)^{-1} (X^r)' Y^r \tag{2.21}$$

with the condition that $\mathbf{X}^r$ has full rank $(p + 1) \leq n$.

For a given observation $\mathbf{X}^* = (X_1^*, ..., X_p^*)$, where $X_j^* = [a_j^*, b_j^*], j = 1, ..., p$, the predicted value of the response is given by

$$\hat{Y}^* = [\hat{Y}_L, \hat{Y}_U] = [\hat{Y}^c - \frac{\hat{Y}^r}{2}, \hat{Y}^c + \frac{\hat{Y}^r}{2}] \tag{2.22}$$

where $\hat{Y}^c = \mathbf{X}^{c*} \hat{\boldsymbol{\beta}}^c$, $\hat{Y}^r = \mathbf{X}^{r*} \hat{\boldsymbol{\beta}}^r$, and $\mathbf{X}^{c*} = (X_1^{c*}, ..., X_p^{c*})$, $X_j^{c*} = \frac{a_j^* + b_j^*}{2}$, $\mathbf{X}^{r*} = (X_1^{r*}, ..., X_p^{r*})$, $X_j^{r*} = b_j^* - a_j^*$, $j = 1, ..., p$.

Bivariate Center and Range Method

For the CRM method, it is assumed that center points and ranges are independent and it builds two regression models on each of them separately. In Billard and Diday (2006) [1], the center points and ranges are considered simultaneously and a bivariate model was created without assuming they are independent. In this approach, we have the model form to be

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{2.23}$$

where $\mathbf{Y} = (\mathbf{Y}^{c'}, \mathbf{Y}^{rT})'$, $\mathbf{Y}^c = (Y_1^c, ..., Y_n^c)'$ and $\mathbf{Y}^r = (Y_1^r, ..., Y_n^r)'$, representing for the center points and the ranges of response values; and $\mathbf{X} = (\mathbf{X}_1, ..., \mathbf{X}_n)' = (\mathbf{X}^c, \mathbf{X}^r)$, $\mathbf{X}_i = (\mathbf{1}, \mathbf{X}_i^c, \mathbf{X}_i^r) = (1, X_{i1}^c, ..., X_{ip}^c, X_{i1}^r, ..., X_{ip}^r)$ for $i = 1, ..., n$, where X_{ij}^c, X_{ij}^r, $i = 1, ..., n; j = 1, ..., p$, represent values of the center points and ranges of the ith observation on the jth predictor, and $\boldsymbol{\beta} = (\beta_0, \beta_1^c, ..., \beta_p^c, \beta_1^r, ..., \beta_p^r)'$. On the condition that $\mathbf{X}$ is of full rank $(2p+1) \leq n$, the least squares estimator of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}^c, \hat{\boldsymbol{\beta}}^r)') = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}. \tag{2.24}$$

For a given observation $\mathbf{X}^* = (X_1^*, ..., X_p^*)$ with the center point and the range to be $\mathbf{X}^{*c} = (1, X_1^{*c}, ..., X_p^{*c})$ and $\mathbf{X}^{*r} = (X_1^{*r}, ..., X_p^{*r})$, the predicted value of the response is

$$\hat{Y}^* = [\hat{Y}_L, \hat{Y}_U] = [\hat{Y}^c - \frac{\hat{Y}^r}{2}, \hat{Y}^c + \frac{\hat{Y}^r}{2}] = [\mathbf{X}^{*c}\hat{\boldsymbol{\beta}}^c - \frac{1}{2}\mathbf{X}^{*r}\hat{\boldsymbol{\beta}}^r, \mathbf{X}^{*c}\hat{\boldsymbol{\beta}}^c + \frac{1}{2}\mathbf{X}^{*r}\hat{\boldsymbol{\beta}}^r]. \tag{2.25}$$

Additionally, Billard and Diday (2006) [1] considers interactions between center points and ranges. In this approach, interaction terms between the center point and range are added into the model. The ith predictor vector of the n observations is expressed as

$$(1, X_{i1}^c, ..., X_{ip}^c, X_{i1}^r, ..., X_{ip}^r, X_{i1}^c \times X_{i1}^r, ..., X_{ip}^c \times X_{ip}^r)'. \tag{2.26}$$

Other settings of the model remains the same as the model without interaction considerations in (2.23).

The center and range methods take both center points and ranges into account in fitting linear regression models on interval-valued data. However, one problem of these methods is the value of the coefficient estimate for range cannot be guaranteed to be always positive. Thus, a negative value will result in the situation that the lower bound of predicted response value is larger than the upper bound.

Constrained Method

For the purpose of solving the problem identified at the end of Section 2.4.3, Neto et al. (2005, 2010) [6] [7] proposed the constrained method (CONM method). Within this framework, the constrained center method was introduced first and then it was further developed as the constrained center and range method.

For the constrained center method, the model has the form

$$\mathbf{Y}^c = \mathbf{X}^c\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{2.27}$$

where $\boldsymbol{\beta} = (\beta_0, ..., \beta_p)'$ with constraints $\beta_j \geq 0, j = 0, ..., p$. For the constrained center and range method, the model settings are the same as those in the bivariate center and range method, except for constraints $\beta_j^r \geq 0, j = 0, ..., p$. To guarantee the constraints of non-negativeness on $\boldsymbol{\beta}$, Neto et al. (2005, 2010) [6] [7] used an algorithm by Lawson and Hanson (1974) [49] and modified it to adapt to the constrained center and range method. This algorithm identifies the negative elements of the least squares estimate of coefficient vector $\boldsymbol{\beta}$ and changes them to non-negative values by a process of re-weighting.

Though this approach prevents the situation that the lower bound of the predicted response value is larger than the upper bound from happening, it fails to discover the real

nature of the data when the true parameter $\beta_j < 0$, so therefore provides improper estimates.

Symbolic Covariance Method

Xu (2010) [8] proposed a symbolic covariance method (SCM) to fit linear regression models on interval-valued data. By utilizing the symbolic sample covariance of (2.8), this method makes full use of both external and internal variations of data. To illustrate this method, we use the same notations as in Xu (2010) [8] below.

In the situation of classical data where p predictor variables are considered, we have the model to be

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \epsilon. \tag{2.28}$$

Let

$$\beta_0 \equiv \bar{Y} - (\beta_1 \bar{X}_1 + \cdots + \beta_p \bar{X}_p). \tag{2.29}$$

Then (2.28) can be written as

$$Y - \bar{Y} = \beta_1 (X_1 - \bar{X}_1) + \cdots + \beta_p (X_p - \bar{X}_p) + \epsilon \tag{2.30}$$

where $\bar{X}_j = \frac{1}{n}\sum_{i=1}^{n} X_i$, $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$, and n is the number of observations. Based on the form (2.30), the least squares coefficient vector estimate is obtained as

$$\hat{\boldsymbol{\beta}} = ((\mathbf{X} - \bar{\mathbf{X}})'(\mathbf{X} - \bar{\mathbf{X}}))^{-1}(\mathbf{X} - \bar{\mathbf{X}})'(\mathbf{Y} - \bar{\mathbf{Y}}) \tag{2.31}$$

where $\mathbf{X} = (\mathbf{1}, \mathbf{X}_1, ..., \mathbf{X}_p), \mathbf{X}_i = (1, X_{i1}, ..., X_{in})', i = 1, ..., p$, and $Y = (Y_1, ..., Y_n)'$. Then, we

have

$$
\begin{aligned}
(\mathbf{X}-\bar{\mathbf{X}})'(\mathbf{X}-\bar{\mathbf{X}}) &= \begin{pmatrix} \sum_{i=1}^n (X_{1i}-\bar{X}_1)^2 & \cdots & \sum_{i=1}^n (X_{1i}-\bar{X}_1)(X_{pi}-\bar{X}_p) \\ \vdots & \vdots & \vdots \\ \sum_{i=1}^n (X_{pi}-\bar{X}_p)(X_{1i}-\bar{X}_1) & \cdots & \sum_{i=1}^n (X_{pi}-\bar{X}_p)^2 \end{pmatrix}_{p\times p} \\
&= \left(\sum_{i=1}^n (X_{j_1 i}-\bar{X}_{j_1})(X_{j_2 i}-\bar{X}_{j_2})\right)_{p\times p} \\
&= (n\times Cov(X_{j_1}, X_{j_2}))_{p\times p},\ j_1, j_2 = 1, ..., p, \qquad (2.32)
\end{aligned}
$$

and

$$
\begin{aligned}
(\mathbf{X}-\bar{\mathbf{X}})'(\mathbf{Y}-\bar{\mathbf{Y}}) &= \left(\sum_{i=1}^n (X_{ji}-\bar{X}_j)(Y_i-\bar{Y})\right)_{p\times 1} \\
&= (n\times Cov(X_j, Y))_{p\times 1},\ i, j = 1, ..., p. \qquad (2.33)
\end{aligned}
$$

Substituting (2.32) and (2.33) to (2.31), we obtain

$$
\hat{\boldsymbol{\beta}} = (n\times Cov(X_{j1}, X_{j2}))^{-1}_{p\times p} \times (n\times Cov(X_j, Y))_{p\times 1}. \qquad (2.34)
$$

Assume for the ith interval-valued observation, the response value and the jth predictor values are $Y_i = [c_i, d_i]$ and $X_{ij} = [a_{ij}, b_{ij}]$, respectively, $i = 1, ..., n, j = 1, ..., p$. By (2.8), we can obtain

$$
Cov(X_{j_1}, X_{j_2}) = \frac{1}{6n}\sum_{i=1}^n [2(a_{ij_1}-\bar{X}_{j_1})(a_{ij_2}-\bar{X}_{j_2})+(a_{ij_1}-\bar{X}_{j_1})(b_{ij_2}-\bar{X}_{j_2})+(b_{ij_1}-\bar{X}_{j_1})(b_{ij_2}-\bar{X}_{j_2})] \qquad (2.35)
$$

and

$$
Cov(X_j, Y) = \frac{1}{6n}\sum_{i=1}^n [2(a_{ij}-\bar{X}_j)(c_i-\bar{Y}) + (a_{ij}-\bar{X}_j)(d_i-\bar{Y}) + (b_{ij}-\bar{X}_j)(d_i-\bar{Y})] \qquad (2.36)
$$

where

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n}\frac{c_i + d_i}{2}, \bar{X}_{j_1} = \frac{1}{n}\sum_{i=1}^{n}\frac{a_{ij_1} + b_{ij_1}}{2}, \bar{X}_{j_2} = \frac{1}{n}\sum_{i=1}^{n}\frac{a_{ij_2} + b_{ij_2}}{2}, \bar{X}_j = \frac{1}{n}\sum_{i=1}^{n}\frac{a_{ij} + b_{ij}}{2}$$

and $i = 1, ..., n; j, j_1, j_2 = 1, ..., p$. Substituting (2.35) and (2.36) into (2.34), we obtain the estimate of $(\beta_1, ..., \beta_p)$, and further the estimate of β_0 from the equation

$$\hat{\beta}_0 = \bar{Y} - (\hat{\beta}_1\bar{X}_1 + \cdots + \hat{\beta}_p\bar{X}_p). \tag{2.37}$$

If we denote the lower bound and upper bound of a new interval-valued realization for the predictor to be $\mathbf{X}_L^{new}$ and $\mathbf{X}_U^{new}$, respectively, and the coefficient estimate to be $\hat{\boldsymbol{\beta}}$, then the SCM method chooses the smaller value and the larger value of $\{\mathbf{X}_L^{new}\hat{\boldsymbol{\beta}},\ \mathbf{X}_U^{new}\hat{\beta}\}$ to be the lower and upper bounds of the predicted response respectively, which avoids the situation that the predicted lower bound is larger than the predicted upper bound for the response. Based on the simulation study by Xu (2010) [8], the proposed method has superior performance in terms of estimation compared to previous methods.

2.5 APPENDIX

R Function to Calculate the Symbolic Variance-Covariance Matrix

```
sym_cov <- function(...){
  vars <- list(...)
  p <- length(vars)
  m <- length(vars[[1]][,1])
  # x <- rep(0, m*p*2)
  # dim(x) <- c(m, 2, p)
  cov <- matrix(0, p, p) # covariance matrix
  corr <- matrix(0, p, p) # correlation matrix
  tmp <- matrix(0, p, 2)

  x_mean <- rep(0, p) # variable means

  # calculate the means
  for (i in 1:p){
    for (j in 1:2){
      tmp[i,j] <- mean(vars[[i]][,j])
    }
    x_mean[i] <- mean(tmp[i,])
  }

  # calculate variance-covariance matrix of all pairs of variables
  for (k in 1:p){
    for (l in 1:p){
```

```
      q <- 0
      for (r in 1:m){
        q <- q + 2*(vars[[k]][r,1]-x_mean[k])*(vars[[l]][r,1]-x_mean[l])
               + (vars[[k]][r,1]-x_mean[k])*(vars[[l]][r,2]-x_mean[l])
               + (vars[[k]][r,2]-x_mean[k])*(vars[[l]][r,1]-x_mean[l])
               + 2*(vars[[k]][r,2]-x_mean[k])*(vars[[l]][r,2]-x_mean[l])
      }
      cov[k,l] <- q/6/m
    }
  }

  # correlation matrix
  for (k in 1:p){
    for (l in 1:p){
      corr[k,l] <- cov[k,l] / sqrt(cov[k,k]) / sqrt(cov[l,l])
    }
  }

  return(list(cov, corr))
  # return(x_mean)
}
```

Chapter 3

LIKELIHOOD METHOD FOR INTERVAL DATA REGRESSION

In Chapter 2, we reviewed previous studies of linear regression methods for symbolic interval-valued data. Only a few of them have dealt with issues surrounding inference on the regression coefficient estimates. In addition, all the information in the data is not fully utilized by those methods. In this chapter, we introduce a novel approach, including point estimation as well as confidence intervals for interval data regression methods.

The remainder of this chapter is arranged as follows. Section 3.1 demonstrates the problem, by outlining the basic settings of interval-valued regression models. Section 3.2 first illustrates the residual forms, with two different assumptions on the residuals for interval-valued regression analyses; then this section demonstrates our approach for point estimation and confidence interval of regression coefficients. Section 3.3 gives predictions by the new approach. Section 3.4 discusses how to determine the likelihood function form used for statistical inference.

3.1 Introduction

Assume we have n observations in a data set with response variable Y and p explanatory variables $X_1, ..., X_p$. Denote X_j to be the jth variable among the p explanatory interval-valued variables, X_{ij} to be the ith observation of the jth variable, and Y_i to be the ith observation of the response. Let the lower cases x_{ij}, y_i denote the realizations of X_{ij} and Y, respectively. Then, $x_{ij} = [x_{Lij}, x_{Uij}]$ and $y_i = [y_{Li}, y_{Ui}]$ with $x_{Lij} \leq x_{Uij}$, $y_{Li} \leq y_{Ui}$ can represent the ith realization of variable X_j and the ith realization of the response, respectively, where $i = 1, ..., n$, $j = 1, ..., p$.

For a linear regression model, based on the above notation the interval-valued design matrix, denoted by $\mathbf{X}$, has the form as follows:

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix} = \begin{pmatrix} 1 & [x_{L11}, x_{U11}] & \cdots & [x_{L1p}, x_{U1p}] \\ 1 & \vdots & \ddots & \vdots \\ 1 & [x_{Ln1}, x_{Un1}] & \cdots & [x_{Lnp}, x_{Unp}] \end{pmatrix} \tag{3.1}$$

where $x_{Lij} \leq x_{Uij}$ for all $i = 1, ..., n$, and $j = 1, ..., p$, with $p < n$. Note that the point value in the first column of $\mathbf{X}$, $x_0 = 1$ can be written as the interval $x_0 = [1, 1]$. The response variable $\mathbf{Y}$ has the form

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} [y_{L1}, y_{U1}] \\ \vdots \\ [y_{Ln}, y_{Un}] \end{pmatrix}. \tag{3.2}$$

The linear regression model for interval-valued data is as follows:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{3.3}$$

where the parameter vector $\boldsymbol{\beta} = (\beta_0, \beta_1, ..., \beta_p)'$ with β_j denoting the effect of the jth explanatory variable X_j to the response variable Y, for $j = 1, ..., p$, and β_0 represents the

intercept. The error term $\boldsymbol{\epsilon}$ has the form

$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix} = \begin{pmatrix} [\epsilon_{L1}, \epsilon_{U1}] \\ \vdots \\ [\epsilon_{Ln}, \epsilon_{Un}] \end{pmatrix}. \tag{3.4}$$

For now we consider the scenario of a simple linear regression model. The proposed idea can be extended without losing generality to multiple regression models.

The simple linear regression model for interval-valued data can be written as:

$$[Y_{Li}, Y_{Ui}] = \beta_0 + [X_{Li}, X_{Ui}]\beta_1 + [\epsilon_{Li}, \epsilon_{Ui}],\ i = 1, ..., n, \tag{3.5}$$

where X_{Li}, Y_{Li} and X_{Ui}, Y_{Ui} represent realizations of lower points and upper points of the explanatory and response variables, respectively; β_0 and β_1 are parameters of the intercept and the slope, respectively; the error term is interval-valued, with ϵ_{Li} being the error for the lower bound of response, and ϵ_{Ui} being the error for the upper bound of response, respectively.

For the error term, there are two different assumptions to be given in the next section. Each of the assumptions can be considered appropriate to describe interval-valued data, depending on different ways the data sets arise.

3.2 Methodology

In this section, we first illustrate two different assumptions on the error term, together with the corresponding forms of residuals. Then, we propose an approach to obtain point estimators and confidence intervals for the regression coefficients in interval-valued regression models by the second assumption. The method is developed utilizing the maximum likelihood principle. By this approach, distributions of the coefficient estimators can be obtained, and

issues surrounding point estimation as well as confidence interval are resolved based on the theoretical results.

Assumptions on the Error Term and Forms of Errors

Assumption I: Order Statistic

For the first assumption, let us suppose for the ith observation, the error for the lower response and the error for the upper response are dependent, for $i = 1, ..., n$. Using the same notation as in (3.5), we assume $\epsilon_{Li} \leq \epsilon_{Ui}$, and further, ϵ_{Li} and ϵ_{Ui} are the order statistics of a random sample of size two from a normal distribution with mean zero and constant variance, for $i = 1, ..., n$; i.e.,

$$\epsilon_{Li} = \epsilon_{(1)i}, \epsilon_{Ui} = \epsilon_{(2)i}, \quad \epsilon_{(1)i}, \epsilon_{(2)i} \sim N(0, \sigma^2), \ i = 1, ..., n. \tag{3.6}$$

We consider the assumption of (3.6) is appropriate according to the following aspects:

1) Based on basic assumptions of interval-valued data, realizations are uniformly distributed within an interval observation but normally distributed across observations. Therefore, the error term of the linear regression model, coming from the difference between the interval-valued response and the linear combination of interval-valued covariates, can be normally distributed with mean zero.

2) Similar to classical linear regression models, from the perspective of residuals, the error $\boldsymbol{\epsilon}$ represents variation in the response variable $\mathbf{Y}$ which is not explained by the predictors. Thus it is advisable to assume normality on the remaining variability after removing the effects of the predictors.

3) Under some circumstances, the lower and upper residuals for the same interval-valued observation by a linear model are dependent, and so we can initially consider that the lower residual is smaller than or equal to the upper residual. The order statistic assumption on

the error term is consistent to this condition.

4) Constant variance of the error term indicates the variance does not change across different levels of the predictors, which is essential to guarantee that the linear model properly describes the relationship between the explanatory variable and the response variable, and provides advisable conclusions.

By (3.5), in order to ensure ϵ_{Li} is no larger than ϵ_{Ui}, for $i = 1, ..., n$, the forms of $[\epsilon_{Li}, \epsilon_{Ui}]$ are given depending on the sign of the slope parameter β_1, under universal conditions as follows:

$$Y_{Li} \leq Y_{Ui},\ X_{Li} \leq X_{Ui},\ \hat{Y}_{Li} = \hat{\beta}_0 + \hat{\beta}_1 X_{Li},\ \hat{Y}_{Ui} = \hat{\beta}_0 + \hat{\beta}_1 X_{Ui}, \tag{3.7}$$

for $i = 1, ..., n$, where $\hat{\beta}_0$ and $\hat{\beta}_1$ are coefficient estimators for the intercept and the slope, respectively. To illustrate how we obtain the forms of residuals, we generate Figures 3.1 - 3.12 below as schematic diagrams. Points on each figure represents a certain interval-valued observation in the sample, with the index i.

(1) For $\boldsymbol{\beta}_1 \geq 0$

Since the slope is not less than zero, $\hat{Y}_{Li}$ and $\hat{Y}_{Ui}$ are the predicted lower and upper points of the ith response value by the linear regression, for $i = 1, ..., n$.

Scenario I

In this scenario, the observed response interval (y_{Li}, y_{Ui}) is contained within the predicted interval $(\hat{y}_{Li}, \hat{y}_{Ui})$ as illustrated in Figure 3.1.

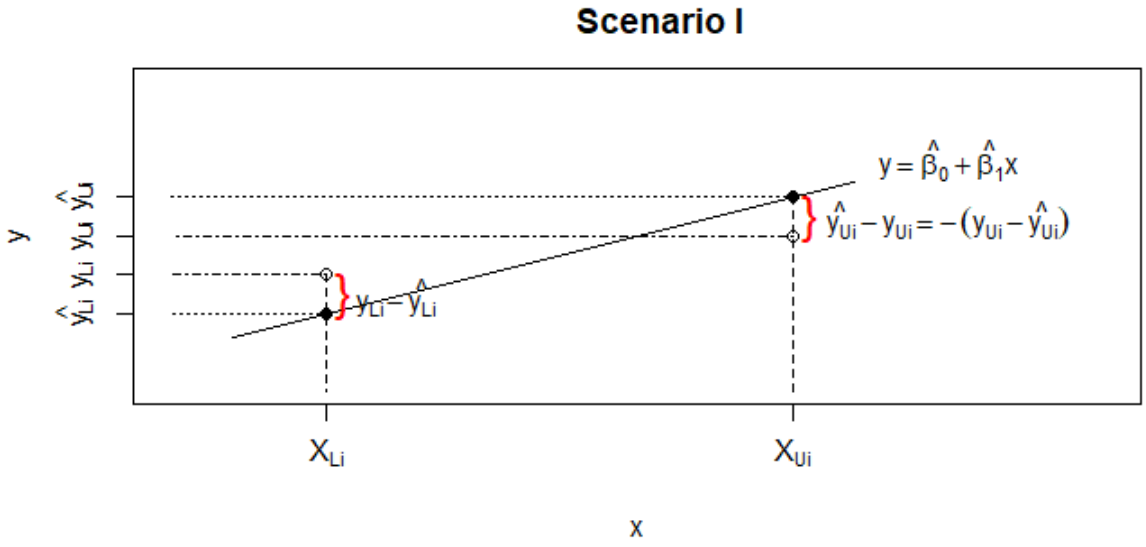

Figure 3.1: Scenario I, when $\beta_1 \geq 0$

From Figure 3.1 above, we have

$$Y_{Ui} - \hat{Y}_{Ui} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui} < 0 < Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li} = Y_{Li} - \hat{Y}_{Li}. \tag{3.8}$$

Then to satisfy the basic assumption that $\epsilon_{Li} \leq \epsilon_{Ui}$, we denote

$$\begin{cases} \epsilon_{Li} \triangleq min\{Y_{Ui} - \hat{Y}_{Ui}, Y_{Li} - \hat{Y}_{Li}\} = Y_{Ui} - \hat{Y}_{Ui} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}, \\ \epsilon_{Ui} \triangleq max\{Y_{Ui} - \hat{Y}_{Ui}, Y_{Li} - \hat{Y}_{Li}\} = Y_{Li} - \hat{Y}_{Li} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}. \end{cases} \tag{3.9}$$

Scenario II

In this scenario, the predicted response interval $(\hat{y}_{Li}, \hat{y}_{Ui})$ is contained within the observed interval (y_{Li}, y_{Ui}) as illustrated in Figure 3.2.

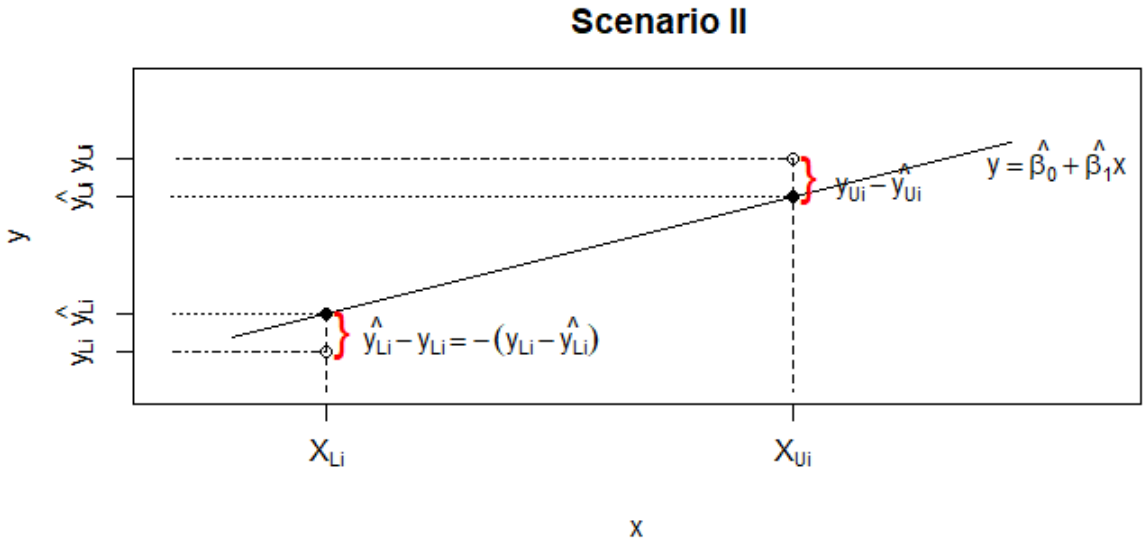

Figure 3.2: Scenario II, when $\beta_1 \geq 0$

From Figure 3.2, we have

$$Y_{Li} - \hat{Y}_{Li} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li} < 0 < Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui} = Y_{Ui} - \hat{Y}_{Ui}. \tag{3.10}$$

To satisfy the basic assumption that $\epsilon_{Li} \leq \epsilon_{Ui}$, we need that

$$\begin{cases} \epsilon_{Li} \triangleq min\{Y_{Ui} - \hat{Y}_{Ui}, Y_{Li} - \hat{Y}_{Li}\} = Y_{Li} - \hat{Y}_{Li} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}, \\ \epsilon_{Ui} \triangleq max\{Y_{Ui} - \hat{Y}_{Ui}, Y_{Li} - \hat{Y}_{Li}\} = Y_{Ui} - \hat{Y}_{Ui} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}. \end{cases} \tag{3.11}$$

Scenario III

In this scenario, both of the predicted response lower and upper points, $\hat{y}_{Li}$ and $\hat{y}_{Ui}$ are larger than the observed response lower and upper points, y_{Li} and y_{Ui}, respectively, and the absolute value of the difference between the predicted and observed response lower points is

greater than the absolute value of the difference between the predicted and observed response upper points, which is as illustrated in Figure 3.3.

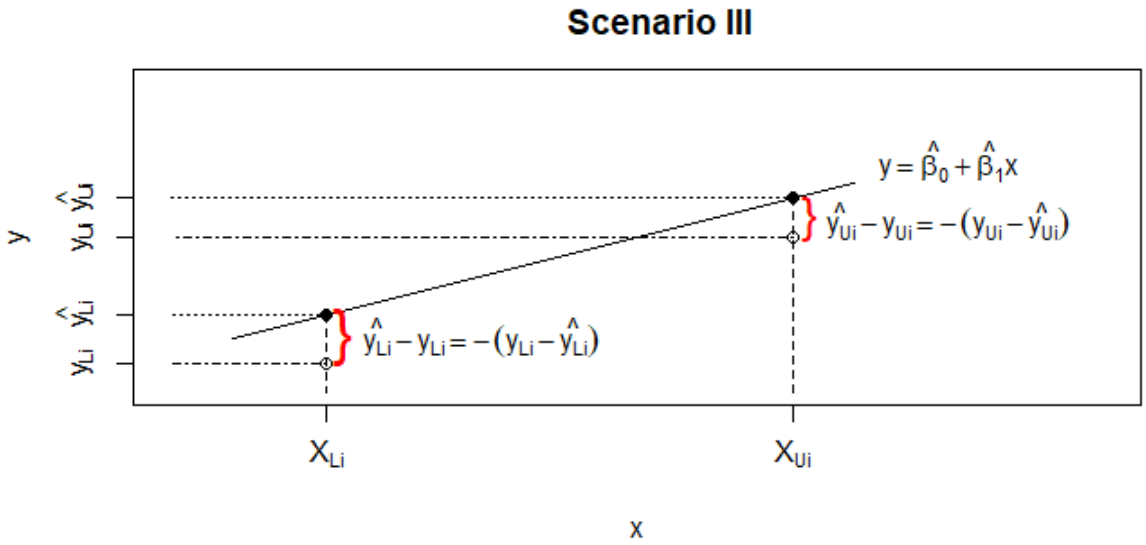

Figure 3.3: Scenario III, when $\beta_1 \geq 0$

From Figure 3.3, we have

$$Y_{Li} - \hat{Y}_{Li} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li} < Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui} = Y_{Ui} - \hat{Y}_{Ui} < 0. \tag{3.12}$$

To satisfy the basic assumption that $\epsilon_{Li} \leq \epsilon_{Ui}$, we have

$$\begin{cases} \epsilon_{Li} \triangleq min\{Y_{Ui} - \hat{Y}_{Ui}, Y_{Li} - \hat{Y}_{Li}\} = Y_{Li} - \hat{Y}_{Li} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}, \\ \epsilon_{Ui} \triangleq max\{Y_{Ui} - \hat{Y}_{Ui}, Y_{Li} - \hat{Y}_{Li}\} = Y_{Ui} - \hat{Y}_{Ui} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}. \end{cases} \tag{3.13}$$

Scenario IV

As in Scenario III, both of the predicted response lower and upper points, $\hat{y}_{Li}$ and $\hat{y}_{Ui}$ are

larger than the observed response lower and upper points, y_{Li} and y_{Ui}, respectively, but the absolute value of the difference between the predicted and observed response lower points is smaller than the absolute value of the difference between the predicted and observed response upper points, which is as illustrated in Figure 3.4.

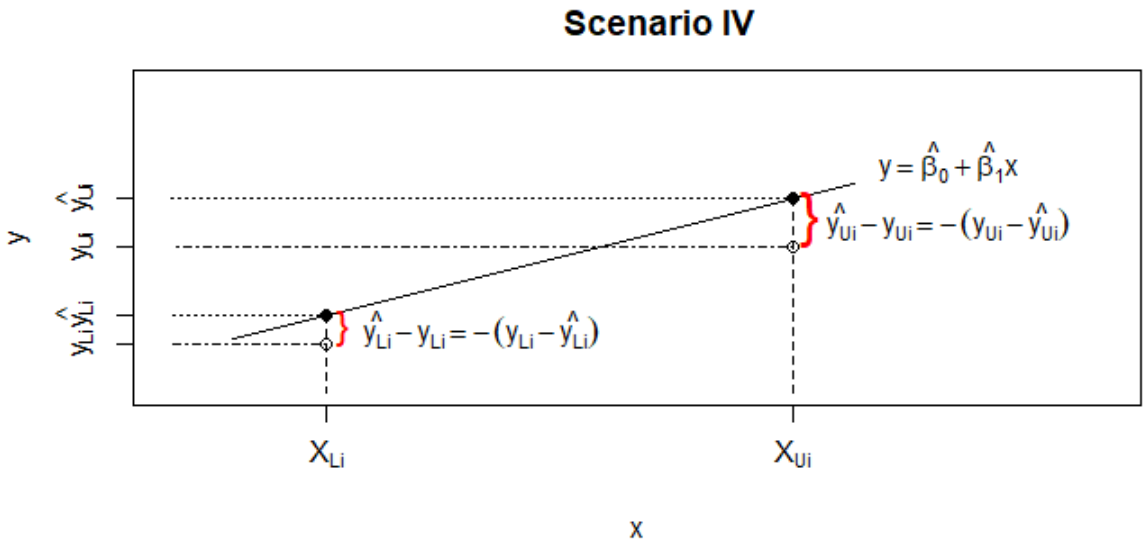

Figure 3.4: Scenario IV, when $\beta_1 \geq 0$

From Figure 3.4 above, we have

$$Y_{Ui} - \hat{Y}_{Ui} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui} < Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li} = Y_{Li} - \hat{Y}_{Li} < 0. \tag{3.14}$$

To satisfy the basic assumption that $\epsilon_{Li} \leq \epsilon_{Ui}$, we need to define

$$\begin{cases} \epsilon_{Li} \triangleq min\{Y_{Ui} - \hat{Y}_{Ui}, Y_{Li} - \hat{Y}_{Li}\} = Y_{Ui} - \hat{Y}_{Ui} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}, \\ \\ \epsilon_{Ui} \triangleq max\{Y_{Ui} - \hat{Y}_{Ui}, Y_{Li} - \hat{Y}_{Li}\} = Y_{Li} - \hat{Y}_{Li} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}. \end{cases} \tag{3.15}$$

Scenario V

In this scenario, both of the predicted response lower and upper points, $\hat{y}_{Li}$ and $\hat{y}_{Ui}$, are smaller than the observed response lower and upper points, y_{Li} and y_{Ui}, respectively, and the absolute value of the difference between the predicted and observed response lower points is smaller than the absolute value of the difference between the predicted and observed response upper points, which is as illustrated in Figure 3.5.

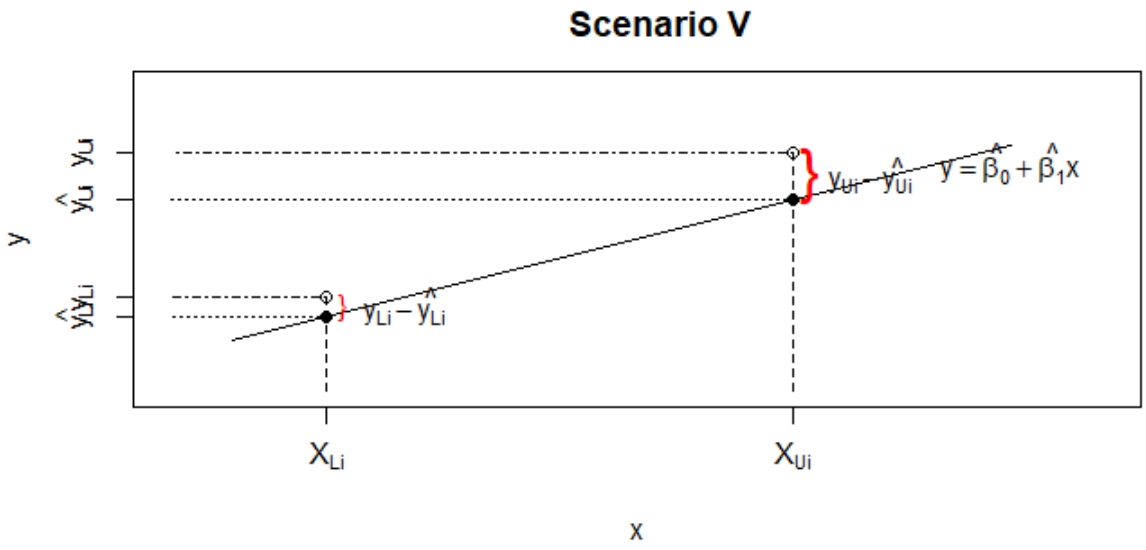

Figure 3.5: Scenario V, when $\beta_1 \geq 0$

From Figure 3.5 above, we have

$$0 < Y_{Li} - \hat{Y}_{Li} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li} < Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui} = Y_{Ui} - \hat{Y}_{Ui}. \tag{3.16}$$

To satisfy $\epsilon_{Li} \leq \epsilon_{Ui}$, we set

$$\begin{cases} \epsilon_{Li} \triangleq min\{Y_{Ui} - \hat{Y}_{Ui}, Y_{Li} - \hat{Y}_{Li}\} = Y_{Li} - \hat{Y}_{Li} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}, \\ \\ \epsilon_{Ui} \triangleq max\{Y_{Ui} - \hat{Y}_{Ui}, Y_{Li} - \hat{Y}_{Li}\} = Y_{Ui} - \hat{Y}_{Ui} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}. \end{cases} \tag{3.17}$$

Scenario VI

As in Scenario V, both of the predicted response lower and upper points, $\hat{y}_{Li}$ and $\hat{y}_{Ui}$, are smaller than the observed response lower and upper points, y_{Li} and y_{Ui}, respectively, but the absolute value of the difference between the predicted and observed response lower points is larger than the absolute value of the difference between the predicted and observed response upper points, which is as illustrated in Figure 3.6.

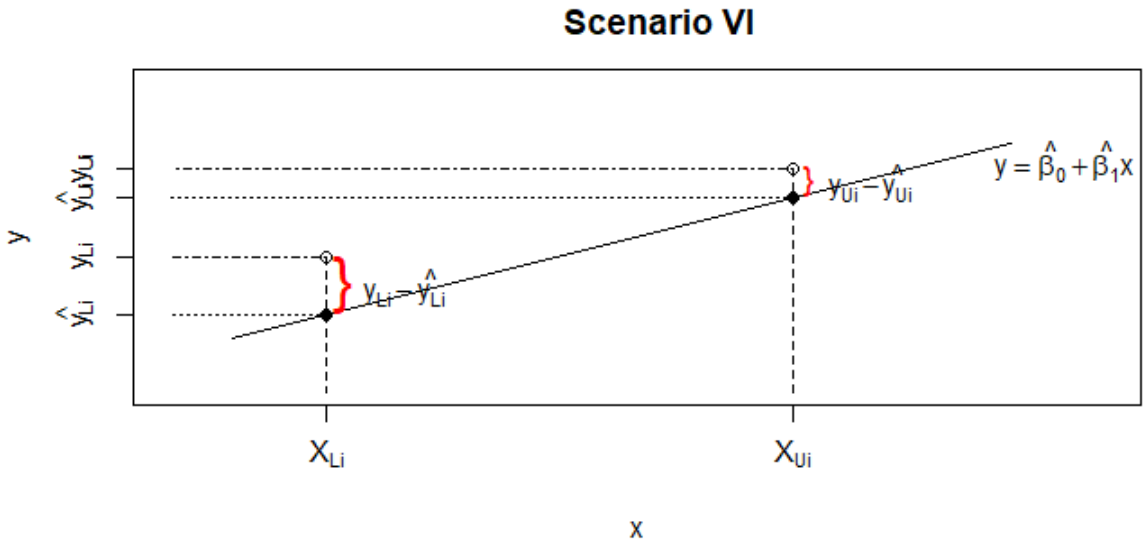

Figure 3.6: Scenario VI, when $\beta_1 \geq 0$

From Figure 3.6, we have

$$0 < Y_{Ui} - \hat{Y}_{Ui} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui} < Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li} = Y_{Li} - \hat{Y}_{Li}. \tag{3.18}$$

To satisfy $\epsilon_{Li} \leq \epsilon_{Ui}$, we have

$$\begin{cases} \epsilon_{Li} \triangleq min\{Y_{Ui} - \hat{Y}_{Ui}, Y_{Li} - \hat{Y}_{Li}\} = Y_{Ui} - \hat{Y}_{Ui} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}, \\ \\ \epsilon_{Ui} \triangleq max\{Y_{Ui} - \hat{Y}_{Ui}, Y_{Li} - \hat{Y}_{Li}\} = Y_{Li} - \hat{Y}_{Li} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}. \end{cases} \tag{3.19}$$

Unifying the results of all these 6 scenarios, we obtain the forms of the lower and upper points of the error as:

$$[\epsilon_{Li}, \epsilon_{Ui}] = [min\{Y_{Li} - X_{Li}\beta_1 - \beta_0, Y_{Ui} - X_{Ui}\beta_1 - \beta_0\}, max\{Y_{Li} - X_{Li}\beta_1 - \beta_0, Y_{Ui} - X_{Ui}\beta_1 - \beta_0\}], \tag{3.20}$$

for $i = 1, ..., n$.

(2) For $\boldsymbol{\beta_1 < 0}$

Note that since the slope is less than zero and $X_{Li} \leq X_{Ui}$, by (3.7), we have

$$\hat{Y}_{Ui} = \hat{\beta}_0 + \hat{\beta}_1 X_{Ui} \leq \hat{\beta}_0 + \hat{\beta}_1 X_{Li} = \hat{Y}_{Li}, \tag{3.21}$$

for $i = 1, ..., n$. Therefore, $\hat{Y}_{Li}$ is the predicted upper point, while $\hat{Y}_{Ui}$ is the predicted lower point of the ith response value by the linear regression model, for $i = 1, ..., n$.

Scenario I

In this scenario, the observed response interval (y_{Li}, y_{Ui}) is contained within the predicted

interval $(\hat{y}_{Li}, \hat{y}_{Ui})$ as illustrated in Figure 3.7.

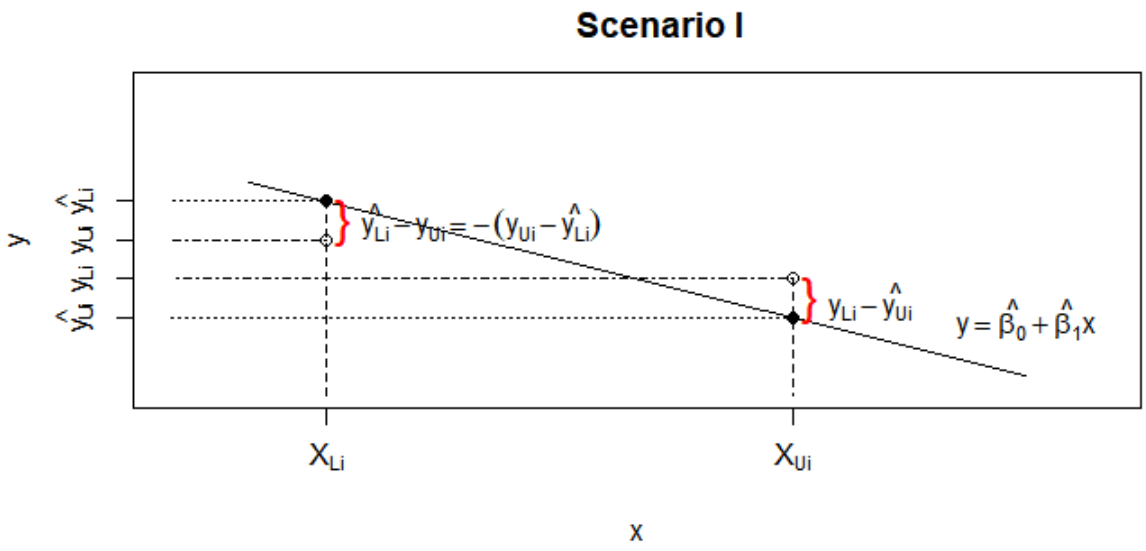

Figure 3.7: Scenario I, when $\beta_1 < 0$

From Figure 3.7, we have

$$Y_{Ui} - \hat{Y}_{Li} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li} < 0 < Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui} = Y_{Li} - \hat{Y}_{Ui}. \tag{3.22}$$

To satisfy the basic assumption that $\epsilon_{Li} \leq \epsilon_{Ui}$, then

$$\begin{cases} \epsilon_{Li} \triangleq min\{Y_{Ui} - \hat{Y}_{Li}, Y_{Li} - \hat{Y}_{Ui}\} = Y_{Ui} - \hat{Y}_{Li} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}, \\ \\ \epsilon_{Ui} \triangleq max\{Y_{Ui} - \hat{Y}_{Li}, Y_{Li} - \hat{Y}_{Ui}\} = Y_{Li} - \hat{Y}_{Ui} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}. \end{cases} \tag{3.23}$$

Scenario II

In this scenario, the predicted response interval $(\hat{y}_{Li}, \hat{y}_{Ui})$ is contained within the observed interval (y_{Li}, y_{Ui}) as illustrated in Figure 3.8.

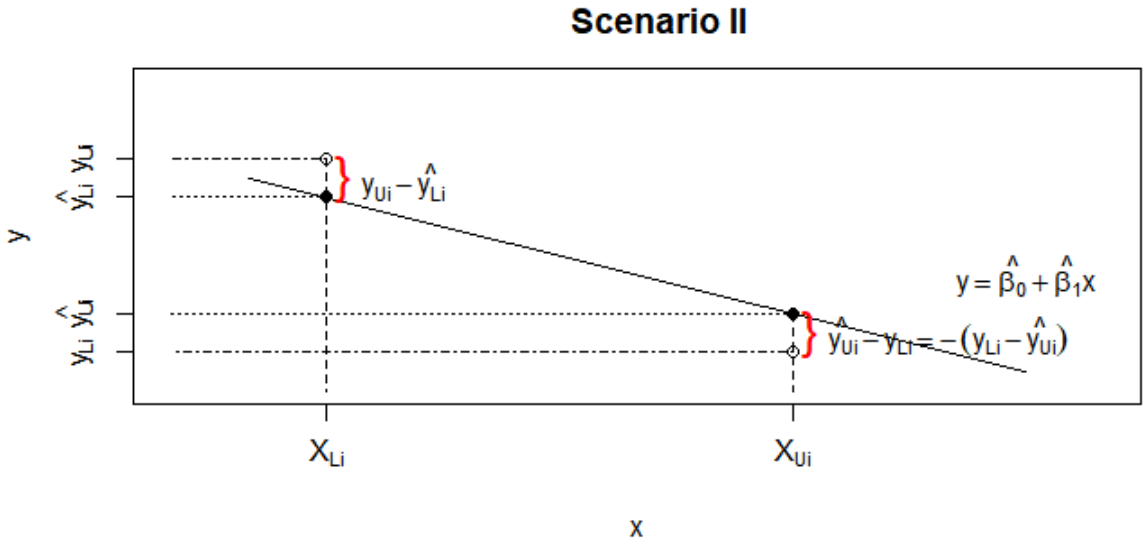

Figure 3.8: Scenario II, when $\beta_1 < 0$

From Figure 3.8, we have

$$Y_{Li} - \hat{Y}_{Ui} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui} < 0 < Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li} = Y_{Ui} - \hat{Y}_{Li}. \tag{3.24}$$

To satisfy the basic assumption that $\epsilon_{Li} \leq \epsilon_{Ui}$, then

$$\begin{cases} \epsilon_{Li} \triangleq min\{Y_{Ui} - \hat{Y}_{Li}, Y_{Li} - \hat{Y}_{Ui}\} = Y_{Li} - \hat{Y}_{Ui} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}, \\ \\ \epsilon_{Ui} \triangleq max\{Y_{Ui} - \hat{Y}_{Li}, Y_{Li} - \hat{Y}_{Ui}\} = Y_{Ui} - \hat{Y}_{Li} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}. \end{cases} \tag{3.25}$$

Scenario III

In this scenario, both of the predicted response lower and upper points, $\hat{y}_{Li}$ and $\hat{y}_{Ui}$, are larger than the observed response lower and upper points, y_{Li} and y_{Ui}, respectively, and the absolute value of the difference between the predicted and observed response lower points is

greater than the absolute value of the difference between the predicted and observed response upper points, which is as illustrated in Figure 3.9.

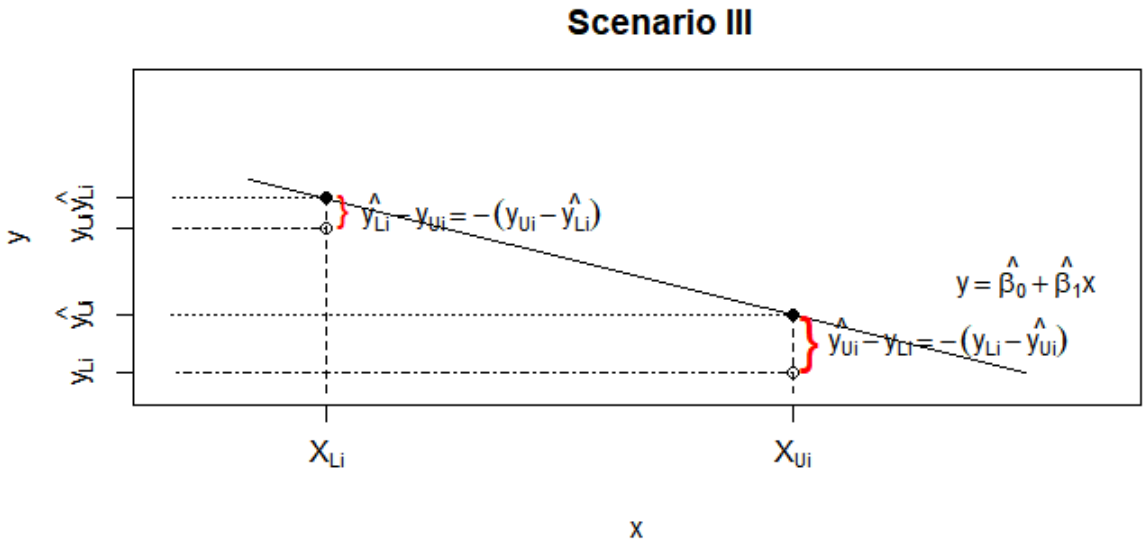

Figure 3.9: Scenario III, when $\beta_1 < 0$

From Figure 3.9, we have

$$Y_{Li} - \hat{Y}_{Ui} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui} < Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li} = Y_{Ui} - \hat{Y}_{Li} < 0. \tag{3.26}$$

To satisfy $\epsilon_{Li} \leq \epsilon_{Ui}$, then we have

$$\begin{cases} \epsilon_{Li} \triangleq min\{Y_{Ui} - \hat{Y}_{Li}, Y_{Li} - \hat{Y}_{Ui}\} = Y_{Li} - \hat{Y}_{Ui} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}, \\ \epsilon_{Ui} \triangleq max\{Y_{Ui} - \hat{Y}_{Li}, Y_{Li} - \hat{Y}_{Ui}\} = Y_{Ui} - \hat{Y}_{Li} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}. \end{cases} \tag{3.27}$$

Scenario IV

As in Scenario III, both of the predicted response lower and upper points, $\hat{y}_{Li}$ and $\hat{y}_{Ui}$, are

larger than the observed response lower and upper points, y_{Li} and y_{Ui}, respectively, but the absolute value of the difference between the predicted and observed response lower points is smaller than the absolute value of the difference between the predicted and observed response upper points, which is as illustrated in Figure 3.10.

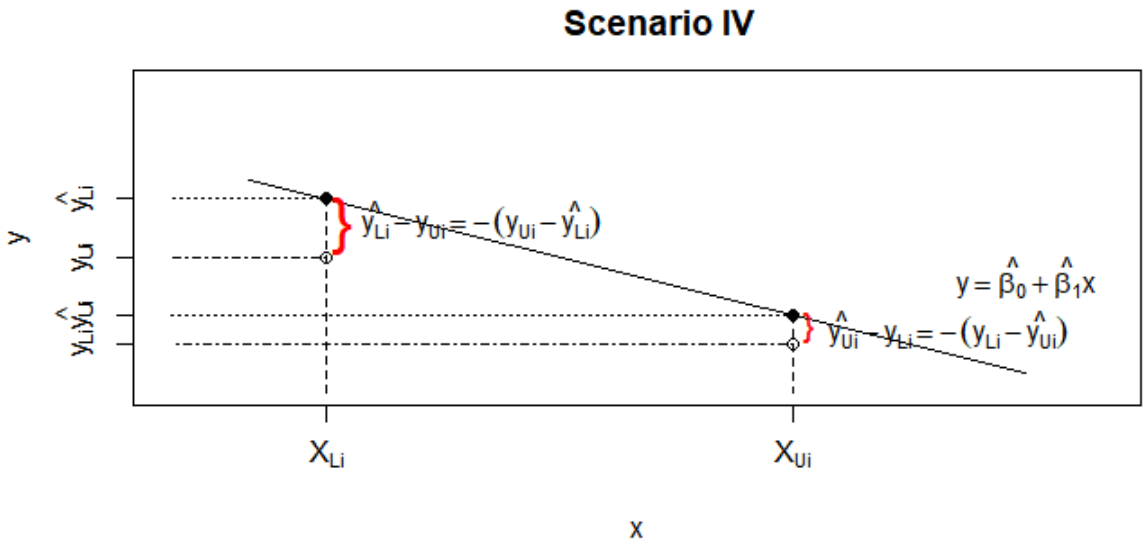

Figure 3.10: Scenario IV, when $\beta_1 < 0$

From Figure 3.10, we have

$$Y_{Ui} - \hat{Y}_{Li} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li} < Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui} = Y_{Li} - \hat{Y}_{Ui} < 0. \tag{3.28}$$

To satisfy $\epsilon_{Li} \leq \epsilon_{Ui}$, then we need

$$\begin{cases} \epsilon_{Li} \triangleq min\{Y_{Ui} - \hat{Y}_{Li}, Y_{Li} - \hat{Y}_{Ui}\} = Y_{Ui} - \hat{Y}_{Li} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}, \\ \\ \epsilon_{Ui} \triangleq max\{Y_{Ui} - \hat{Y}_{Li}, Y_{Li} - \hat{Y}_{Ui}\} = Y_{Li} - \hat{Y}_{Ui} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}. \end{cases} \tag{3.29}$$

Scenario V

In this scenario, both of the predicted response lower and upper points, $\hat{y}_{Li}$, and $\hat{y}_{Ui}$ are smaller than the observed response lower and upper points, y_{Li} and y_{Ui}, respectively, and the absolute value of the difference between the predicted and observed response lower points is smaller than the absolute value of the difference between the predicted and observed response upper points, which is as illustrated in Figure 3.11.

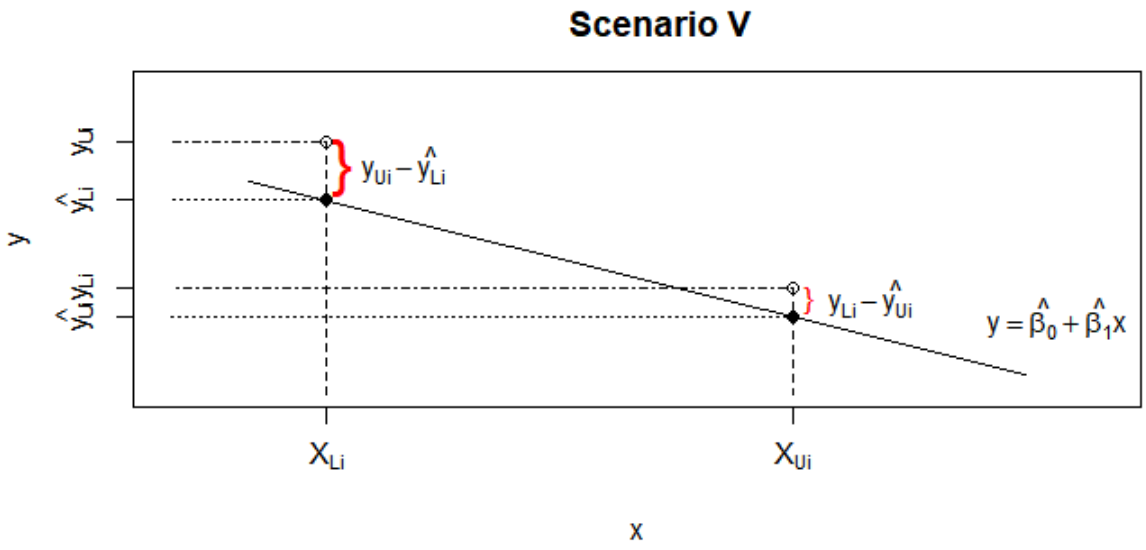

Figure 3.11: Scenario V, when $\beta_1 < 0$

From Figure 3.11, we have

$$0 < Y_{Li} - \hat{Y}_{Ui} < Y_{Ui} - \hat{Y}_{Li}. \tag{3.30}$$

To satisfy $\epsilon_{Li} \leq \epsilon_{Ui}$, then we define

$$\begin{cases} \epsilon_{Li} \triangleq min\{Y_{Ui} - \hat{Y}_{Li}, Y_{Li} - \hat{Y}_{Ui}\} = Y_{Li} - \hat{Y}_{Ui} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}, \\ \\ \epsilon_{Ui} \triangleq max\{Y_{Ui} - \hat{Y}_{Li}, Y_{Li} - \hat{Y}_{Ui}\} = Y_{Ui} - \hat{Y}_{Li} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}. \end{cases} \tag{3.31}$$

Scenario VI

As in the Scenario V, both of the predicted response lower and upper points, $\hat{y}_{Li}$ and $\hat{y}_{Ui}$, are smaller than the observed response lower and upper points, y_{Li} and y_{Ui}, respectively, but the absolute value of the difference between the predicted and observed response lower points is larger than the absolute value of the difference between the predicted and observed response upper points, which is as illustrated in Figure 3.12.

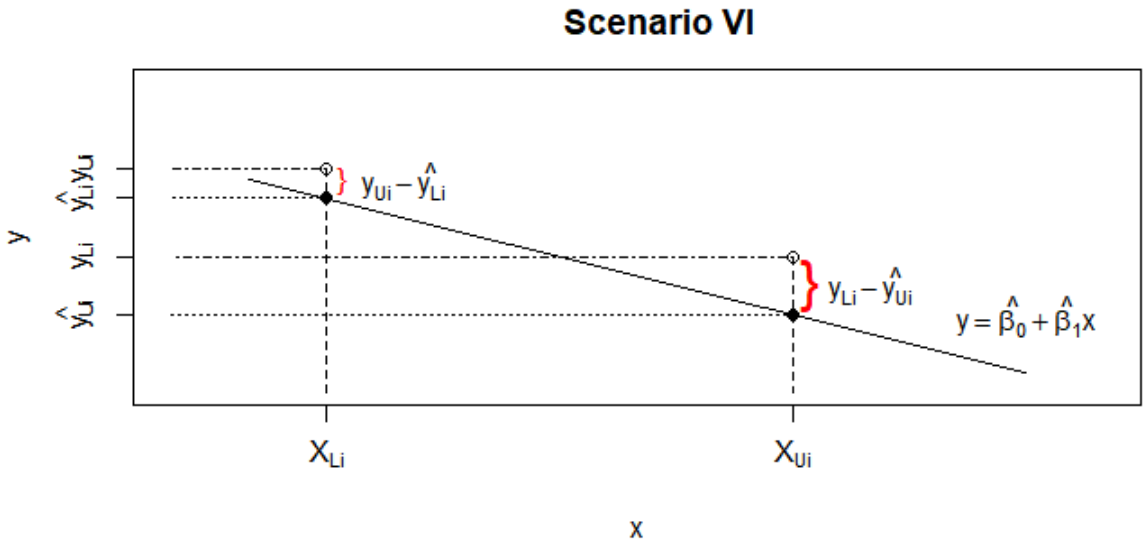

Figure 3.12: Scenario VI, when $\beta_1 < 0$

From Figure 3.12, we have

$$0 < Y_{Ui} - \hat{Y}_{Li} < Y_{Li} - \hat{Y}_{Ui}. \tag{3.32}$$

To satisfy $\epsilon_{Li} \leq \epsilon_{Ui}$, then we have

$$\begin{cases} \epsilon_{Li} \triangleq min\{Y_{Ui} - \hat{Y}_{Li}, Y_{Li} - \hat{Y}_{Ui}\} = Y_{Ui} - \hat{Y}_{Li} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}, \\ \\ \epsilon_{Ui} \triangleq max\{Y_{Ui} - \hat{Y}_{Li}, Y_{Li} - \hat{Y}_{Ui}\} = Y_{Li} - \hat{Y}_{Ui} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}. \end{cases} \tag{3.33}$$

Combining the results of all these 6 scenarios, we have:

$$[\epsilon_{Li}, \epsilon_{Ui}] = [min\{Y_{Li} - X_{Ui}\beta_1 - \beta_0, Y_{Ui} - X_{Li}\beta_1 - \beta_0\}, max\{Y_{Li} - X_{Ui}\beta_1 - \beta_0, Y_{Ui} - X_{Li}\beta_1 - \beta_0\}], \tag{3.34}$$

for $i = 1, ..., n$.

Next, we generate the likelihood function of the errors for the linear regression model (3.5), based on the order statistic assumption.

First we illustrate the joint distribution of $(\epsilon_{Li}, \epsilon_{Ui})$, for $i = 1, ..., n$. Based on the assumption (3.6), the probability density function of ϵ_i is

$$f_{\epsilon_i}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, -\infty < x < \infty, \ i = 1, ..., n. \tag{3.35}$$

We use the theorem on the joint probability density of order statistics of a random sample (Casella and Berger, 2002 [50], Theorem 5.4.6):

Let $X_{(1)}, ..., X_{(n)}$ denote the order statistics of a random sample, $X_1, ..., X_n$, from a continuous population with cumulative density function $F_X(x)$ and probability density function

$f_X(x)$. *Then, the joint probability density function of* $X_{(i)}$ *and* $X_{(j)}$, $1 \leq i < j \leq n$, *is*

$$f_{X_{(i)},X_{(j)}}(u,v) = \frac{n!}{(i-1)!(j-1-i)!(n-j)!} f_X(u) f_X(v) [F_X(u)]^{i-1} \times [F_X(v) - F_X(u)]^{j-1-i} [1 - F_X(v)]^{n-j}, \ -\infty < u < v < \infty. \tag{3.36}$$

In our case, we have $n = 2$, and we are interested in the order statistics $(i) = (1)$ and $(j) = (n)$. Thus, we can derive the joint distribution of $(\epsilon_{Li}, \epsilon_{Ui}), i = 1, ..., n$, with probability density function to be

$$\begin{aligned} g(\epsilon_{Li}, \epsilon_{Ui}) &= \frac{2!}{1!0!0!} f_{\epsilon_i}(\epsilon_{Li}) f_{\epsilon_i}(\epsilon_{Ui}) \\ &= \frac{1}{\pi\sigma^2} e^{-\frac{1}{2\sigma^2}(\epsilon_{Li}^2 + \epsilon_{Ui}^2)}, -\infty < \epsilon_{Li} < \epsilon_{Ui} < \infty. \end{aligned} \tag{3.37}$$

Assuming the errors $[\epsilon_{Li}, \epsilon_{Ui}]$ are independent across observations, $i = 1, ..., n$, by (3.35), we can show that the likelihood function of the errors is as follows:

$$\begin{aligned} L(\epsilon_{L1}, \epsilon_{U1}, ..., \epsilon_{Ln}, \epsilon_{Un}) &= \prod_{i=1}^{n} g(\epsilon_{Li}, \epsilon_{Ui}) \\ &= (\pi\sigma^2)^{-n} exp(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (\epsilon_{Li}^2 + \epsilon_{Ui}^2)) \prod_{i=1}^{n} I(\epsilon_{Li} \leq \epsilon_{Ui}) \end{aligned} \tag{3.38}$$

where $g(\epsilon_{Li}, \epsilon_{Ui})$ is obtained in (3.37).

Assumption II: Independence

For the second assumption, suppose ϵ_{Li} and ϵ_{Ui} are independent for $i = 1, ..., n$, and are normally distributed with mean zero and constant variance, i.e.,

$$\epsilon_{Li}, \epsilon_{Ui} \overset{iid}{\sim} N(0, \sigma^2), i = 1, ..., n. \tag{3.39}$$

Compared with the first assumption, independence between the error for the lower bound of response and the error for the upper bound of response is the only difference. This assumption is appropriate from the point of view that corresponding to the error term, both of the residuals by the lower and the upper bounds of response are caused by a random component that failed to be explained by the linear combination of explanatory variables. Therefore, the residuals are random, and can be considered as independent both within and across observations.

By (3.5), we have the universal conditions: $Y_{Li} \leq Y_{Ui}$ and $X_{Li} \leq X_{Ui}$. The forms of Y_{Li} and Y_{Ui} by the linear regression model vary depending on the sign of the slope parameter β_1. For $\beta_1 \geq 0$,

$$Y_{Li} = \beta_0 + \beta_1 X_{Li} + \epsilon_{Li},\ Y_{Ui} = \beta_0 + \beta_1 X_{Ui} + \epsilon_{Ui}, \tag{3.40}$$

while for $\beta_1 < 0$,

$$Y_{Li} = \beta_0 + \beta_1 X_{Ui} + \epsilon_{Li},\ Y_{Ui} = \beta_0 + \beta_1 X_{Li} + \epsilon_{Ui} \tag{3.41}$$

for $i = 1, ..., n$, where ϵ_{Li} and ϵ_{Ui} represent the errors for the lower bound and upper bound of the response, respectively.

When $\beta_1 \geq 0$, we denote $\hat{Y}_{Li}$ and $\hat{Y}_{Ui}$ to be the predicted lower and upper points of the ith response value by the linear regression, and r_{Li} and r_{Ui} to be the lower and upper residuals for the ith response, $i = 1, ..., n$. With the assumption given by (3.39), and by (3.7), we have

$$\hat{Y}_{Li} = \hat{\beta}_0 + \hat{\beta}_1 X_{Li},\ r_{Li} = Y_{Li} - \hat{Y}_{Li} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}, \tag{3.42}$$

$$\hat{Y}_{Ui} = \hat{\beta}_0 + \hat{\beta}_1 X_{Ui},\ r_{Ui} = Y_{Ui} - \hat{Y}_{Ui} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}. \tag{3.43}$$

When $\beta_1 < 0$, by (3.7), since $\hat{Y}_{Li} = \hat{\beta}_0 + \hat{\beta}_1 X_{Li} > \hat{\beta}_0 + \hat{\beta}_1 X_{Ui} = \hat{Y}_{Ui}$, we denote the predicted upper point of the ith response value to be $\hat{Y}_{Li}$, and the predicted lower point of the ith response value to be $\hat{Y}_{Ui}$. Therefore, letting r_{Li} and r_{Ui} denote the lower and upper

residuals for the ith response, we have

$$r_{Li} = Y_{Li} - \hat{Y}_{Ui} = Y_{Li} - \hat{\beta}_0 - \hat{\beta}_1 X_{Ui}, \tag{3.44}$$

$$r_{Ui} = Y_{Ui} - \hat{Y}_{Li} = Y_{Ui} - \hat{\beta}_0 - \hat{\beta}_1 X_{Li}, \tag{3.45}$$

for $i = 1, ..., n$.

Figure 3.13 displays the relations between the interval predicted response value $[\hat{Y}_{Li}, \hat{Y}_{Ui}]$ and the interval realization of the explanatory variable $[X_{Li}, X_{Ui}]$ by different signs of the slope parameter.

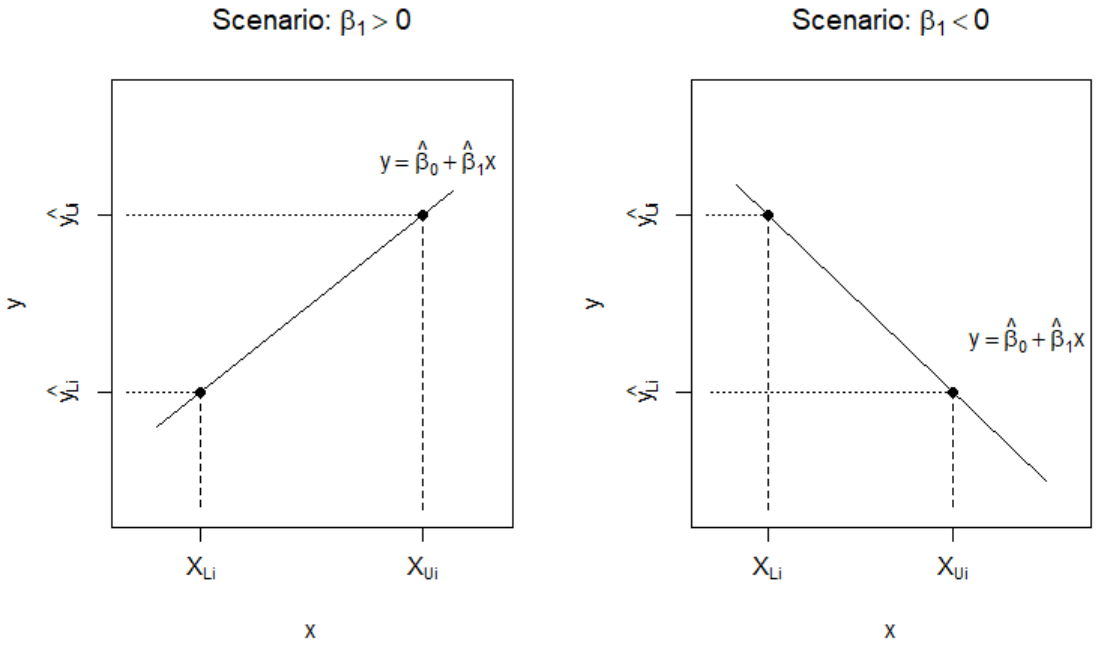

Figure 3.13: Relations between $[\hat{Y}_{Li}, \hat{Y}_{Ui}]$ and $[X_{Li}, X_{Ui}]$

Based on (3.39), we generate the likelihood function of the random error for the linear regression model (3.5) under the second assumption.

By (3.35), the joint distribution of $(\epsilon_{Li}, \epsilon_{Ui}), i = 1, ..., n$, with probability density function

is

$$\begin{aligned} g(\epsilon_{Li}, \epsilon_{Ui}) &= f_{\epsilon_i}(\epsilon_{Li}) f_{\epsilon_i}(\epsilon_{Ui}) \\ &= \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}(\epsilon_{Li}^2 + \epsilon_{Ui}^2)}, -\infty < \epsilon_{Li}, \epsilon_{Ui} < \infty. \end{aligned} \tag{3.46}$$

Since $[\epsilon_{Li}, \epsilon_{Ui}]$ are independent across observations, $i = 1, ..., n$, the likelihood function of the random error is:

$$\begin{aligned} L(\epsilon_{L1}, \epsilon_{U1}, ..., \epsilon_{Ln}, \epsilon_{Un}) &= \prod_{i=1}^{n} g(\epsilon_{Li}, \epsilon_{Ui}) \\ &= (2\pi\sigma^2)^{-n} exp(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (\epsilon_{Li}^2 + \epsilon_{Ui}^2)) \end{aligned} \tag{3.47}$$

where $g(\epsilon_{Li}, \epsilon_{Ui})$ is obtained in (3.46).

Point Estimation for Regression Coefficients

In this section, we first generate the point estimators for the regression coefficients by means of the independence assumption and the maximum likelihood principle, and then study their properties. There are two reasons why we choose the second assumption on the error term. The first is that the independence between the error for the lower bound of response and the error for the upper bound of response is more appropriate in terms of the establishments and interpretations of linear regression models, especially for interval-valued data sets obtained by data aggregation. The second is that by choosing the first assumption, we have to consider the non-negative correlation between the lower and upper residuals within each observation, which is complicated, in order to do statistical inference on the linear regression model. As we are at the initial stage of the methodology development, it is more feasible to choose the second assumption. Further development of the first option will be deferred for future research.

Maximum Likelihood Estimators of $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$

Let us consider obtaining the maximum likelihood estimator (MLE) of the parameters β_0 and β_1. By (3.40) and (3.41),

$$\epsilon_{Li} = Y_{Li} - \beta_0 - \beta_1 X_{Li},\ \epsilon_{Ui} = Y_{Ui} - \beta_0 - \beta_1 X_{Ui},\ \text{when } \beta_1 \geq 0, \tag{3.48}$$

$$\epsilon_{Li} = Y_{Li} - \beta_0 - \beta_1 X_{Ui},\ \epsilon_{Ui} = Y_{Ui} - \beta_0 - \beta_1 X_{Li},\ \text{when } \beta_1 < 0, \tag{3.49}$$

for $i = 1, ..., n$. Replacing ϵ_{Li} and ϵ_{Ui} in (3.46) by (3.48) or (3.49), depending on whether β_1 is positive or negative, we can express the likelihood function of the random error as a function of the intercept and the slope parameters. We consider each case in turn.

(1) For $\boldsymbol{\beta}_1 \geq 0$

When the slope parameter $\beta_1 \geq 0$, we have, from (3.47) and (3.48),

$$L(\epsilon_{L1}, \epsilon_{U1}, ..., \epsilon_{Ln}, \epsilon_{Un}) = (\pi\sigma^2)^{-n} exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[(Y_{Li} - X_{Li}\beta_1 - \beta_0)^2 + (Y_{Ui} - X_{Ui}\beta_1 - \beta_0)^2]\}. \tag{3.50}$$

Then, the log-likelihood function is

$$\begin{aligned} l = logL(\epsilon_{L1}, \epsilon_{U1}, ..., \epsilon_{Ln}, \epsilon_{Un}) &= -nlog\pi - nlog\sigma^2 \\ &- \frac{1}{2\sigma^2}\sum_{i=1}^{n}[(Y_{Li} - X_{Li}\beta_1 - \beta_0)^2 + (Y_{Ui} - X_{Ui}\beta_1 - \beta_0)^2]. \end{aligned}$$

Taking the first derivative of l with respect to β_1 and β_0 gives

$$\frac{\partial l}{\partial \beta_1} = \frac{1}{\sigma^2}\sum_{i=1}^{n}[X_{Li}Y_{Li} + X_{Ui}Y_{Ui} - \beta_1(X_{Li}^2 + X_{Ui}^2) - \beta_0(X_{Li} + X_{Ui})], \tag{3.51}$$

$$\frac{\partial l}{\partial \beta_0} = \frac{1}{\sigma^2}[\sum_{i=1}^{n}(Y_{Li} + Y_{Ui}) - \sum_{i=1}^{n}(X_{Li} + X_{Ui})\beta_1 - 2n\beta_0]. \tag{3.52}$$

Then setting the derivatives evaluated at $\underline{\beta} = \hat{\underline{\beta}}$ to be 0, respectively, i.e., $\frac{\partial l}{\partial \beta_1}|_{\underline{\beta}=\hat{\underline{\beta}}} = 0$ and $\frac{\partial l}{\partial \beta_0}|_{\underline{\beta}=\hat{\underline{\beta}}} = 0$, we have:

$$\begin{cases} \frac{\partial l}{\partial \beta_1}|_{\underline{\beta}=\hat{\underline{\beta}}} = \frac{1}{\sigma^2}\sum_{i=1}^{n}[X_{Li}Y_{Li} + X_{Ui}Y_{Ui} - \beta_1(X_{Li}^2 + X_{Ui}^2) - \beta_0(X_{Li} + X_{Ui})]|_{\underline{\beta}=\hat{\underline{\beta}}} \stackrel{\triangle}{=} 0, \\ \\ \frac{\partial l}{\partial \beta_0}|_{\underline{\beta}=\hat{\underline{\beta}}} = \frac{1}{\sigma^2}[\sum_{i=1}^{n}(Y_{Li} + Y_{Ui}) - \sum_{i=1}^{n}(X_{Li} + X_{Ui})\beta_1 - 2n\beta_0]|_{\underline{\beta}=\hat{\underline{\beta}}} \stackrel{\triangle}{=} 0. \end{cases} \tag{3.53}$$

Solving (3.53) for $\hat{\beta}_1$ and $\hat{\beta}_0$, we obtain

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_{Li} + X_{Ui})\sum_{i=1}^{n}(Y_{Li} + Y_{Ui}) - 2n\sum_{i=1}^{n}(X_{Li}Y_{Li} + X_{Ui}Y_{Ui})}{[\sum_{i=1}^{n}(X_{Li} + X_{Ui})]^2 - 2n\sum_{i=1}^{n}(X_{Li}^2 + X_{Ui}^2)}, \tag{3.54}$$

$$\begin{aligned} \hat{\beta}_0 = & \frac{1}{2n}\sum_{i=1}^{n}(Y_{Li} + Y_{Ui}) \\ & - \frac{1}{2n}\frac{[\sum_{i=1}^{n}(X_{Li} + X_{Ui})]^2\sum_{i=1}^{n}(Y_{Li} + Y_{Ui}) - 2n\sum_{i=1}^{n}(X_{Li}Y_{Li} + X_{Ui}Y_{Ui})\sum_{i=1}^{n}(X_{Li} + X_{Ui})}{[\sum_{i=1}^{n}(X_{Li} + X_{Ui})]^2 - 2n\sum_{i=1}^{n}(X_{Li}^2 + X_{Ui}^2)}. \end{aligned} \tag{3.55}$$

Since $\frac{\partial^2 l}{\partial \beta_1^2}|_{\beta_1=\hat{\beta}_1} < 0$, and $\frac{\partial^2 l}{\partial \beta_0^2}|_{\beta_0=\hat{\beta}_0} < 0$, the estimators $\hat{\beta}_1$ and $\hat{\beta}_0$ are indeed MLE estimators.

(2) For $\boldsymbol{\beta_1 < 0}$

When the slope parameter $\beta_1 < 0$, we proceed as we did for $\beta_1 > 0$. By (3.47) and (3.49),

$$L(\epsilon_{L1}, \epsilon_{U1}, ..., \epsilon_{Ln}, \epsilon_{Un}) = (\pi\sigma^2)^{-n}exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[(Y_{Li} - X_{Ui}\beta_1 - \beta_0)^2 + (Y_{Ui} - X_{Li}\beta_1 - \beta_0)^2]\}, \tag{3.56}$$

and the log-likelihood function is

$$\begin{aligned} l = logL(\epsilon_{L1}, \epsilon_{U1}, ..., \epsilon_{Ln}, \epsilon_{Un}) &= -nlog\pi - nlog\sigma^2 \\ &- \frac{1}{2\sigma^2}\sum_{i=1}^{n}[(Y_{Li} - X_{Ui}\beta_1 - \beta_0)^2 + (Y_{Ui} - X_{Li}\beta_1 - \beta_0)^2]. \end{aligned}$$

Setting $\frac{\partial l}{\partial \beta_1}|_{\underline{\beta}=\hat{\underline{\beta}}} \stackrel{\triangle}{=} 0$, we have

$$\hat{\beta}_1 \sum_{i=1}^{n}(X_{Li}^2 + X_{Ui}^2) + \hat{\beta}_0 \sum_{i=1}^{n}(X_{Li} + X_{Ui}) = \sum_{i=1}^{n}(X_{Ui}Y_{Li} + X_{Li}Y_{Ui}); \tag{3.57}$$

and setting $\frac{\partial l}{\partial \beta_0}|_{\underline{\beta}=\hat{\underline{\beta}}} \stackrel{\triangle}{=} 0$, we have

$$\hat{\beta}_1 \sum_{i=1}^{n}(X_{Li} + X_{Ui}) + 2n\hat{\beta}_0 = \sum_{i=1}^{n}(Y_{Li} + Y_{Ui}). \tag{3.58}$$

Similar to the case when $\beta_1 \geq 0$, by (3.57) and (3.58), we obtain the MLEs of β_1 and β_0 as

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_{Li} + X_{Ui}) \sum_{i=1}^{n}(Y_{Li} + Y_{Ui}) - 2n \sum_{i=1}^{n}(X_{Li}Y_{Ui} + X_{Ui}Y_{Li})}{[\sum_{i=1}^{n}(X_{Li} + X_{Ui})]^2 - 2n \sum_{i=1}^{n}(X_{Li}^2 + X_{Ui}^2)}, \tag{3.59}$$

$$\begin{aligned}\hat{\beta}_0 = {} & \frac{1}{2n} \sum_{i=1}^{n}(Y_{Li} + Y_{Ui}) \\ & - \frac{1}{2n} \frac{[\sum_{i=1}^{n}(X_{Li} + X_{Ui})]^2 \sum_{i=1}^{n}(Y_{Li} + Y_{Ui}) - 2n \sum_{i=1}^{n}(X_{Li}Y_{Ui} + X_{Ui}Y_{Li}) \sum_{i=1}^{n}(X_{Li} + X_{Ui})}{[\sum_{i=1}^{n}(X_{Li} + X_{Ui})]^2 - 2n \sum_{i=1}^{n}(X_{Li}^2 + X_{Ui}^2)}.\end{aligned} \tag{3.60}$$

Properties of $\hat{\boldsymbol{\beta}}_0$ and $\hat{\boldsymbol{\beta}}_1$

Next, we study properties of the point estimators. First, we prove that $\hat{\beta}_1$ and $\hat{\beta}_0$ are unbiased estimators. For $\beta_1 \geq 0$, from (3.40), we know

$$E(Y_{Li}) = X_{Li}\beta_1 + \beta_0, E(Y_{Ui}) = X_{Ui}\beta_1 + \beta_0, \ i = 1, ..., n. \tag{3.61}$$

For $\beta_1 \geq 0$, by (3.54), we have

$$
\begin{aligned}
E(\hat{\beta}_1) &= \frac{[\sum_{i=1}^n (X_{Li}+X_{Ui})]^2\beta_1 + 2n\sum_{i=1}^n (X_{Li}+X_{Ui})\beta_0}{[\sum_{i=1}^n (X_{Li}+X_{Ui})]^2 - 2n\sum_{i=1}^n (X_{Li}^2+X_{Ui}^2)} \\
&- \frac{2n\sum_{i=1}^n (X_{Li}^2+X_{Ui}^2)\beta_1 + 2n\sum_{i=1}^n (X_{Li}+X_{Ui})\beta_0}{[\sum_{i=1}^n (X_{Li}+X_{Ui})]^2 - 2n\sum_{i=1}^n (X_{Li}^2+X_{Ui}^2)} \\
&= \frac{\{[\sum_{i=1}^n (X_{Li}+X_{Ui})]^2 - 2n\sum_{i=1}^n (X_{Li}^2+X_{Ui}^2)\}\beta_1}{[\sum_{i=1}^n (X_{Li}+X_{Ui})]^2 - 2n\sum_{i=1}^n (X_{Li}^2+X_{Ui}^2)} \\
&= \beta_1. \qquad (3.62)
\end{aligned}
$$

Since, by (3.55),

$$\hat{\beta}_0 = \frac{1}{2n}[\sum_{i=1}^n (Y_{Li}+Y_{Ui}) - \hat{\beta}_1 \sum_{i=1}^n (X_{Li}+X_{Ui})], \qquad (3.63)$$

we have

$$E(\hat{\beta}_0) = \frac{1}{2n}[\sum_{i=1}^n (X_{Li}\beta_1 + X_{Ui}\beta_1 + 2\beta_0) - \beta_1 \sum_{i=1}^n (X_{Li}+X_{Ui})] = \beta_0. \qquad (3.64)$$

Therefore, $\hat{\beta}_1$ and $\hat{\beta}_0$ are unbiased estimators.

Similarly, for $\beta_1 < 0$, by (3.59) and (3.60), we also have

$$E(\hat{\beta}_1) = \beta_1, \ E(\hat{\beta}_0) = \beta_0. \qquad (3.65)$$

Therefore, both $\hat{\beta}_1$ and $\hat{\beta}_0$ are unbiased.

Distributions of $\hat{\boldsymbol{\beta}}_0$ and $\hat{\boldsymbol{\beta}}_1$

Next we give the distributions of the estimators $\hat{\beta}_1$ and $\hat{\beta}_0$. For $\beta_1 \geq 0$, since by (3.54) and (3.55), both $\hat{\beta}_1$ and $\hat{\beta}_0$ are linear functions of the lower and upper points of the response. Let us first obtain the variances and covariance of Y_{Li} and Y_{Ui}, for $i = 1, ..., n$.

By (3.40), we have

$$Var(Y_{Li}) = Var(\beta_0 + \beta_1 X_{Li} + \epsilon_{Li}) = Var(\epsilon_{Li}) = \sigma^2, \tag{3.66}$$

$$Var(Y_{Ui}) = Var(\beta_0 + \beta_1 X_{Ui} + \epsilon_{Ui}) = Var(\epsilon_{Ui}) = \sigma^2, \tag{3.67}$$

for $i = 1, ..., n$.

Now since ϵ_{Li} and ϵ_{Ui} are independent, for $i = 1, ..., n$,

$$Cov(Y_{Li}, Y_{Ui}) = Cov(X_{Li}\beta_1 + \beta_0 + \epsilon_{Li}, X_{Ui}\beta_1 + \beta_0 + \epsilon_{Ui}) = Cov(\epsilon_{Li}, \epsilon_{Ui}) = 0. \tag{3.68}$$

Therefore, the distribution of $(Y_{Li}, Y_{Ui})^T$ is a bivariate normal distribution, i.e.,

$$\begin{pmatrix} Y_{Li} \\ Y_{Ui} \end{pmatrix} \sim N_2 \left(\begin{pmatrix} X_{Li}\beta_1 + \beta_0 \\ X_{Ui}\beta_1 + \beta_0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \right), \ i = 1, ..., n. \tag{3.69}$$

By (3.54), since $\hat{\beta}_1$ is a linear function of Y_{Li} and Y_{Ui} for $i = 1, ..., n$, and because of the property that a linear combination of normal random variables has a normal distribution, we have that $\hat{\beta}_1$ is a normally distributed random variable; likewise, by (3.55) $\hat{\beta}_0$ is also a normally distributed random variable.

Let us now derive the variances of $\hat{\beta}_1$ and $\hat{\beta}_0$. By (3.54) and (3.68), we have

$$\begin{aligned} Var(\hat{\beta}_1) &= \frac{Var[\sum_{i=1}^n (X_{Li} + X_{Ui}) \sum_{i=1}^n (Y_{Li} + Y_{Ui})] + 4n^2 Var[\sum_{i=1}^n (X_{Li}Y_{Li} + X_{Ui}Y_{Ui})]}{\{[\sum_{i=1}^n (X_{Li} + X_{Ui})]^2 - 2n \sum_{i=1}^n (X_{Li}^2 + X_{Ui}^2)\}^2} \\ &- \frac{2Cov[\sum_{i=1}^n (X_{Li} + X_{Ui}) \sum_{i=1}^n (Y_{Li} + Y_{Ui}), 2n \sum_{i=1}^n (X_{Li}Y_{Li} + X_{Ui}Y_{Ui})]}{\{[\sum_{i=1}^n (X_{Li} + X_{Ui})]^2 - 2n \sum_{i=1}^n (X_{Li}^2 + X_{Ui}^2)\}^2} \\ &= \frac{-2n\sigma^2[\sum_{i=1}^n (X_{Li} + X_{Ui})]^2 + 4n^2\sigma^2 \sum_{i=1}^n (X_{Li}^2 + X_{Ui}^2)}{\{[\sum_{i=1}^n (X_{Li} + X_{Ui})]^2 - 2n \sum_{i=1}^n (X_{Li}^2 + X_{Ui}^2)\}^2} \\ &= \frac{2n\sigma^2}{2n \sum_{i=1}^n (X_{Li}^2 + X_{Ui}^2) - [\sum_{i=1}^n (X_{Li} + X_{Ui})]^2} \triangleq V_{\hat{\beta}_1}. \end{aligned} \tag{3.70}$$

For $\hat{\beta}_0$, by (3.55) and (3.68), we have

$$\begin{aligned}
Var(\hat{\beta}_0) &= \frac{1}{4n^2}\sum_{i=1}^{n} Var(Y_{Li}+Y_{Ui}) + \frac{1}{4n^2}[\sum_{i=1}^{n}(X_{Li}+X_{Ui})]^2 Var(\hat{\beta}_1) \\
&- \frac{\sum_{i=1}^{n}(X_{Li}+X_{Ui})}{2n^2} Cov[\sum_{i=1}^{n}(Y_{Li}+Y_{Ui}), \hat{\beta}_1] \\
&= \frac{\sigma^2}{2n} + \frac{\sigma^2}{2n}\frac{[\sum_{i=1}^{n}(X_{Li}+X_{Ui})]^2}{2n\sum_{i=1}^{n}(X_{Li}^2+X_{Ui}^2) - [\sum_{i=1}^{n}(X_{Li}+X_{Ui})]^2} \\
&- \frac{\sum_{i=1}^{n}(X_{Li}+X_{Ui})}{2n^2}\frac{\sum_{i=1}^{n}(X_{Li}+X_{Ui})\sum_{i=1}^{n} Var(Y_{Li}+Y_{Ui})}{[\sum_{i=1}^{n}(X_{Li}+X_{Ui})]^2 - 2n\sum_{i=1}^{n}(X_{Li}^2+X_{Ui}^2)} \\
&+ \frac{1}{n}\frac{\sum_{i=1}^{n} Cov(Y_{Li}+Y_{Ui}, X_{Li}Y_{Li}+X_{Ui}Y_{Ui})}{[\sum_{i=1}^{n}(X_{Li}+X_{Ui})]^2 - 2n\sum_{i=1}^{n}(X_{Li}^2+X_{Ui}^2)} \\
&= \frac{\sigma^2}{2n} + \frac{\sigma^2}{2n}\frac{[\sum_{i=1}^{n}(X_{Li}+X_{Ui})]^2}{2n\sum_{i=1}^{n}(X_{Li}^2+X_{Ui}^2) - [\sum_{i=1}^{n}(X_{Li}+X_{Ui})]^2} \\
&- \frac{\sigma^2\sum_{i=1}^{n}(X_{Li}+X_{Ui})}{2n^2}\frac{2n\sum_{i=1}^{n}(X_{Li}+X_{Ui}) - 2n\sum_{i=1}^{n}(X_{Li}+X_{Ui})}{2n\sum_{i=1}^{n}(X_{Li}^2+X_{Ui}^2) - [\sum_{i=1}^{n}(X_{Li}+X_{Ui})]^2} \\
&= \frac{\sum_{i=1}^{n}(X_{Li}^2+X_{Ui}^2)\sigma^2}{2n\sum_{i=1}^{n}(X_{Li}^2+X_{Ui}^2) - [\sum_{i=1}^{n}(X_{Li}+X_{Ui})]^2} \stackrel{\triangle}{=} V_{\hat{\beta}_0}. \qquad (3.71)
\end{aligned}$$

Based on the deviations of $Var(\hat{\beta}_1)$ and $Var(\hat{\beta}_0)$ in (3.70) and (3.71), respectively, and the derivations of $E(\hat{\beta}_1)$ and $E(\hat{\beta}_0)$ in (3.62) and (3.64), we conclude that

$$\hat{\beta}_1 \sim N(\beta_1, V_{\hat{\beta}_1}),\ \hat{\beta}_0 \sim N(\beta_0, V_{\hat{\beta}_0}). \qquad (3.72)$$

Similarly, for $\beta_1 < 0$, by (3.41), we can show that

$$Var(Y_{Li}) = Var(\beta_0 + \beta_1 X_{Ui} + \epsilon_{Li}) = Var(\epsilon_{Li}) = \sigma^2, \qquad (3.73)$$

$$Var(Y_{Ui}) = Var(\beta_0 + \beta_1 X_{Li} + \epsilon_{Ui}) = Var(\epsilon_{Ui}) = \sigma^2, \qquad (3.74)$$

for $i = 1, ..., n$.

Since ϵ_{Li} and ϵ_{Ui} are independent, for $i = 1, ..., n$, we have

$$Cov(Y_{Li}, Y_{Ui}) = Cov(X_{Ui}\beta_1 + \beta_0 + \epsilon_{Li}, X_{Li}\beta_1 + \beta_0 + \epsilon_{Ui}) = Cov(\epsilon_{Li}, \epsilon_{Ui}) = 0. \tag{3.75}$$

Therefore, the distribution of $(Y_{Li}, Y_{Ui})^T$ is a bivariate normal distribution, i.e.,

$$\begin{pmatrix} Y_{Li} \\ Y_{Ui} \end{pmatrix} \sim N_2 \left(\begin{pmatrix} X_{Ui}\beta_1 + \beta_0 \\ X_{Li}\beta_1 + \beta_0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \right), \ i = 1, ..., n. \tag{3.76}$$

By (3.59), (3.60) and (3.76), we also obtain that $Var(\hat{\beta}_1) = V_{\hat{\beta}_1}$, and $Var(\hat{\beta}_0) = V_{\hat{\beta}_0}$. With the unbiasedness of $\hat{\beta}_1$ and $\hat{\beta}_0$ by (3.65), we have

$$\hat{\beta}_1 \sim N(\beta_1, V_{\hat{\beta}_1}), \ \hat{\beta}_0 \sim N(\beta_0, V_{\hat{\beta}_0}). \tag{3.77}$$

Confidence Intervals for Regression Coefficients

Based on the theoretical results generated in Section 3.2.2, we are able to give confidence intervals for the point estimators of the intercept and the slope parameters.

First, we estimate the variance of the error term, i.e., σ^2 in (3.39) through the residuals. From (3.77), recall (3.42), (3.43), (3.44), and (3.45), when the estimate $\hat{\beta}_1$ is positive, we predict the lower and upper points of the response by X_{Li} and X_{Ui}, respectively; while when the estimate $\hat{\beta}_1$ is negative, we predict the lower points of Y_i by the upper bound X_{Ui}, and predict the upper points of Y_i by the lower bound X_{Li}, for $i = 1, ..., n$.

Based on the definition of the covariance in (2.8), the variance $var(\boldsymbol{\epsilon})$ can be estimated by

$$\widehat{Var(\boldsymbol{\epsilon})} = \frac{1}{6n}\sum_{i=1}^{n}[2(r_{Li}-\bar{r})(r_{Li}-\bar{r}) + 2(r_{Li}-\bar{r})(r_{Ui}-\bar{r}) + 2(r_{Ui}-\bar{r})(r_{Ui}-\bar{r})]$$
$$= \frac{1}{3n}\sum_{i=1}^{n}[(r_{Li}-\bar{r})^2 + (r_{Li}-\bar{r})(r_{Ui}-\bar{r}) + (r_{Ui}-\bar{r})^2] \tag{3.78}$$

where $\bar{r} = \frac{1}{n}\sum_{i=1}^{n}\frac{r_{Li}+r_{Ui}}{2}$.

By (3.78), the standard deviation of the error can be estimated by

$$\hat{\sigma} = \sqrt{\widehat{Var(\boldsymbol{\epsilon})}}. \tag{3.79}$$

Since by (3.72) and (3.77), both $\hat{\beta}_1$ and $\hat{\beta}_0$ are normally distributed, confidence intervals for the parameters β_1 and β_0 can be obtained by inverting the $t-$test statistic. By (3.70) and (3.71), we know that the variances of $\hat{\beta}_1$ and $\hat{\beta}_0$, i.e., V_{β_1} and V_{β_0}, are functions of σ, which can be estimated by (3.79). Let the significance level of the $t-$tests be α. Then, we have the $(1-\alpha)100\%$ confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$ to be

$$[\hat{\beta}_1 - t_{n-2,\frac{\alpha}{2}}\sqrt{V_{\hat{\beta}_1}(\hat{\sigma})}, \hat{\beta}_1 + t_{n-2,\frac{\alpha}{2}}\sqrt{V_{\hat{\beta}_1}(\hat{\sigma})}], \tag{3.80}$$
$$[\hat{\beta}_0 - t_{n-2,\frac{\alpha}{2}}\sqrt{V_{\hat{\beta}_0}(\hat{\sigma})}, \hat{\beta}_1 + t_{n-2,\frac{\alpha}{2}}\sqrt{V_{\hat{\beta}_0}(\hat{\sigma})}] \tag{3.81}$$

where $V_{\hat{\beta}_1}(\hat{\sigma})$ and $V_{\hat{\beta}_0}(\hat{\sigma})$ represent the estimated variances of $\hat{\beta}_1$ and $\hat{\beta}_0$ by $\hat{\sigma}$, respectively, given in (3.70) and (3.71), respectively.

3.3 Predictions

In addition to confidence intervals for the point estimators of the regression coefficients, the predicted value and the confidence interval for the response variable also need to be addressed

in the interval-valued data regression model. An application for this topic is illustrated by the following example. Suppose a market research firm has the records of the consumers' age, gender, income and annual flight expenses from a number of airlines. The firm wants to discover relationships between annual flight expenses and covariates by a linear model built on the aggregated interval-valued age and income levels. Assume we have a new consumer with age, e.g., 30 years and income, e.g., \$45,000/year. It would be quite meaningful to find out the predicted value and confidence interval of this consumer's annual travel expenditure.

For the simple regression specified in (3.5), assume we have a new observation with realization of the explanatory variable to be $X^{new} = (X_L^{new}, X_U^{new})$. Then, the lower and upper points of a predicted response $\hat{Y}$ can be obtained as

$$\hat{Y}_L = min\{\hat{\beta}_0 + \hat{\beta}_1 X_L^{new}, \hat{\beta}_0 + \hat{\beta}_1 X_U^{new}\}, \hat{Y}_U = max\{\hat{\beta}_0 + \hat{\beta}_1 X_L^{new}, \hat{\beta}_0 + \hat{\beta}_1 X_U^{new}\}. \quad (3.82)$$

By setting the lower point and the upper point equal to the minimum and the maximum values in (3.82), respectively, we guarantee that the predicted lower bound is always not greater than the predicted upper bound.

Next we consider confidence intervals of the predicted response. Again we need to consider the two cases, namely when the slope parameter is positive or negative.

For the slope parameter $\beta_1 \geq 0$, by (3.46),

$$\hat{Y}_L = \hat{\beta}_0 + \hat{\beta}_1 X_L^{new}, \hat{Y}_U = \hat{\beta}_0 + \hat{\beta}_1 X_U^{new}. \quad (3.83)$$

Then, by (3.72) and (3.77), we have

$$
\begin{aligned}
E(\hat{Y}_L) &= \beta_0 + \beta_1 X_L^{new},\ E(\hat{Y}_U) = \beta_0 + \beta_1 X_U^{new}, \\
Var(\hat{Y}_L) &= Var(\hat{\beta}_0) + (X_L^{new})^2 Var(\hat{\beta}_1) = Var(\hat{\beta}_0) + (X_L^{new})^2 \hat{\beta}_1 = V_{\hat{\beta}_0} + (X_L^{new})^2 V_{\hat{\beta}_1}, \\
Var(\hat{Y}_U) &= Var(\hat{\beta}_0) + (X_U^{new})^2 Var(\hat{\beta}_1) = Var(\hat{\beta}_0) + (X_U^{new})^2 \hat{\beta}_1 = V_{\hat{\beta}_0} + (X_U^{new})^2 V_{\hat{\beta}_1}.
\end{aligned}
\tag{3.84}
$$

By (3.83),

$$
\begin{aligned}
Cov(\hat{Y}_L, \hat{Y}_U) &= Cov(\hat{\beta}_0 + \hat{\beta}_1 X_L^{new}, \hat{\beta}_0 + \hat{\beta}_1 X_U^{new}) \\
&= E[(\hat{\beta}_0 + \hat{\beta}_1 X_L^{new})(\hat{\beta}_0 + \hat{\beta}_1 X_U^{new})] - E(\hat{\beta}_0 + \hat{\beta}_1 X_L^{new}) E(\hat{\beta}_0 + \hat{\beta}_1 X_U^{new}) \\
&= V_{\beta_0} + V_{\beta_1} X_L^{new} X_U^{new} + (X_L^{new} + X_U^{new}) Cov(\hat{\beta}_0, \hat{\beta}_1) \\
&\triangleq \delta.
\end{aligned}
\tag{3.85}
$$

Therefore, since both $\hat{\beta}_0$ and $\hat{\beta}_1$ are normally distributed, and $\hat{Y}_L$ and $\hat{Y}_U$ are linear functions of $\hat{\beta}_0$ and $\hat{\beta}_1$, by (3.84) and (3.85), we have

$$
\begin{pmatrix} \hat{Y}_L \\ \hat{Y}_U \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \beta_0 + \beta_1 X_L^{new} \\ \beta_0 + \beta_1 X_U^{new} \end{pmatrix}, \begin{pmatrix} V_{\hat{\beta}_0} + (X_L^{new})^2 V_{\hat{\beta}_1} & \delta \\ \delta & V_{\hat{\beta}_0} + (X_U^{new})^2 V_{\hat{\beta}_1} \end{pmatrix} \right). \tag{3.86}
$$

Then, since both variances $V_{\hat{\beta}_0}$ and $V_{\hat{\beta}_1}$ need to be estimated by $\hat{\sigma}^2$, by (3.86),

$$
\begin{aligned}
\frac{\hat{Y}_L - (\beta_0 + \beta_1 X_L^{new})}{\sqrt{\hat{V}_{\hat{\beta}_0} + (X_L^{new})^2 \hat{V}_{\hat{\beta}_1}}} &\sim t_{n-2}, \\
\frac{\hat{Y}_U - (\beta_0 + \beta_1 X_U^{new})}{\sqrt{\hat{V}_{\hat{\beta}_0} + (X_U^{new})^2 \hat{V}_{\hat{\beta}_1}}} &\sim t_{n-2}
\end{aligned}
$$

where t_{n-2} represents the t distribution with degree of freedom $n-2$.

Setting the significance level to be α, and inverting the $t-$statistic above, we obtain the

$(1-\alpha)100\%$ confidence intervals for $\hat{Y}_L$ and $\hat{Y}_U$:

$$\hat{Y}_L: \ [\hat{Y}_L - t_{n-2,\frac{\alpha}{2}}\sqrt{\hat{V}_{\hat{\beta}_0} + (X_L^{new})^2\hat{V}_{\hat{\beta}_1}}, \hat{Y}_L + t_{n-2,\frac{\alpha}{2}}\sqrt{\hat{V}_{\hat{\beta}_0} + (X_L^{new})^2\hat{V}_{\hat{\beta}_1}}\ ], \tag{3.87}$$

$$\hat{Y}_U: \ [\hat{Y}_U - t_{n-2,\frac{\alpha}{2}}\sqrt{\hat{V}_{\hat{\beta}_0} + (X_U^{new})^2\hat{V}_{\hat{\beta}_1}}, \hat{Y}_U + t_{n-2,\frac{\alpha}{2}}\sqrt{\hat{V}_{\hat{\beta}_0} + (X_U^{new})^2\hat{V}_{\hat{\beta}_1}}\ ]. \tag{3.88}$$

Similarly, for the slope parameter $\beta_1 < 0$, we have the $(1-\alpha)100\%$ confidence intervals for $\hat{Y}_L$ and $\hat{Y}_U$, respectively, as

$$\hat{Y}_L: \ [\hat{Y}_L - t_{n-2,\frac{\alpha}{2}}\sqrt{\hat{V}_{\hat{\beta}_0} + (X_U^{new})^2\hat{V}_{\hat{\beta}_1}}, \hat{Y}_L + t_{n-2,\frac{\alpha}{2}}\sqrt{\hat{V}_{\hat{\beta}_0} + (X_U^{new})^2\hat{V}_{\hat{\beta}_1}}\ ], \tag{3.89}$$

$$\hat{Y}_U: \ [\hat{Y}_U - t_{n-2,\frac{\alpha}{2}}\sqrt{\hat{V}_{\hat{\beta}_0} + (X_L^{new})^2\hat{V}_{\hat{\beta}_1}}, \hat{Y}_U + t_{n-2,\frac{\alpha}{2}}\sqrt{\hat{V}_{\hat{\beta}_0} + (X_L^{new})^2\hat{V}_{\hat{\beta}_1}}\ ]. \tag{3.90}$$

3.4 Measurement of Model Fit

Similar to classical data regression, we can use the R-square statistic, which is the fraction of explained variation by the linear regression model over the total variation of the response to measure if the model fits the data well.

In Xu (2010) [8], the R-square is calculated by

$$R^2 = \frac{Var(\hat{Y})}{Var(Y)} \tag{3.91}$$

where Y indicates the observed interval valued response, and $\hat{Y}$ indicates the predicted interval valued response. The variances are calculated by (2.3).

3.5 Determination of Likelihood Function Form

Since we have observed that the likelihood functions of residuals for $\beta_1 \geq 0$ and $\beta_1 < 0$ are different, it is of importance to choose the correct likelihood function in order to obtain

valid point estimates. Two approaches can be considered to use, which will be illustrated as follows.

The first approach is to select the likelihood form with the larger value. Given the data, by (3.50) and (3.56), we can calculate both values of the likelihood functions, after obtaining the maximum likelihood estimators of β_1 and β_0 under $\beta_1 \geq 0$ and $\beta_1 < 0$, respectively. Then, we compare the two values, and choose to use the likelihood function with the larger value, and the corresponding MLEs, i.e., $\hat{\beta}_1$ and $\hat{\beta}_0$, are treated as the point estimators for the slope and the intercept parameters.

The second approach is by checking the sign of the correlation between the response and the explanatory variable. By (2.8) and (2.9), we can calculate the correlation coefficient between the response variable Y and the explanatory variable X, denoted by $r = Corr(Y, X)$. If r is no less than zero, we can assume the slope parameter β_1 is non-negative, and therefore we choose the likelihood function in (3.50); on the other hand, if r is less than zero, we then assume the slope parameter β_1 is negative, and choose the likelihood function in (3.56).

Compared to the second approach, the first approach is far less efficient, since we have to first calculate the MLEs for β_1 and β_0 under the assumptions that $\beta_1 \geq 0$ and $\beta_1 < 0$, respectively, in order to obtain the values of the likelihood functions under the two scenarios and compare. For the second approach, there is no need to calculate the MLEs before selecting the appropriate likelihood form. In Chapter 5, we choose the second approach to determine the likelihood function, and give the MLEs for the regression coefficients in our data analysis of two real data sets.

3.6 APPENDIX

R Function of the Maximum Likelihood Method for Statistical Inference on Interval-Valued Data Regression

```
# Input:
# X_L: vector of lower point of X
# X_U: vector of upper point of X
# Y_L: vector of lower point of Y
# Y_U: vector of upper point of Y



est_ord1 <- function(X_L, X_U, Y_L, Y_U){
  n <- length(X_L)
  sum_X <- sum(X_L + X_U)
  sum_Y <- sum(Y_L + Y_U)
  sum_XY <- sum(X_L * Y_L + X_U * Y_U)
  sum_X2 <- sum_X^2
  sum_X22 <- sum(X_L^2 + X_U^2)

  # 1) Point estimates of parameters
  beta1_h <- ((sum_X * sum_Y) - 2*n*sum_XY) / (sum_X2 - 2*n*sum_X22)
  beta0_h <- 1 / (2*n) * (sum_Y - beta1_h * sum_X)

  # 2) estimate sd of error term
  n <- length(X_L)
```

```
# calculate residuals
r_L <- Y_L - beta0_h - beta1_h * X_L
r_U <- Y_U - beta0_h - beta1_h * X_U

r_mean <- sum(r_L + r_U) / (2*n)

var_e <- 1/(3*n)*sum((r_L - r_mean)^2 + (r_L - r_mean)*(r_U - r_mean) +
                       (r_U - r_mean)^2)
sd_e <- sqrt(var_e)

# 3) Variances of parameter estimates
var_beta1 <- 2*n*sd_e^2 / (2*n*sum_X22 - sum_X2)
var_beta0 <- sum_X22 * sd_e^2 / (2*n*sum_X22 - sum_X2)

# 4) calculate 95% C.I.
lower_beta1 <- beta1_h - qt(.975, n-2)*sqrt(var_beta1)
upper_beta1 <- beta1_h + qt(.975, n-2)*sqrt(var_beta1)

lower_beta0 <- beta0_h - qt(.975, n-2)*sqrt(var_beta0)
upper_beta0 <- beta0_h + qt(.975, n-2)*sqrt(var_beta0)

# 5) calculate R square
Y_L_hat <- beta0_h + beta1_h * X_L
Y_U_hat <- beta0_h + beta1_h * X_U

var_Y_hat <- 1 / (3*n) * sum(Y_L_hat^2 + Y_L_hat * Y_U_hat + Y_U_hat^2) -
```

```
                1 / (4*n^2) * sum(Y_L_hat + Y_U_hat)^2
  var_Y <- 1 / (3*n) * sum(Y_L^2 + Y_L * Y_U + Y_U^2) -
           1 / (4*n^2) * sum(Y_L + Y_U)^2
  ssm <- n*var_Y_hat
  sst <- n*var_Y
  R2 <- var_Y_hat / var_Y



  res <- paste("beta1: ", round(beta1_h, 3), "beta0: "
        , round(beta0_h, 3), "var(beta1): "
        , round(var_beta1, 4), "var(beta0): "
        , round(var_beta0, 4), "C.I. for beta_1: ["
        , round(lower_beta1, 3), ",", round(upper_beta1, 3),"]",
        "C.I. for beta_0: [", round(lower_beta0, 3), ",",
        round(upper_beta0, 3),"]", "standard deviation of error: ",
        round(sd_e, 4), "R Square: ", round(R2, 3))

  return(res)
}
```

Chapter 4

SIMULATION

In this chapter, two simulation methods for interval-valued data are introduced in Section 4.1. Then in Section 4.2, we implement the proposed approach for different parameter settings, and evaluate the performances under various settings of parameters.

4.1 Simulation: Methodology

To study the performance of the proposed method, we first illustrate the two simulation methods in this section.

Method I

First, we simulate data for applying the proposed method as follows. At the beginning, we generate the interval means of the explanatory variable X, $X^{(c)}$, by randomly sampling from a normal distribution $N(\mu, \sigma^2)$, and the interval ranges of X, $X^{(r)}$, from a uniform distribution with positive support, i.e., $Uniform[a, b]$. This range value can also be generated from some other distributions, such as the exponential, the chi-square, or the log-normal distribution. Then, interval-valued observations for the explanatory variable can be simulated as $x_i =$

$[x_{Li}, x_{Ui}] = [x_i^{(c)} - 0.5x_i^{(r)}, x_i^{(c)} + 0.5x_i^{(r)}]$, for $i = 1, ..., n$, where n is the number of observations. By (3.5), we know that the interval-valued realization of the response variable, $[y_{Li}, y_{Ui}]$, is composed of two elements, the systematic component $\beta_0 + [x_{Li}, x_{Ui}]\beta_1$ and the random error $[\epsilon_{Li}, \epsilon_{Ui}]$, for $i = 1, ..., n$. Based on the assumption shown by (3.39), for each observation, the error, ϵ_{Li}, for the lower bound response y_{Li}, and the error, ϵ_{Ui}, for the upper bound response y_{Ui}, are independently generated from a normal distribution $N(0, \sigma_e^2)$, respectively. Then, the intervals for the response variable are obtained by

$$[y_{Li}, y_{Ui}] = [\beta_0 + x_{Li}\beta_1 + \epsilon_{Li}, \beta_0 + x_{Ui}\beta_1 + \epsilon_{Ui}] \tag{4.1}$$

for $i = 1, ..., n$.

This method to simulate interval-valued data guarantees that the expected interval response Y and the interval explanatory variable X follow a linear relationship.

The drawback about this simulation method is that the ranges of Y are always positively correlated with the ranges of X, which may not always be true in reality.

Method II

The other simulation method for interval-valued data originates from a common way that interval data sets arise, which is by aggregating classical data. Similar to Method I, we first generate the interval means, $X^{(c)}$, and ranges, $X^{(r)}$, for the explanatory variable X by randomly sampling from a normal distribution, $N(\mu, \sigma^2)$ and a truncated normal distribution, denoted by $Trun - N(a, b, \mu_0, \sigma_0^2)$, where a is the lower bound of support, b is the upper bound of support, μ_0 is the mean value, and σ_0 is the standard deviation value, respectively. Then, the interval for X can be obtained by $[x_L, x_U] = [x^{(c)} - 0.5x^{(r)}, x^{(c)} + 0.5x^{(r)}]$.

By the basic assumption that the distribution within each interval is uniform, for the ith observation, a certain number of values, i.e., m values are randomly drawn from the uniform

distribution $U(x_{Li}, x_{Ui})$, for $i = 1, ..., n$, where m is pre-specified, or a value drawn from a distribution with integer support for more general cases, each of the m values is denoted by x_{il}. As in the first method, for the ith observation, we generate the random errors for the lower and upper bounds of the response, i.e., ϵ_{Li} and ϵ_{Ui}, respectively, by randomly sampling from a normal distribution $N(0, \sigma_e^2)$. Then, the interval realization for the response variable of the ith observation $y_i = [y_{Li}, y_{Ui}]$ can be determined by

$$\begin{aligned} y_{Li} &= \min_{l \in \{1,...,m\}} \{\beta_0 + x_{il}\beta_1\} + \epsilon_{Li}, \\ y_{Ui} &= \max_{l \in \{1,...,m\}} \{\beta_0 + x_{il}\beta_1\} + \epsilon_{Ui}, \end{aligned} \tag{4.2}$$

where $\epsilon_{Li}, \epsilon_{Ui} \overset{iid}{\sim} N(0, \sigma_e^2)$, for $i = 1, ..., n$, and $l = 1, ..., m$.

This simulation method is closer to how interval data arise in practice. For example, the daily temperature is described by an interval, for the lower bound to be the minimum temperature, while the upper bound is the maximum temperature among a number of measurements during a day, respectively. A drawback of this method is that the obtained intervals for the response variable y_i, $i = 1, ..., n$, cannot be guaranteed to follow a uniform distribution internally.

4.2 Simulation: Case Study

In this section, we conduct simulations by the two methods described in Section 4.1, respectively, to investigate the performance of the proposed approach for interval data regression. For each of the two simulation methods, we give different settings on pre-determined parameters to compare and analyse corresponding results.

Simulation: Method I

The settings for simulation studies by the first method are as follows:

1. μ: the values of the means of the normal distributions from which the interval means $X^{(c)}$ are generated, -35, -25, -15, -5, 5, 15, 25, 35, 45, 55, 65, 75, 85, 95, 105, 115, 125;

2. σ: the standard deviation of the normal distributions from which the interval means $X^{(c)}$ are generated; this is set to be equal to 7;

3. n_0: the numbers of observations with the interval means at each of the seventeen values in Step **1**. The n_0 is randomly sampled from the discrete uniform distribution $Uniform(5, 9)$ with support $k \in \{5, 6, 7, 8, 9\}$;

4. σ_e: the standard deviation of the error term to be generated, 3, 7, and 10;

5. (a, b): the lower and upper bounds of the uniform distribution from which the interval ranges $X^{(r)}$ are generated, (6.5, 9.25) and (10, 12.45);

6. β_1: the true slope parameters, 0.64, 2.15, -3.21;

7. β_0: the true intercept parameters, 68.57 and -43.29;

8. B repetition times of drawing samples: 2000, 5000, and 10000.

We conduct simulations based on each of the settings delineated here, and use the proposed method to estimate the slope parameter β_1 and the intercept parameter β_0, both by point estimation and confidence interval. Then, we compare the results with the true parameter values, and we calculate variances, mean square errors (MSE) as well as empirical confidence intervals for the point estimators $\hat{\beta}_1$ and $\hat{\beta}_0$, respectively.

To illustrate, we first take the setting $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$, $\beta_1 = 2.15$ and $\beta_0 = 68.57$ as an example. Figure 4.1 shows the scatter plot with the fitted regression line given by the average values of $\hat{\beta}_1$ and $\hat{\beta}_0$; and Figure 4.2 shows histograms of the observations along the two point estimators. The results in Figure 4.2 are based on 10000 repetitions of sampling.

EXAMPLE I: When we set simulations for parameter values $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$, $\beta_1 = 2.15$, $\beta_0 = 68.57$, we obtain the data shown in Figure 4.1. The histogram plots of the resulting estimates for the slope parameter $(\hat{\beta}_1)$ and for the intercept parameter $(\hat{\beta}_0)$ are shown in Figure 4.2.

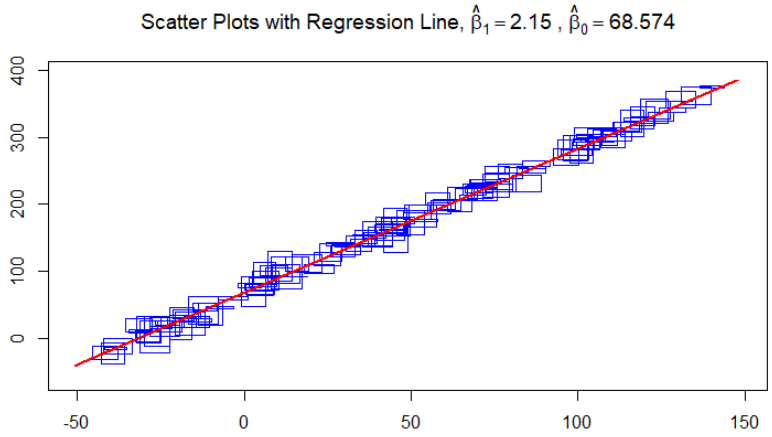

Figure 4.1: Scatter plot - $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$, $\beta_1 = 2.15$, $\beta_0 = 68.57$

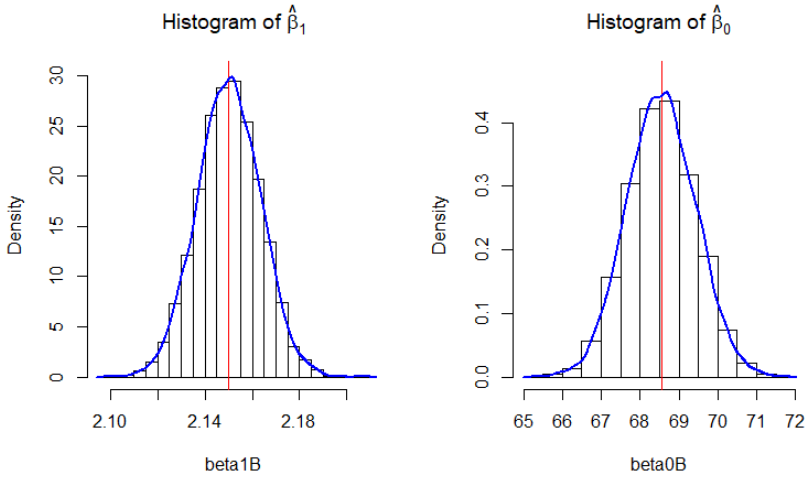

Figure 4.2: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$, $\beta_1 = 2.15$, $\beta_0 = 68.57$

In Figure 4.2, the values of $\hat{\beta}_1$ and $\hat{\beta}_0$ obtained by different repetitions are indicated on the $x-$axis by beta1B and beta0B, respectively. The two vertical red lines on Figure 4.2 display the positions of the true parameter values, i.e., $\beta_1 = 2.15$ and $\beta_0 = 68.57$. The following table summarizes the simulation results for the setting in EXAMPLE I for the different numbers of repetitions, $B = (2000, 5000, 10000)$.

Table 4.1: *Summary of Simulation by Method I* : $\beta_1 = 2.15,\ \beta_0 = 68.57$
$\sigma_e = 10,\ (a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	68.560	2×10^{-4}	0.717	[2.124, 2.176]	[66.894, 70.240]
5000	2.150	68.571	2×10^{-4}	0.761	[2.125 , 2.175]	[66.861, 70.288]
10000	2.150	68.586	2×10^{-4}	0.756	[2.124, 2.175]	[66.848, 70.289]

From Table 4.1, we can observe that the averages of the point estimators for both β_1 and β_0 are equal to or quite close to the true values, with small MSEs, indicating the proposed approach gives accurate estimations for the regression coefficients, especially for β_1 with different repetitions from 2000 to 10000. From the histograms in Figure 4.2, it can be observed that both of $\hat{\beta}_1$ and $\hat{\beta}_0$ are distributed with shapes consistent with normality, which verifies the normal property shown in (3.72) and (3.77). The 95% confidence intervals of both $\hat{\beta}_1$ and $\hat{\beta}_0$ given in Table 4.1 cover the true values of β_1 and β_0 and have almost the same length from the lower bound to the true value, and from the upper bound to the true value.

We now will use different sets of values for $(\beta_1,\ \beta_0)$. There are six sets of values for the error standard deviation σ_e, and the uniform distribution bounds (a, b). For each of these sets, there are six different pairs of values for the regression parameters β_1 and β_0. The simulation results for each of these set $\times$ pair ($6\times$ 6) combinations are provided along the same lines above as illustrated in Figure 4.1 and Table 4.1 for EXAMPLE I. These are

briefly described as follows. Then, for each (β_1, β_0) pairing, comparisons of these results are discussed and presented in Tables 4.37-4.42.

Method 1 - <u>Set 1</u>: $\sigma_e = 3$, $(a, b) = (6.5,\ 9.25)$

(1) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.3. Table 4.2 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

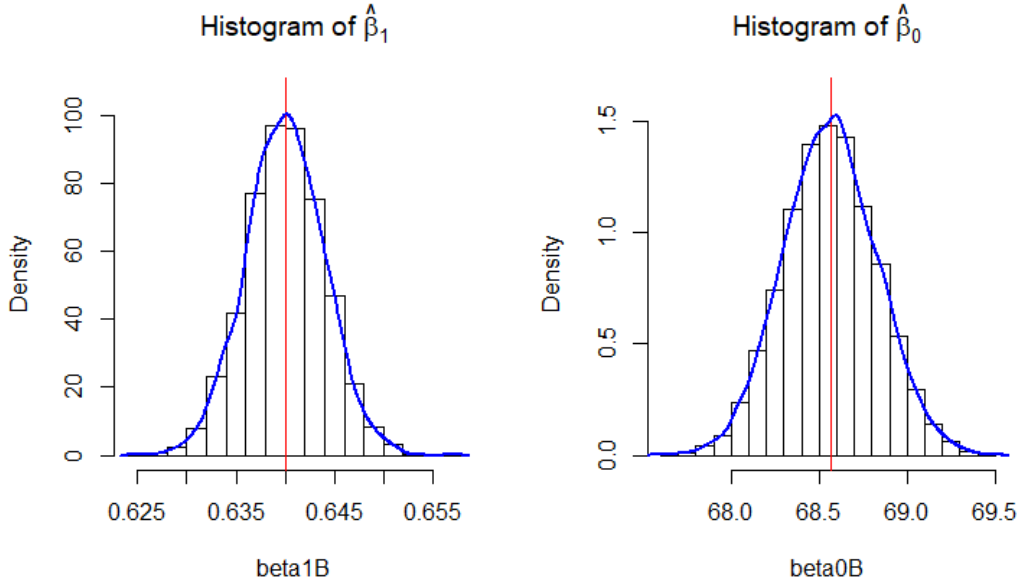

Figure 4.3: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3$, $(a, b) = (6.5, 9.25)$, $\beta_1 = 0.64$, $\beta_0 = 68.57$

Table 4.2: *Summary of Simulation by Method I* : $\beta_1 = 0.64,\ \beta_0 = 68.57$
$\sigma_e = 3,\ (a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	68.570	1×10^{-5}	0.068	[0.633, 0.647]	[68.058, 69.079]
5000	0.640	68.566	2×10^{-5}	0.069	[0.632, 0.648]	[68.047, 69.080]
10000	0.640	68.565	2×10^{-5}	0.070	[0.632, 0.648]	[68.051, 69.083]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.565, with the MSE to be 0.070.

(2) When $\boldsymbol{\beta}_1 = 0.64,\ \boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.4. Table 4.3 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

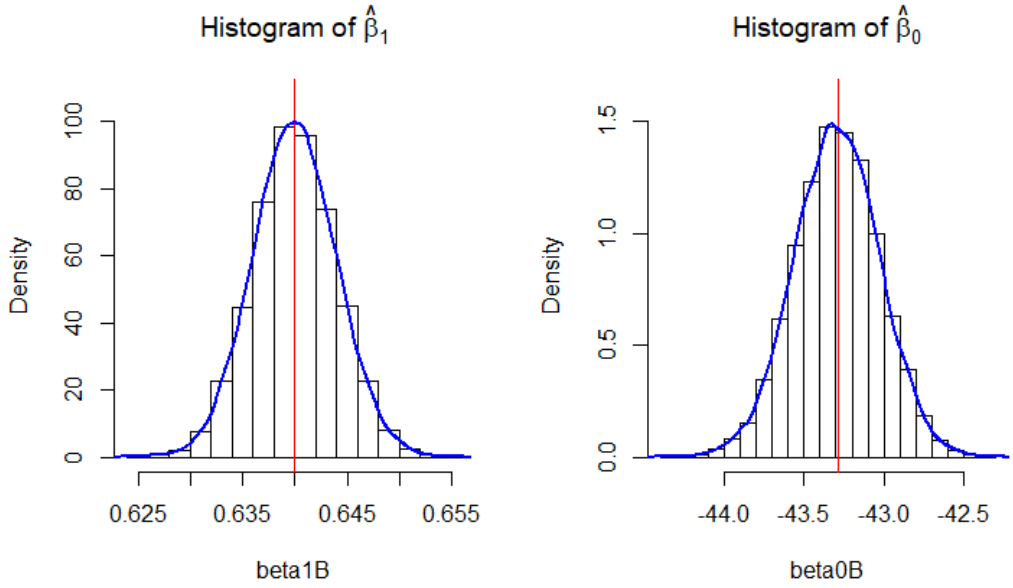

Figure 4.4: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3$, $(a, b) = (6.5, 9.25)$, $\beta_1 = 0.64$, $\beta_0 = -43.29$

Table 4.3: *Summary of Simulation by Method I* : $\beta_1 = 0.64$, $\beta_0 = -43.29$ $\sigma_e = 3$, $(a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	-43.289	2×10^{-5}	0.067	[0.632, 0.647]	[-43.811, -42.792]
5000	0.640	-43.286	1×10^{-5}	0.072	[0.632, 0.647]	[-43.821, -42.779]
10000	0.640	-43.294	2×10^{-5}	0.071	[0.632, 0.648]	[-43.815, -42.780]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.294, with the MSE to be 0.071.

(3) When $\boldsymbol{\beta}_1 = 2.15$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.5. Table 4.4 provides the overall estimates for each parame-

ter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

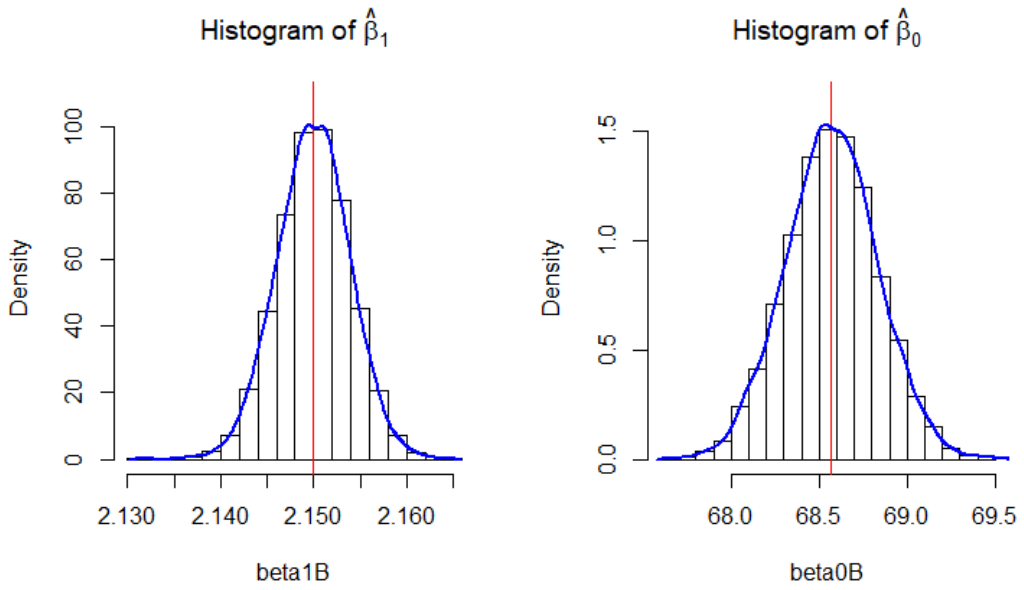

Figure 4.5: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 3$, $(a, b) = (6.5, 9.25)$, $\beta_1 = 2.15$, $\beta_0 = 68.57$

Table 4.4: *Summary of Simulation by Method I* : $\beta_1 = 2.15$, $\beta_0 = 68.57$
$\sigma_e = 3$, $(a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	68.572	2×10^{-5}	0.069	[2.143, 2.157]	[68.051, 69.108]
5000	2.150	68.570	2×10^{-5}	0.068	[2.142, 2.158]	[68.055, 69.086]
10000	2.150	68.573	2×10^{-5}	0.068	[2.142, 2.157]	[68.066, 69.085]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.573, with the MSE to be 0.068.

(4) When $\boldsymbol{\beta}_1 = 2.15,\ \boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.6. Table 4.5 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

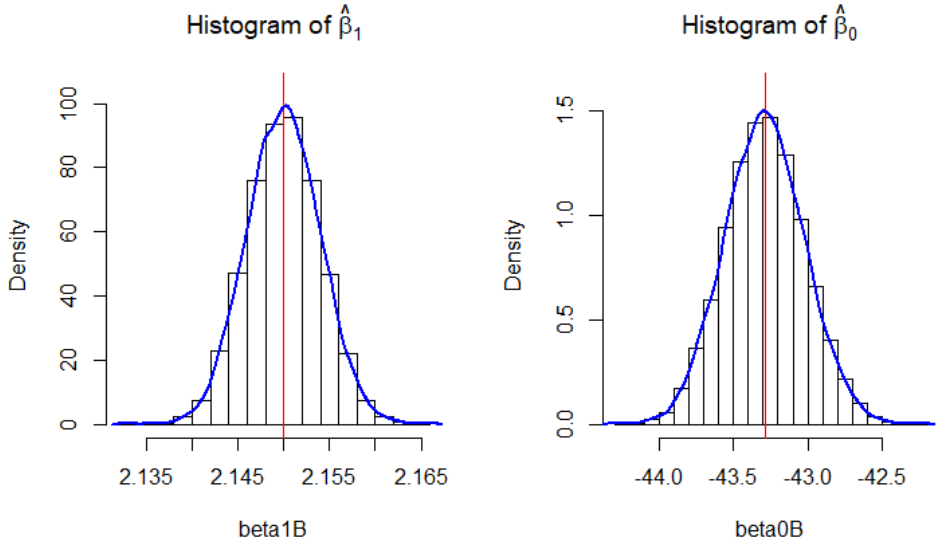

Figure 4.6: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3,\ (a, b) = (6.5, 9.25),\ \beta_1 = 2.15,\ \beta_0 = -43.29$

Table 4.5: *Summary of Simulation by Method I* : $\beta_1 = 2.15,\ \beta_0 = -43.29$ $\sigma_e = 3,\ (a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	-43.282	2×10^{-5}	0.074	[2.142, 2.157]	[-43.836, -42.771]
5000	2.150	-43.294	2×10^{-5}	0.070	[2.142, 2.158]	[-43.800, -42.769]
10000	2.150	-43.288	2×10^{-5}	0.071	[2.142, 2.158]	[-43.803, -42.763]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.288, with the MSE to be 0.071.

(5) When $\boldsymbol{\beta}_1 = -3.21,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.7. Table 4.6 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

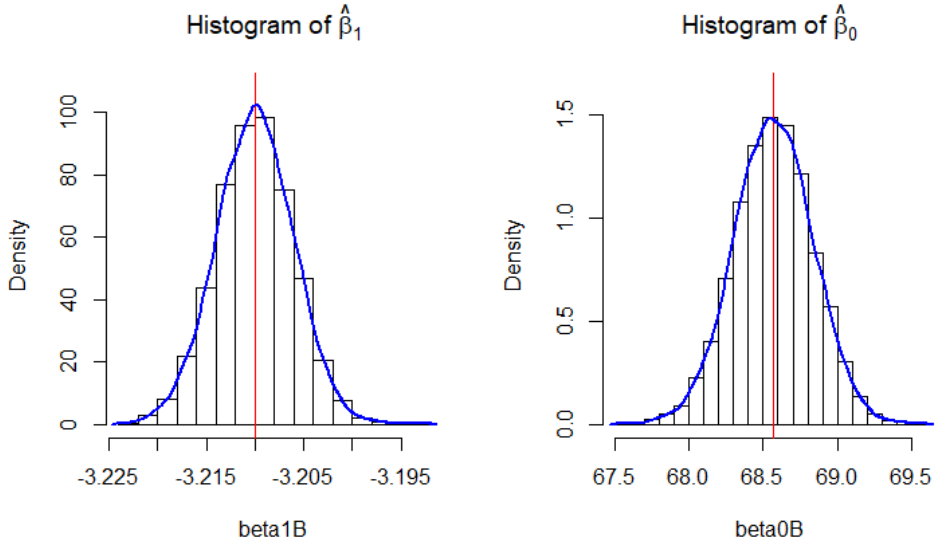

Figure 4.7: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3,\ (a, b) = (6.5, 9.25),\ \beta_1 = -3.21,\ \beta_0 = 68.57$

Table 4.6: *Summary of Simulation by Method I* : $\beta_1 = -3.21,\ \beta_0 = 68.57$ $\sigma_e = 3,\ (a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	68.573	2×10^{-5}	0.069	[-3.218, -3.202]	[68.074, 69.091]
5000	-3.210	68.573	2×10^{-5}	0.068	[-3.218, -3.202]	[68.052, 69.081]
10000	-3.210	68.571	2×10^{-5}	0.069	[-3.218, -3.202]	[68.047, 69.082]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.571, with the MSE to be 0.069.

(6) When $\boldsymbol{\beta}_1 = -3.21,\ \boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.8. Table 4.7 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

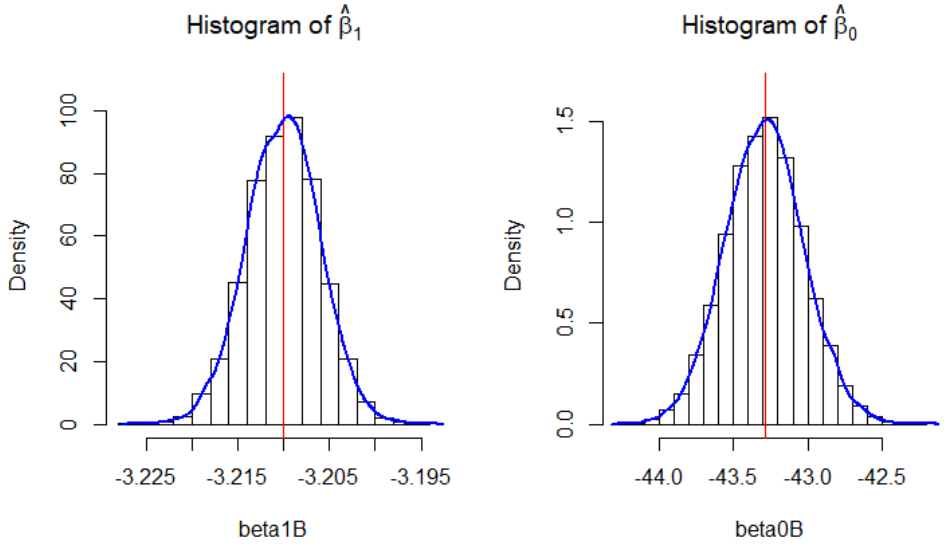

Figure 4.8: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3$, $(a, b) = (6.5, 9.25)$, $\beta_1 = -3.21$, $\beta_0 = -43.29$

Table 4.7: *Summary of Simulation by Method I* : $\beta_1 = -3.21,\ \beta_0 = -43.29$, $\sigma_e = 3$, $(a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	-43.281	1×10^{-5}	0.072	[-3.218, -3.202]	[-43.804, -42.768]
5000	-3.210	-43.294	2×10^{-5}	0.070	[-3.217, -3.202]	[-43.814, -42.779]
10000	-3.210	-43.289	2×10^{-5}	0.069	[-3.218, -3.202]	[-43.800, -42.774]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.289, with the MSE to be 0.069.

Method 1 - Set 2: $\sigma_e = 7,\ (a, b) = (6.5,\ 9.25)$

(1) When $\boldsymbol{\beta}_1 = 0.64,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.9. Table 4.8 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

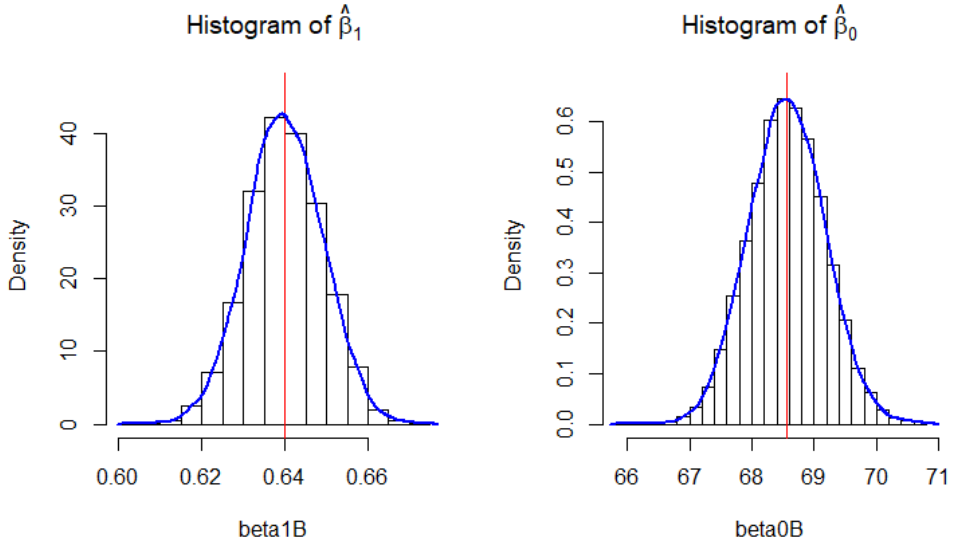

Figure 4.9: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7,\ (a, b) = (6.5, 9.25),\ \beta_1 = 0.64,\ \beta_0 = 68.57$

Table 4.8: *Summary of Simulation by Method I* : $\beta_1 = 0.64,\ \beta_0 = 68.57$ $\sigma_e = 7,\ (a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	68.586	9×10^{-5}	0.399	[0.621, 0.659]	[67.342, 69.837]
5000	0.640	68.571	8×10^{-5}	0.371	[0.623, 0.658]	[67.374, 69.781]
10000	0.640	68.566	8×10^{-5}	0.370	[0.622, 0.658]	[67.397, 69.760]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.566, with the MSE to be 0.370.

(2) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.10. Table 4.9 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

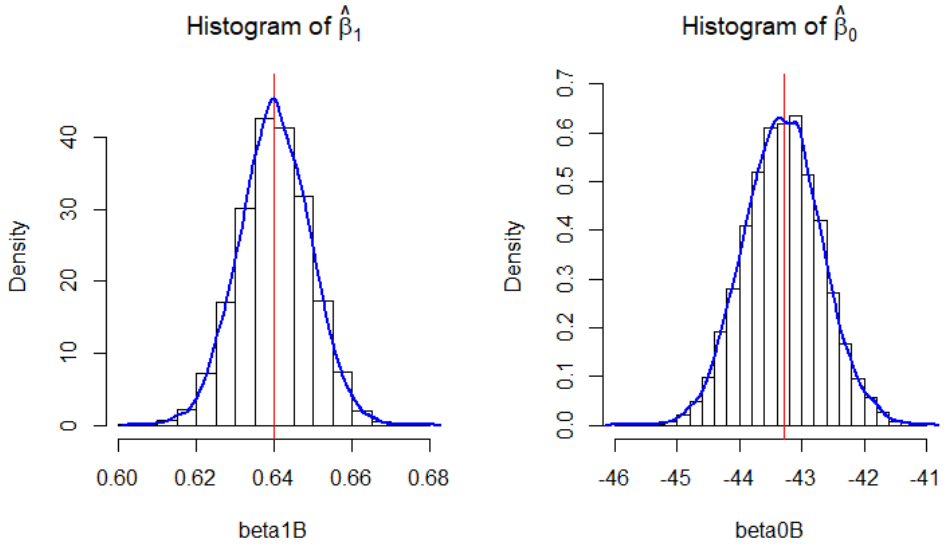

Figure 4.10: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $(a, b) = (6.5, 9.25)$, $\beta_1 = 0.64$, $\beta_0 = -43.29$

Table 4.9: *Summary of Simulation by Method I* : $\beta_1 = 0.64,\ \beta_0 = -43.29$
$\sigma_e = 7,\ (a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	-43.285	8×10^{-5}	0.372	[0.622, 0.658]	[-44.535, -42.133]
5000	0.640	-43.281	9×10^{-5}	0.378	[0.622, 0.659]	[-44.498, -42.097]
10000	0.640	-43.303	8×10^{-5}	0.382	[0.622, 0.658]	[-44.498, -42.080]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.303, with the MSE to be 0.382.

(3) When $\boldsymbol{\beta}_1 = 2.15,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.11. Table 4.10 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

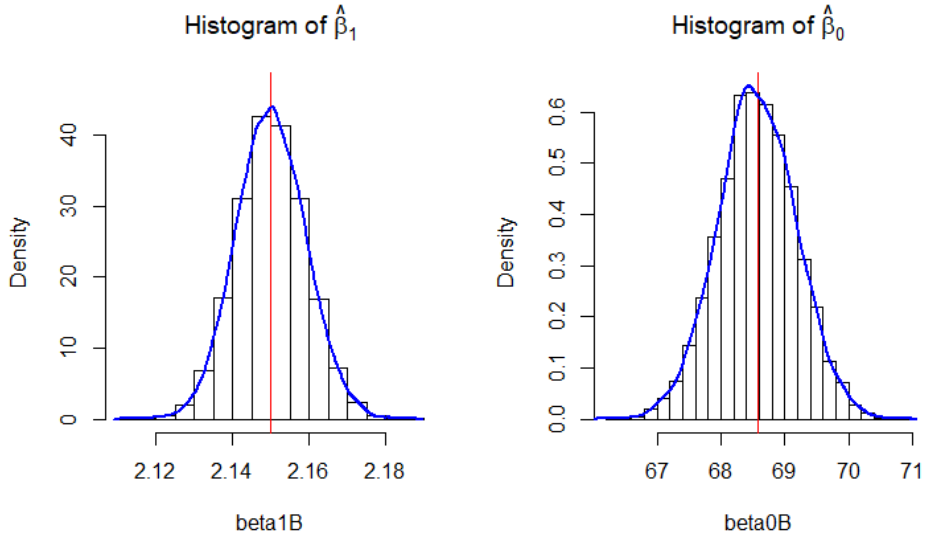

Figure 4.11: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 7,\ (a, b) = (6.5, 9.25),\ \beta_1 = 2.15,\ \beta_0 = 68.57$

Table 4.10: *Summary of Simulation by Method I* : $\beta_1 = 2.15,\ \beta_0 = 68.57$
$\sigma_e = 7,\ (a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	68.586	8×10^{-5}	0.376	[2.131, 2.167]	[67.316, 69.774]
5000	2.150	68.576	8×10^{-5}	0.375	[2.132, 2.168]	[67.391, 69.755]
10000	2.150	68.569	8×10^{-5}	0.375	[2.133, 2.168]	[67.370, 69.772]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.569, with the MSE to be 0.375.

(4) When $\boldsymbol{\beta}_1 = 2.15,\ \boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.12. Table 4.11 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

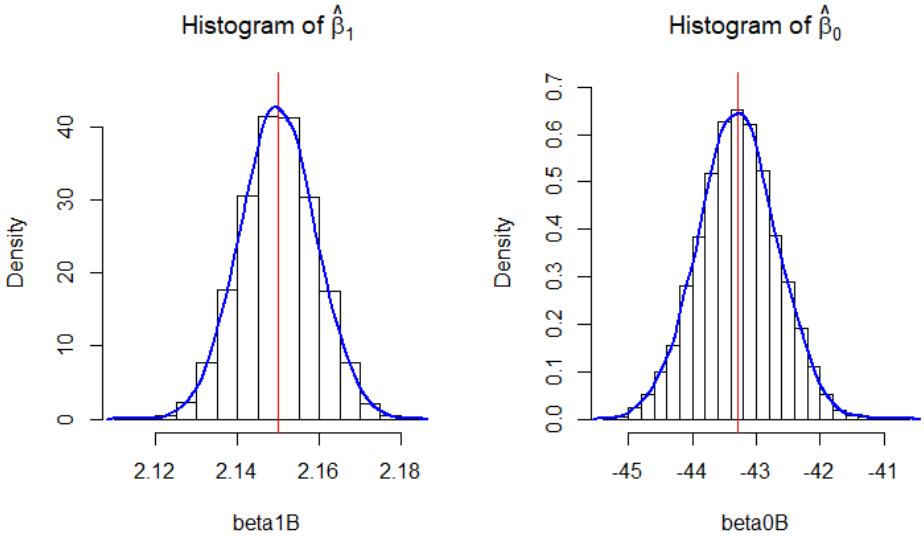

Figure 4.12: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 7,\ (a, b) = (6.5, 9.25),\ \beta_1 = 2.15,\ \beta_0 = -43.29$

Table 4.11: *Summary of Simulation by Method I* : $\beta_1 = 2.15,\ \beta_0 = -43.29$ $\sigma_e = 7,\ (a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	-43.281	8×10^{-5}	0.367	[2.132, 2.168]	[-44.451, -42.123]
5000	2.150	-43.290	8×10^{-5}	0.377	[2.132, 2.168]	[-44.492, -42.100]
10000	2.150	-43.288	8×10^{-5}	0.375	[2.132, 2.168]	[-44.495, -42.099]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.288, with the MSE to be 0.375.

(5) When $\boldsymbol{\beta}_1 = -3.21,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.13. Table 4.12 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

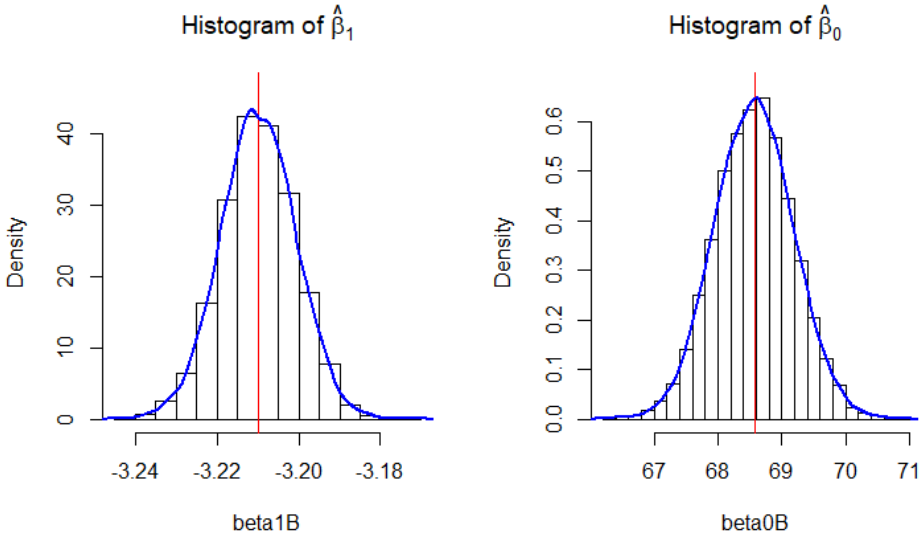

Figure 4.13: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7,\ (a, b) = (6.5, 9.25),\ \beta_1 = -3.21,\ \beta_0 = 68.57$

Table 4.12: *Summary of Simulation by Method I* : $\beta_1 = -3.21,\ \beta_0 = 68.57$ $\sigma_e = 7,\ (a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	68.553	8×10^{-5}	0.374	[-3.227, -3.192]	[67.304, 69.734]
5000	-3.210	68.573	9×10^{-5}	0.402	[-3.228, -3.192]	[67.322, 69.800]
10000	-3.210	68.570	8×10^{-5}	0.376	[-3.228, -3.192]	[67.387, 69.770]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.570, with the MSE to be 0.376.

(6) When $\boldsymbol{\beta}_1 = -3.21,\ \boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.14. Table 4.13 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

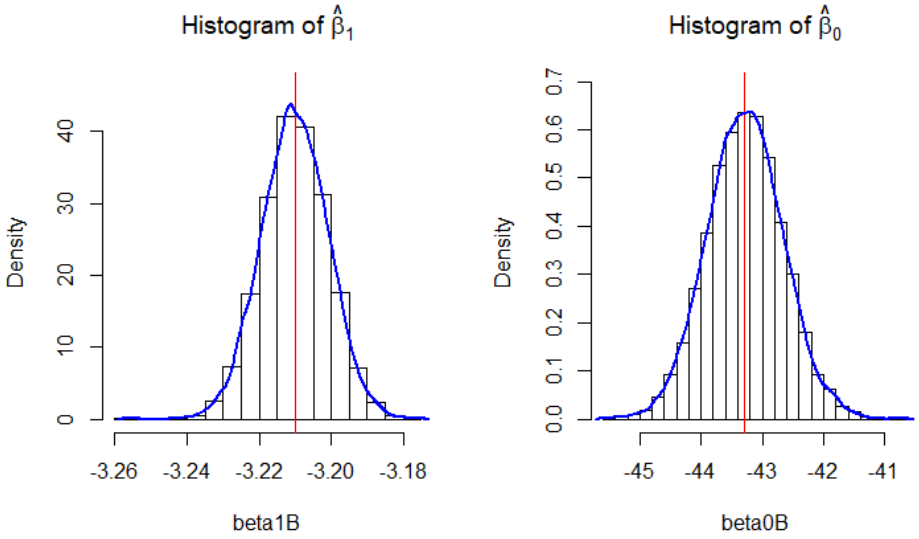

Figure 4.14: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7,\ (a, b) = (6.5, 9.25),\ \beta_1 = -3.21,\ \beta_0 = -43.29$

Table 4.13: *Summary of Simulation by Method I* : $\beta_1 = -3.21,\ \beta_0 = -43.29$ $\sigma_e = 7,\ (a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	-43.271	8×10^{-5}	0.402	[-3.228, -3.192]	[-44.515, -42.058]
5000	-3.210	-43.292	8×10^{-5}	0.370	[-3.228, -3.192]	[-44.501, -42.106]
10000	-3.210	-43.278	8×10^{-5}	0.379	[-3.228, -3.192]	[-44.472, -42.057]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.278, with the MSE to be 0.379.

Method 1 - <u>Set 3</u>: $\sigma_e = 10,\ (a, b) = (6.5,\ 9.25)$

(1) When $\boldsymbol{\beta}_1 = 0.64,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.15. Table 4.14 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

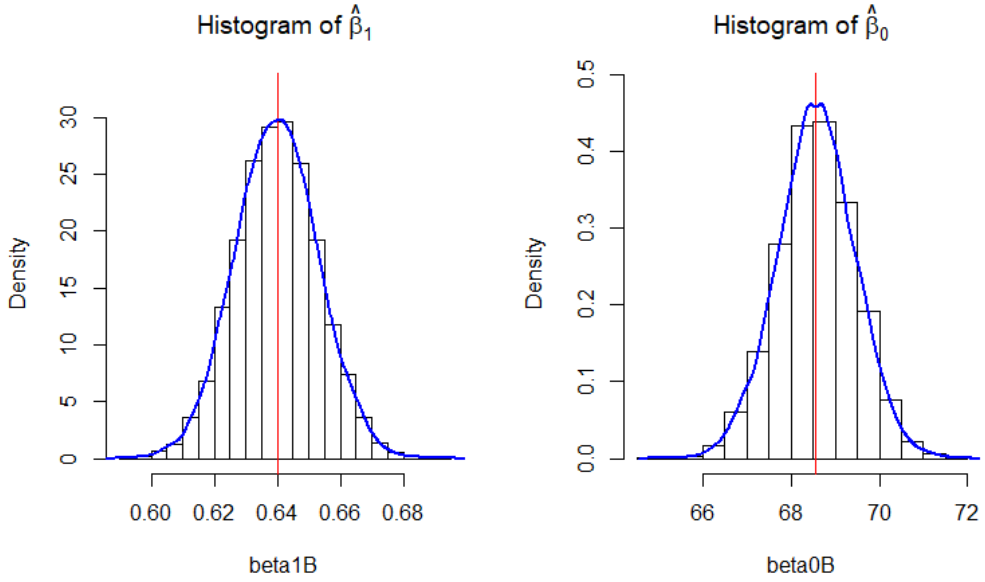

Figure 4.15: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$, $\beta_1 = 0.64$, $\beta_0 = 68.57$

Table 4.14: *Summary of Simulation by Method I* : $\beta_1 = 0.64$, $\beta_0 = 68.57$ $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	68.584	2×10^{-4}	0.770	[0.613, 0.667]	[66.864, 70.277]
5000	0.640	68.569	2×10^{-4}	0.785	[0.613, 0.665]	[66.823, 70.292]
10000	0.640	68.565	2×10^{-4}	0.771	[0.615, 0.666]	[66.835, 70.272]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 2×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.565, with the MSE to be 0.771.

(2) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.16. Table 4.15 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

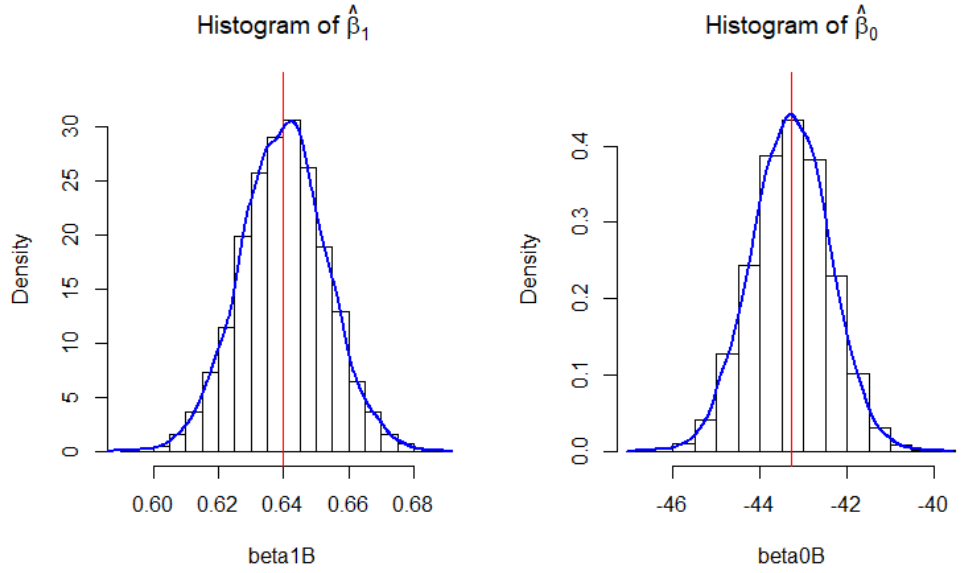

Figure 4.16: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$, $\beta_1 = 0.64$, $\beta_0 = -43.29$

Table 4.15: *Summary of Simulation by Method I* : $\beta_1 = 0.64$, $\beta_0 = -43.29$ $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	-43.297	2×10^{-4}	0.751	[0.615, 0.666]	[-44.928, -41.540]
5000	0.640	-43.289	2×10^{-4}	0.796	[0.614 , 0.666]	[-45.027, -41.494]
10000	0.640	-43.296	2×10^{-4}	0.790	[0.614, 0.666]	[-45.035, -41.583]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 2×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.296, with the MSE to be 0.790.

(3) When $\boldsymbol{\beta}_1 = 2.15$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.17. Table 4.16 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

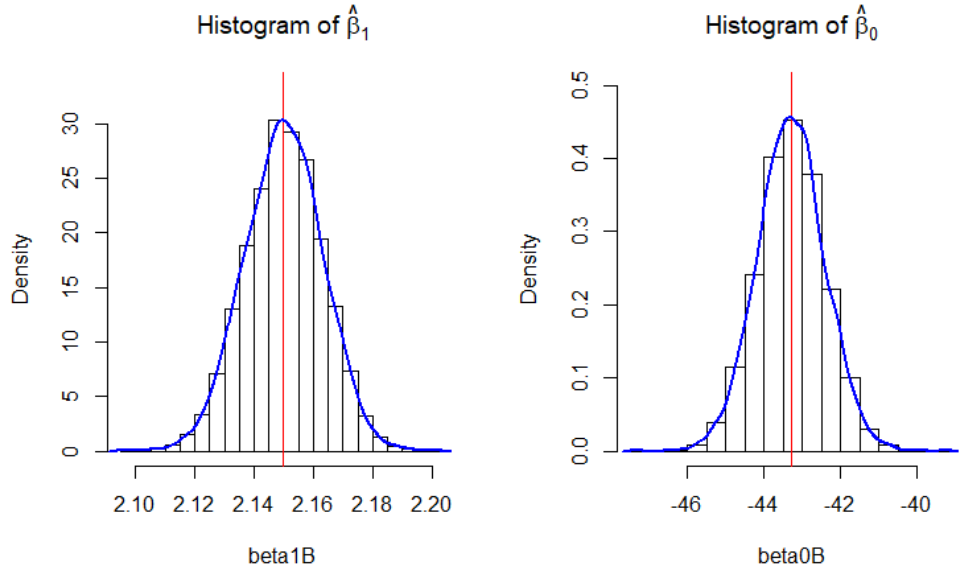

Figure 4.17: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$, $\beta_1 = 2.15$, $\beta_0 = -43.29$

Table 4.16: *Summary of Simulation by Method I* : $\beta_1 = 2.15$, $\beta_0 = -43.29$ $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	-43.314	2×10^{-4}	0.797	[2.124, 2.176]	[-44.993, -41.586]
5000	2.150	-43.296	2×10^{-4}	0.766	[2.124, 2.176]	[-45.005, -41.601]
10000	2.150	-43.289	2×10^{-4}	0.762	[2.124, 2.175]	[-44.998, -41.572]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 2×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.289, with the MSE to be 0.762.

(4) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.18. Table 4.17 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

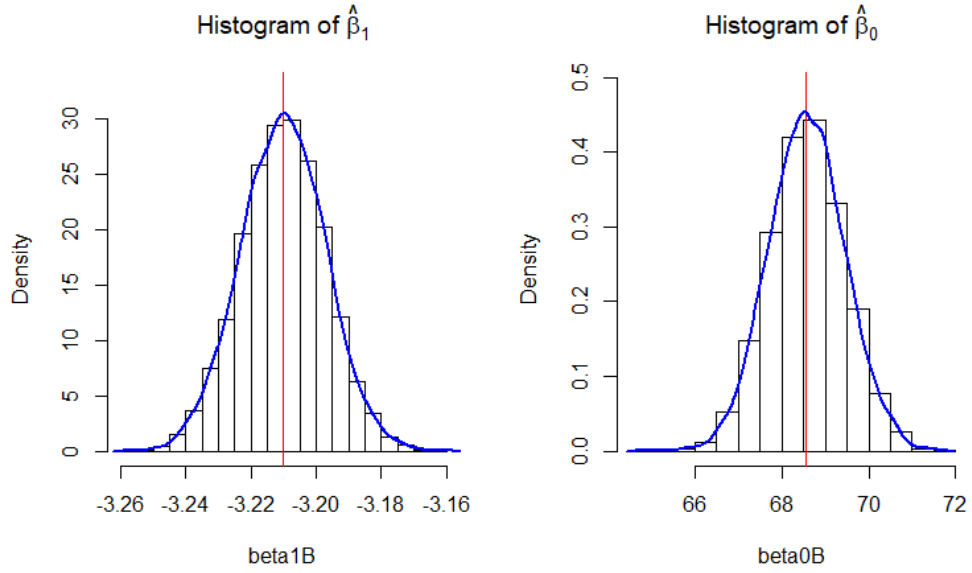

Figure 4.18: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$, $\beta_1 = -3.21$, $\beta_0 = 68.57$

Table 4.17: *Summary of Simulation by Method I* : $\beta_1 = -3.21$, $\beta_0 = 68.57$ $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	68.540	2×10^{-4}	0.745	[-3.234, -3.185]	[66.754, 70.230]
5000	-3.210	68.578	2×10^{-4}	0.762	[-3.236 , -3.184]	[66.870, 70.278]
10000	-3.210	68.581	2×10^{-4}	0.762	[-3.236, -3.185]	[66.880, 70.315]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 2×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.581, with the MSE to be 0.762.

(5) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.19. Table 4.18 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

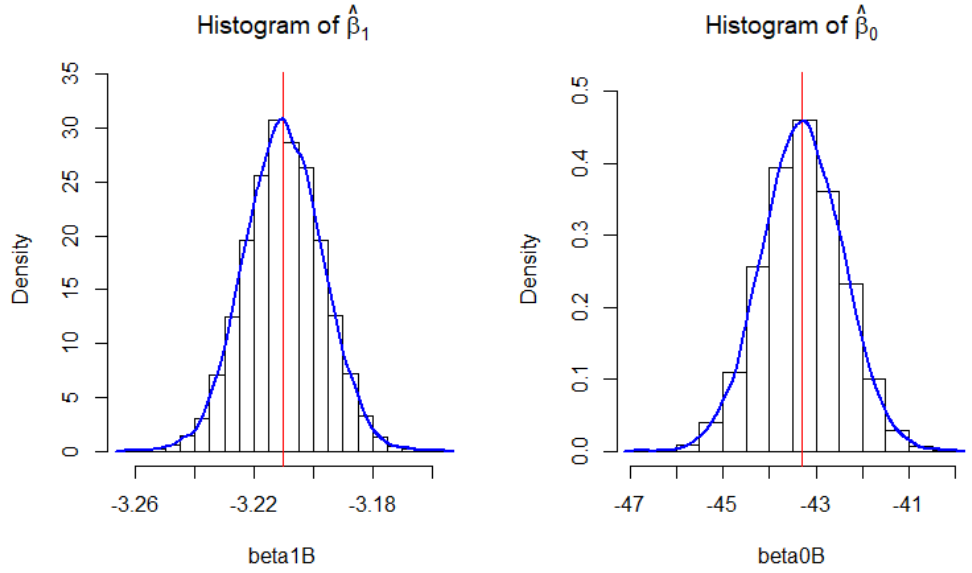

Figure 4.19: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$, $\beta_1 = -3.21$, $\beta_0 = -43.29$

Table 4.18: *Summary of Simulation by Method I* : $\beta_1 = -3.21$, $\beta_0 = -43.29$ $\sigma_e = 10$, $(a, b) = (6.5, 9.25)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	-43.298	2×10^{-4}	0.772	[-3.235, -3.184]	[-45.083, -41.597]
5000	-3.210	-43.301	2×10^{-4}	0.759	[-3.236 , -3.185]	[-44.993, -41.583]
10000	-3.210	-43.289	2×10^{-4}	0.747	[-3.235, -3.185]	[-44.992, -41.593]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 2×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.289, with the MSE to be 0.747.

Method 1 - Set 4: $\sigma_e = 3,\ (a, b) = (10,\ 12.45)$

(1) When $\boldsymbol{\beta}_1 = 0.64,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.20. Table 4.19 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

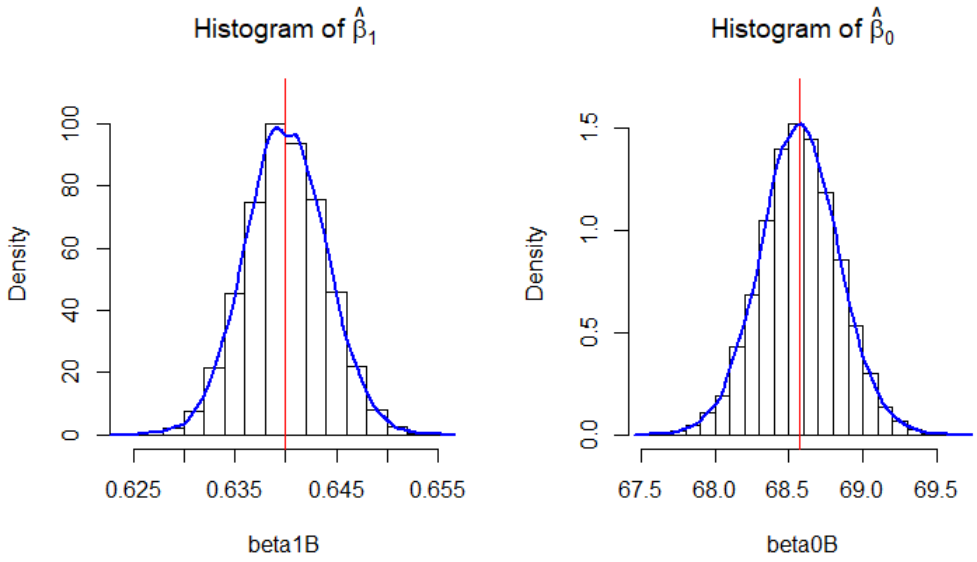

Figure 4.20: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3,\ (a, b) = (10, 12.45),\ \beta_1 = 0.64,\ \beta_0 = 68.57$

Table 4.19: *Summary of Simulation by Method I* : $\beta_1 = 0.64,\ \beta_0 = 68.57$ $\sigma_e = 3,\ (a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	68.576	2×10^{-5}	0.073	[0.632, 0.648]	[68.027, 69.104]
5000	0.640	68.573	1×10^{-5}	0.068	[0.632, 0.647]	[68.061, 69.099]
10000	0.640	68.571	2×10^{-5}	0.070	[0.632, 0.648]	[68.046, 69.089]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 2×10^{-5}; the

average value of $\hat{\beta}_0$ for 10000 repetitions is 68.571, with the MSE to be 0.070.

(2) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.21. Table 4.20 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

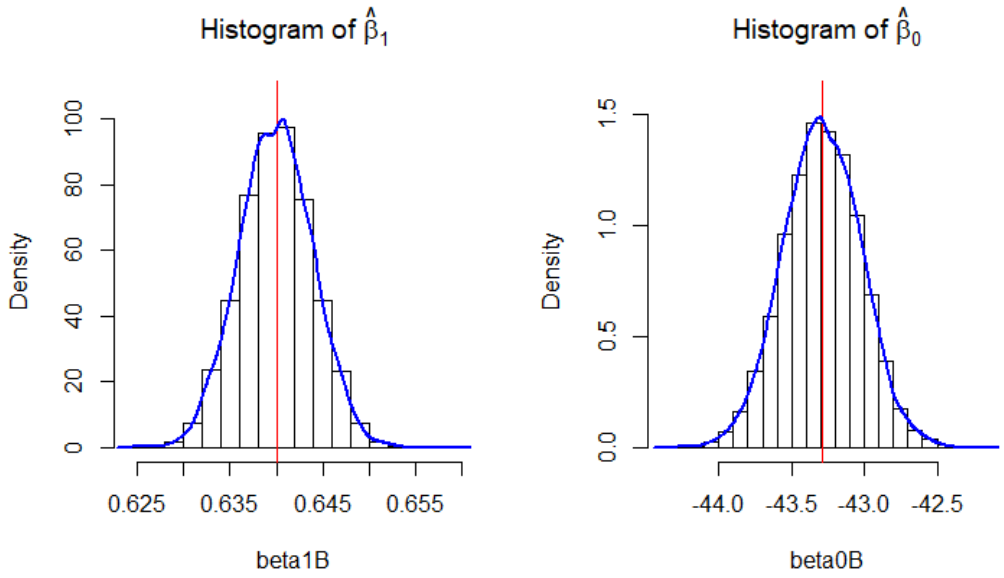

Figure 4.21: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3$, $(a, b) = (10, 12.45)$, $\beta_1 = 0.64$, $\beta_0 = -43.29$

Table 4.20: *Summary of Simulation by Method I* : $\beta_1 = 0.64$, $\beta_0 = -43.29$ $\sigma_e = 3$, $(a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	-43.301	2×10^{-5}	0.070	[0.632, 0.648]	[-43.823, -42.770]
5000	0.640	-43.294	2×10^{-5}	0.071	[0.632, 0.648]	[-43.803, -42.770]
10000	0.640	-43.292	2×10^{-5}	0.070	[0.632, 0.648]	[-43.807, -42.768]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.292, with the MSE to be 0.070.

(3) When $\boldsymbol{\beta}_1 = 2.15$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.22. Table 4.21 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

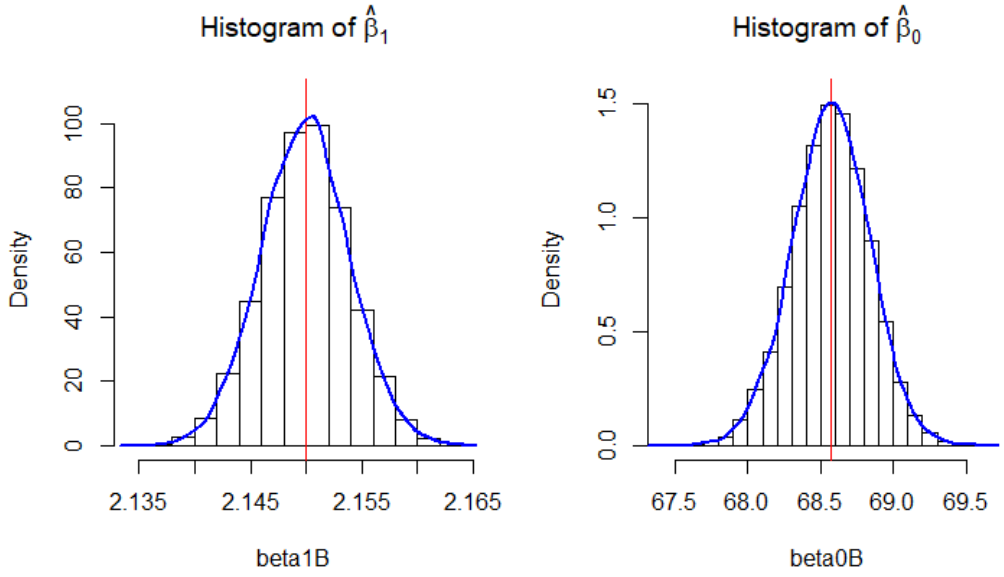

Figure 4.22: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3$, $(a, b) = (10, 12.45)$, $\beta_1 = 2.15$, $\beta_0 = 68.57$

Table 4.21: *Summary of Simulation by Method I* : $\beta_1 = 2.15,\ \beta_0 = 68.57$ $\sigma_e = 3,\ (a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	68.573	2×10^{-5}	0.071	[2.142, 2.158]	[68.058, 69.091]
5000	2.150	68.573	2×10^{-5}	0.070	[2.142, 2.158]	[68.045, 69.080]
10000	2.150	68.571	2×10^{-5}	0.068	[2.142, 2.158]	[68.053, 69.086]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.571, with the MSE to be 0.068.

(4) When $\boldsymbol{\beta}_1 = 2.15, \boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.23. Table 4.22 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

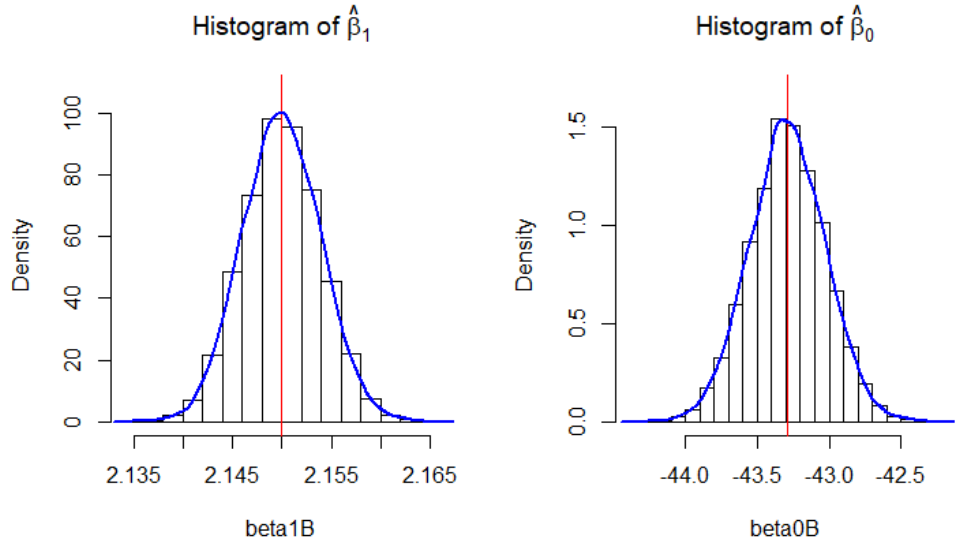

Figure 4.23: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3,\ (a, b) = (10, 12.45),\ \beta_1 = 2.15,\ \beta_0 = -43.29$

Table 4.22: *Summary of Simulation by Method I* : $\beta_1 = 2.15,\ \beta_0 = -43.29$ $\sigma_e = 3,\ (a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	-43.293	2×10^{-5}	0.067	[2.142, 2.158]	[-43.805, -42.799]
5000	2.150	-43.294	2×10^{-5}	0.069	[2.142, 2.158]	[-43.815, -42.787]
10000	2.150	-43.296	2×10^{-5}	0.069	[2.142, 2.158]	[-43.805, -42.782]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.296, with the MSE to be 0.069.

(5) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.24. Table 4.23 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

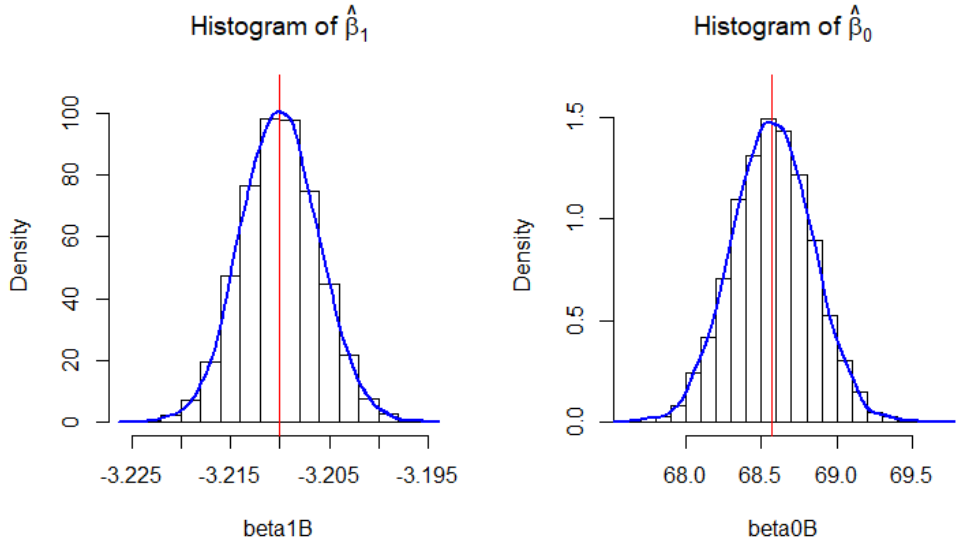

Figure 4.24: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3,\ (a, b) = (10, 12.45),\ \beta_1 = -3.21,\ \beta_0 = 68.57$

Table 4.23: *Summary of Simulation by Method I* : $\beta_1 = -3.21,\ \beta_0 = 68.57$ $\sigma_e = 3,\ (a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	68.580	1×10^{-5}	0.069	[-3.217, -3.202]	[68.064, 69.089]
5000	-3.210	68.569	2×10^{-5}	0.070	[-3.218, -3.202]	[68.057, 69.089]
10000	-3.210	68.569	2×10^{-5}	0.071	[-3.218, -3.202]	[68.050, 69.089]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.569, with the MSE to be 0.071.

(6) When $\boldsymbol{\beta}_1 = -3.21,\ \boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.25. Table 4.24 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

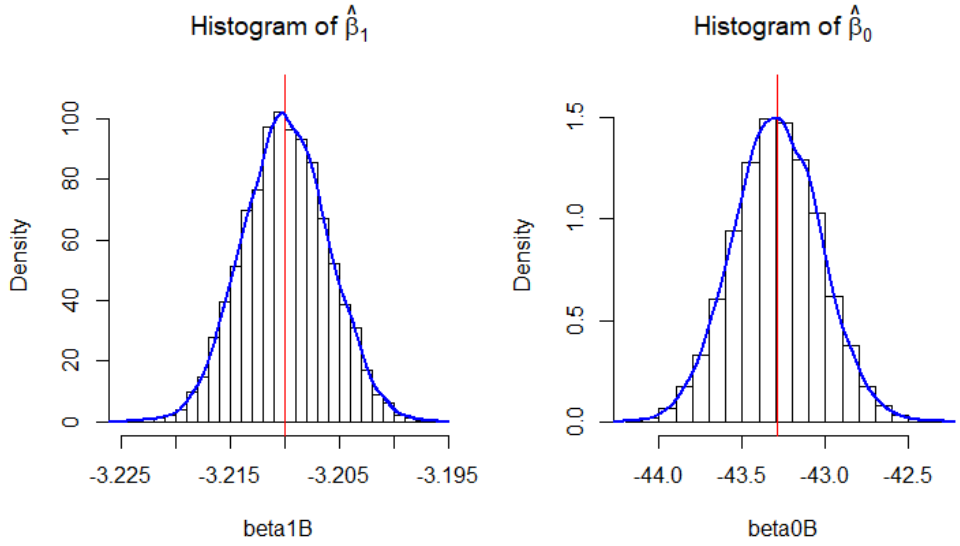

Figure 4.25: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3,\ (a, b) = (10, 12.45),\ \beta_1 = -3.21,\ \beta_0 = -43.29$

Table 4.24: *Summary of Simulation by Method I* : $\beta_1 = -3.21,\ \beta_0 = -43.29$ $\sigma_e = 3,\ (a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	-43.290	2×10^{-5}	0.068	[-3.218, -3.202]	[-43.821, -42.797]
5000	-3.210	-43.293	2×10^{-5}	0.370	[-3.218, -3.202]	[-43.800, -42.780]
10000	-3.210	-43.290	2×10^{-5}	0.069	[-3.218, -3.202]	[-43.804, -42.778]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.290, with the MSE to be 0.069.

Method 1 - <u>Set 5</u>: $\sigma_e = 7,\ (a, b) = (10,\ 12.45)$

(1) When $\boldsymbol{\beta}_1 = 0.64,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.26. Table 4.25 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

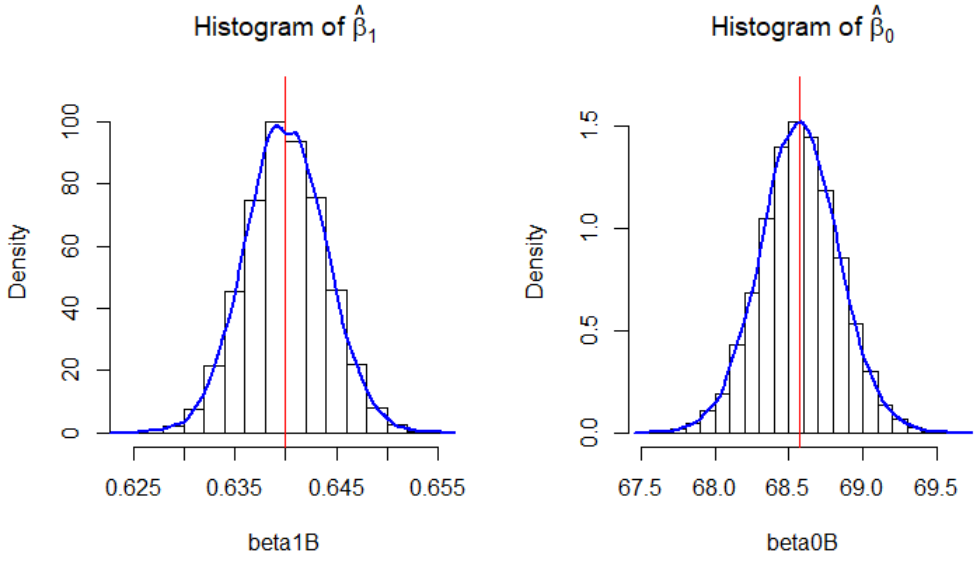

Figure 4.26: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $(a, b) = (10, 12.45)$, $\beta_1 = 0.64$, $\beta_0 = 68.57$

Table 4.25: *Summary of Simulation by Method I* : $\beta_1 = 0.64$, $\beta_0 = 68.57$ $\sigma_e = 7$, $(a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.641	68.556	9×10^{-5}	0.380	[0.622, 0.659]	[67.385, 69.763]
5000	0.640	68.561	8×10^{-5}	0.383	[0.622, 0.658]	[67.385, 69.824]
10000	0.640	68.580	8×10^{-5}	0.371	[0.622, 0.658]	[67.367, 69.777]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.580, with the MSE to be 0.371.

(2) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.27. Table 4.26 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

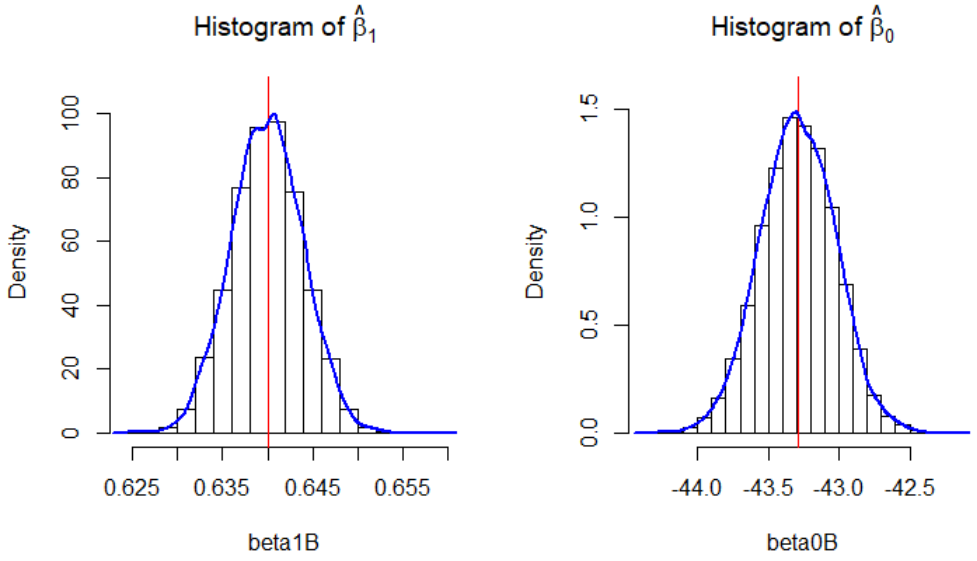

Figure 4.27: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $(a, b) = (10, 12.45)$, $\beta_1 = 0.64$, $\beta_0 = -43.29$

Table 4.26: *Summary of Simulation by Method I* : $\beta_1 = 0.64$, $\beta_0 = -43.29$ $\sigma_e = 7$, $(a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	-43.315	8×10^{-5}	0.378	[0.623, 0.659]	[-44.512, -42.117]
5000	0.640	-43.275	8×10^{-5}	0.381	[0.622, 0.657]	[-44.458, -42.029]
10000	0.640	-43.292	9×10^{-5}	0.384	[0.622, 0.658]	[-44.491, -42.075]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 9×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.292, with the MSE to be 0.384.

(3) When $\boldsymbol{\beta}_1 = 2.15,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.28. Table 4.27 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

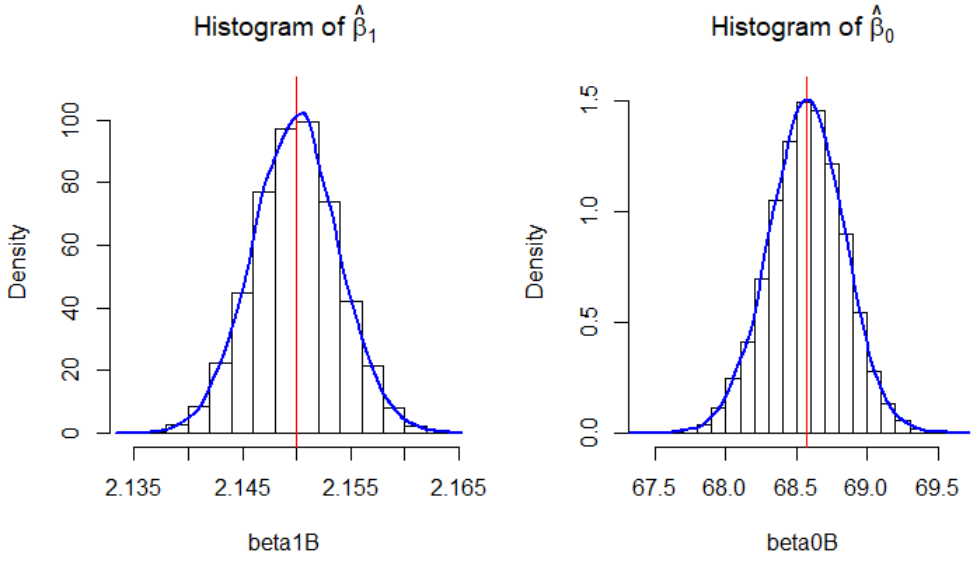

Figure 4.28: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7,\ (a, b) = (10, 12.45),\ \beta_1 = 2.15,\ \beta_0 = 68.57$

Table 4.27: *Summary of Simulation by Method I* : $\beta_1 = 2.15,\ \beta_0 = 68.57$ $\sigma_e = 7,\ (a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	68.590	9×10^{-5}	0.375	[2.132, 2.167]	[67.397, 69.774]
5000	2.150	68.573	8×10^{-5}	0.373	[2.132, 2.168]	[67.378, 69.752]
10000	2.150	68.561	8×10^{-5}	0.380	[2.132, 2.168]	[67.346, 69.745]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.561, with the MSE to be 0.380.

(4) When $\boldsymbol{\beta}_1 = 2.15$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.29. Table 4.28 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

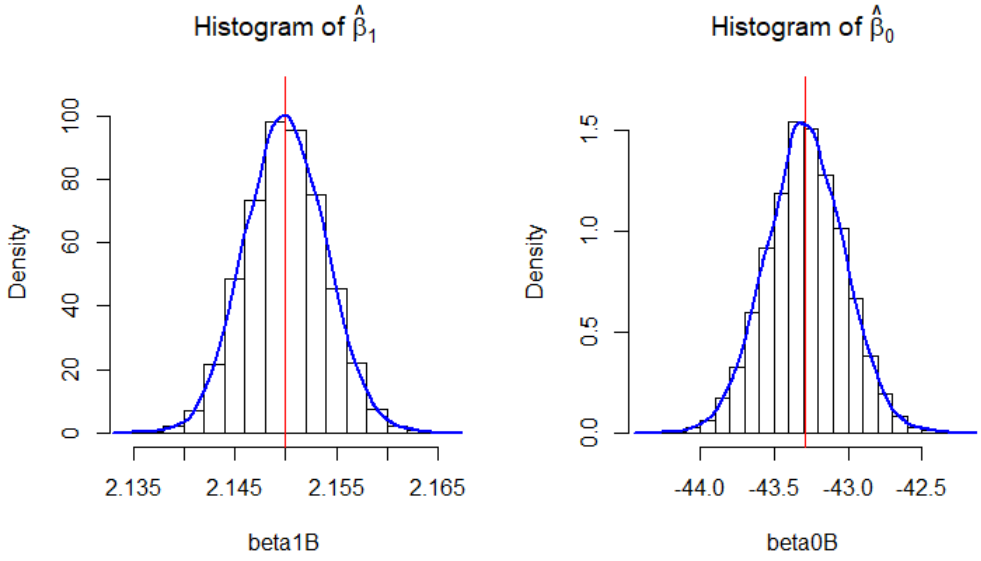

Figure 4.29: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $(a, b) = (10, 12.45)$, $\beta_1 = 2.15$, $\beta_0 = -43.29$

Table 4.28: *Summary of Simulation by Method I* : $\beta_1 = 2.15$, $\beta_0 = -43.29$ $\sigma_e = 7$, $(a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	-43.316	8×10^{-5}	0.374	[2.133, 2.168]	[-44.512, -42.169]
5000	2.150	-43.281	8×10^{-5}	0.378	[2.132, 2.168]	[-44.499, -42.087]
10000	2.150	-43.281	8×10^{-5}	0.377	[2.132, 2.168]	[-44.484, -42.072]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.281, with the MSE to be 0.377.

(5) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.30. Table 4.29 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

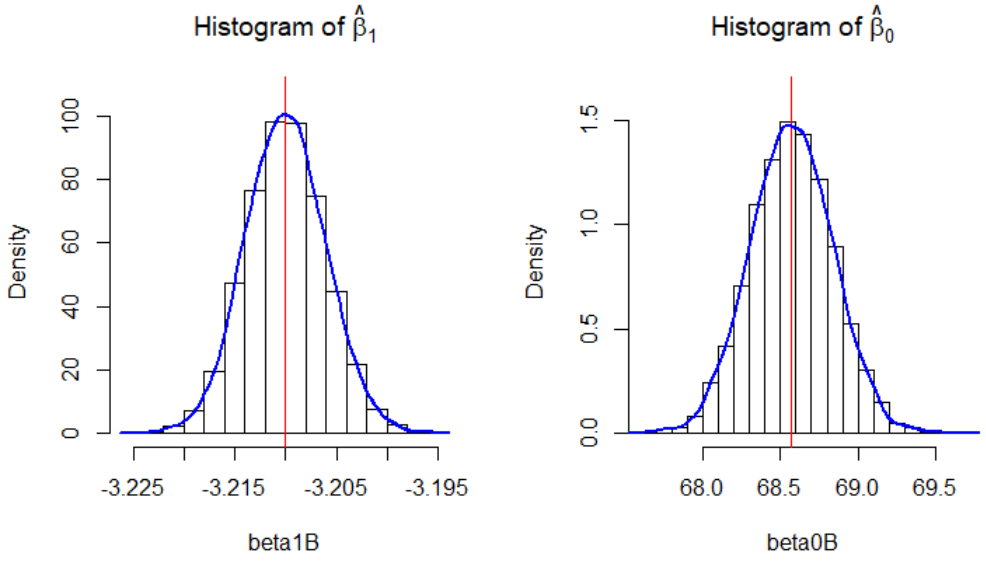

Figure 4.30: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $(a, b) = (10, 12.45)$, $\beta_1 = -3.21$, $\beta_0 = 68.57$

Table 4.29: *Summary of Simulation by Method I* : $\beta_1 = -3.21$, $\beta_0 = 68.57$ $\sigma_e = 7$, $(a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	68.567	8×10^{-5}	0.374	[-3.228, -3.192]	[67.384, 69.775]
5000	-3.210	68.573	9×10^{-5}	0.376	[-3.228, -3.192]	[67.352, 69.754]
10000	-3.210	68.567	9×10^{-5}	0.389	[-3.228, -3.192]	[67.371, 69.809]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 9×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.567, with the MSE to be 0.389.

(6) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.31. Table 4.30 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

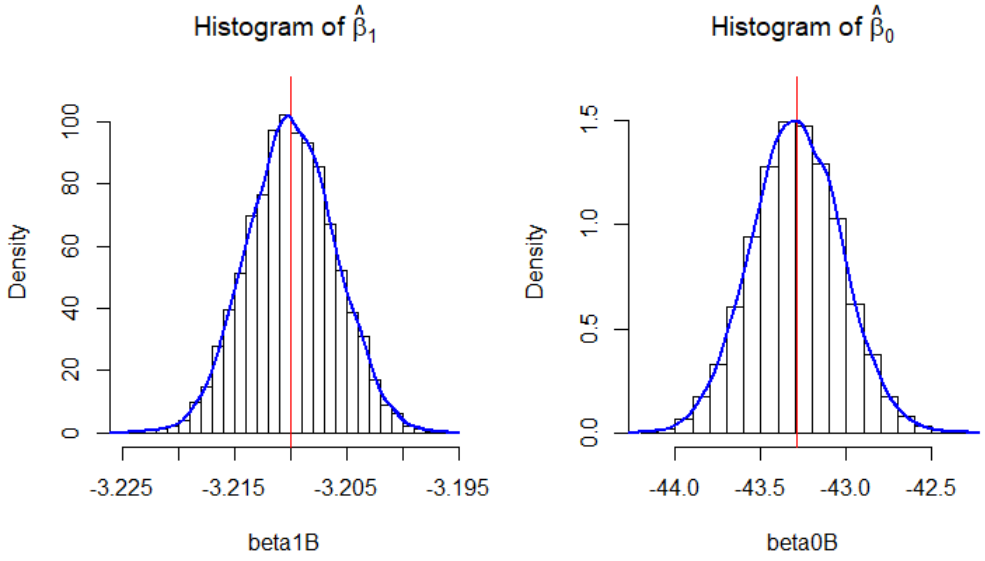

Figure 4.31: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $(a, b) = (10, 12.45)$, $\beta_1 = -3.21$, $\beta_0 = -43.29$

Table 4.30: *Summary of Simulation by Method I* : $\beta_1 = -3.21$, $\beta_0 = -43.29$ $\sigma_e = 7$, $(a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	-43.303	9×10^{-5}	0.388	[-3.228, -3.193]	[-44.512, -42.076]
5000	-3.210	-43.291	8×10^{-5}	0.376	[-3.228, -3.192]	[-44.475, -42.051]
10000	-3.210	-43.289	9×10^{-5}	0.387	[-3.228, -3.192]	[-44.526, -42.071]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 9×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.289, with the MSE to be 0.387.

Method 1 - Set 6: $\sigma_e = 10$, $(a, b) = (10,\ 12.45)$

(1) When $\boldsymbol{\beta}_1 = 0.64,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.32. Table 4.31 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

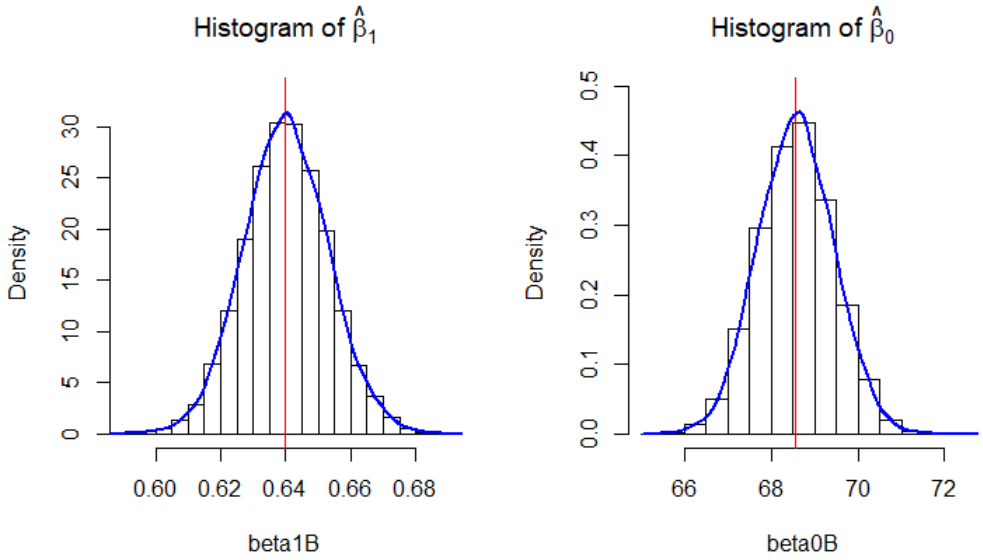

Figure 4.32: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $(a, b) = (10, 12.45)$, $\beta_1 = 0.64$, $\beta_0 = 68.57$

Table 4.31: *Summary of Simulation by Method I* : $\beta_1 = 0.64,\ \beta_0 = 68.57$ $\sigma_e = 10$, $(a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.639	68.597	1.7×10^{-4}	0.705	[0.614, 0.665]	[66.961, 70.170]
5000	0.640	68.558	1.7×10^{-4}	0.775	[0.615, 0.665]	[66.814, 70.315]
10000	0.640	68.580	1.7×10^{-4}	0.767	[0.615, 0.665]	[66.869, 70.272]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.580, with the MSE to be 0.767.

(2) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.33. Table 4.32 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

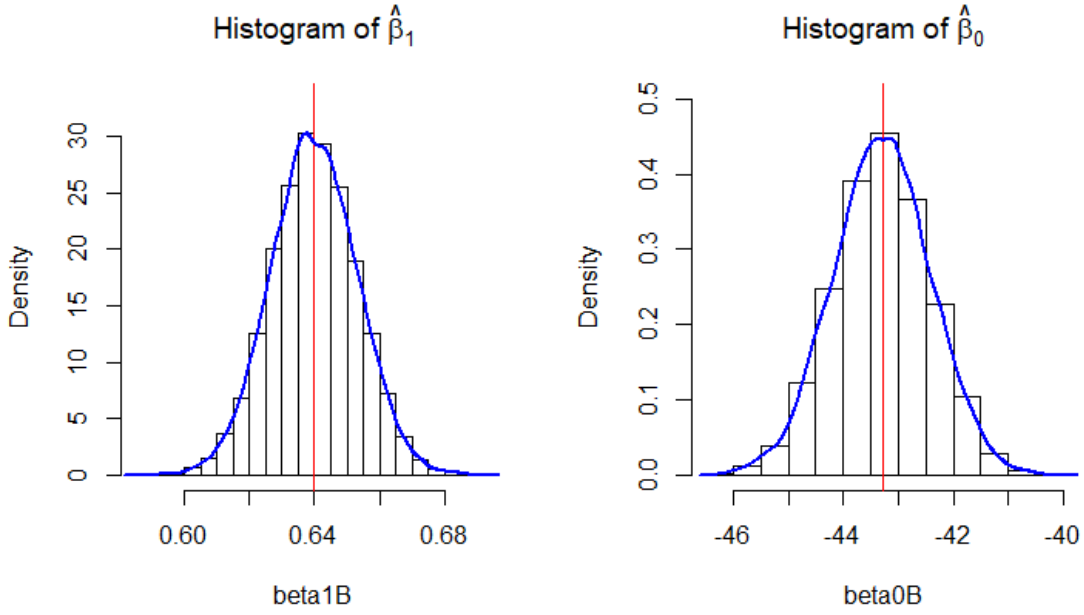

Figure 4.33: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $(a, b) = (10, 12.45)$, $\beta_1 = 0.64$, $\beta_0 = -43.29$

Table 4.32: *Summary of Simulation by Method I* : $\beta_1 = 0.64,\ \beta_0 = -43.29$ $\sigma_e = 10,\ (a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	-43.278	1.6×10^{-4}	0.754	[0.615, 0.665]	[-43.018, -41.601]
5000	0.640	-43.303	1.7×10^{-4}	0.778	[0.615, 0.665]	[-45.022, -41.539]
10000	0.640	-43.296	1.7×10^{-4}	0.752	[0.615, 0.666]	[-45.000, -41.605]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.296, with the MSE to be 0.752.

(3) When $\boldsymbol{\beta}_1 = 2.15,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.34. Table 4.33 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

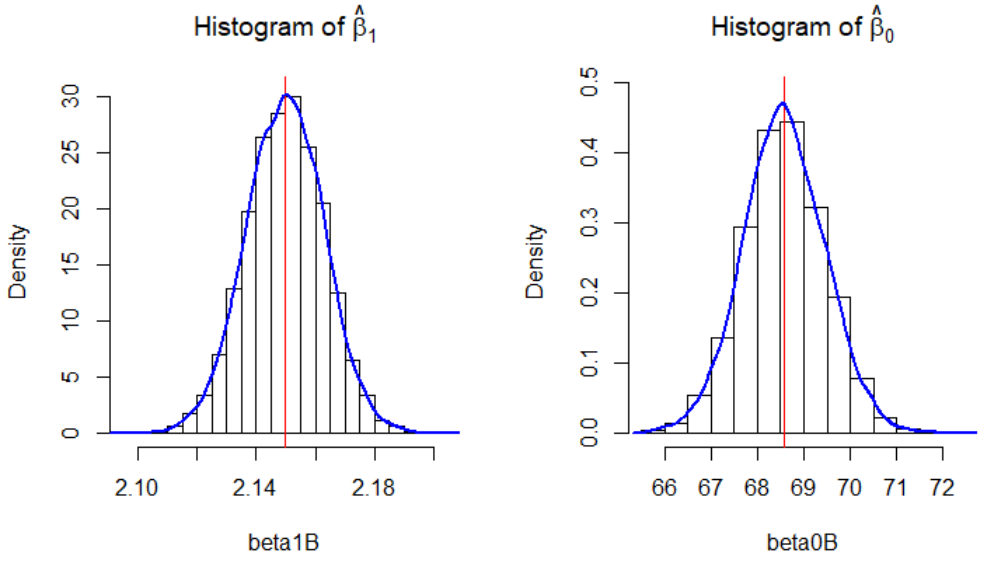

Figure 4.34: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $(a, b) = (10, 12.45)$, $\beta_1 = 2.15$, $\beta_0 = 68.57$

Table 4.33: *Summary of Simulation by Method I* : $\beta_1 = 2.15$, $\beta_0 = 68.57$ $\sigma_e = 10$, $(a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	68.576	1.7×10^{-4}	0.776	[2.124, 2.176]	[66.913, 70.361]
5000	2.150	68.576	1.7×10^{-4}	0.769	[2.125, 2.176]	[66.886, 70.298]
10000	2.150	68.569	1.7×10^{-4}	0.768	[2.125, 2.175]	[66.849, 70.271]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.569, with the MSE to be 0.768.

(4) When $\boldsymbol{\beta}_1 = 2.15$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.35. Table 4.34 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

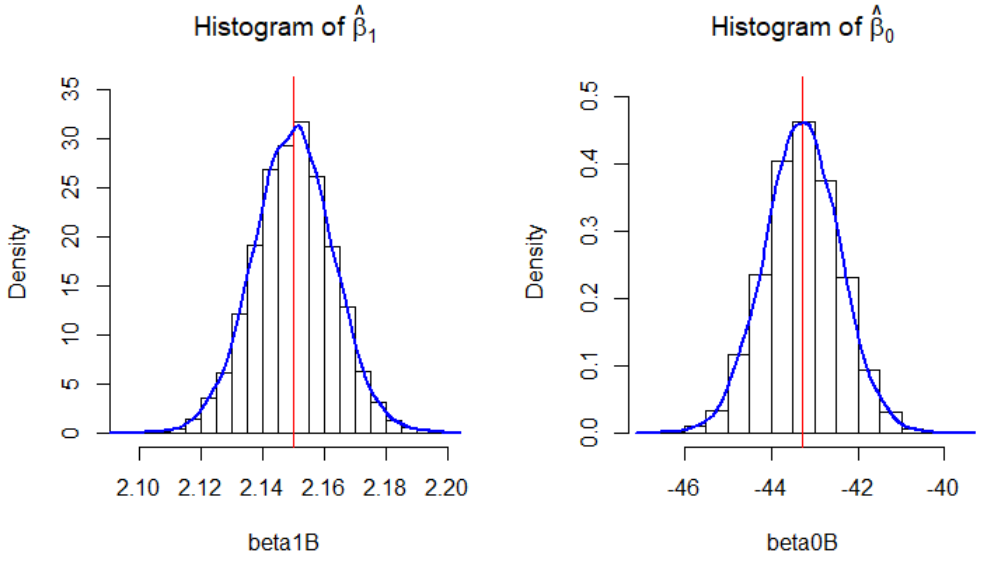

Figure 4.35: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $(a, b) = (10, 12.45)$, $\beta_1 = 2.15$, $\beta_0 = -43.29$

Table 4.34: *Summary of Simulation by Method I* : $\beta_1 = 2.15$, $\beta_0 = -43.29$ $\sigma_e = 10$, $(a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	-43.269	1.6×10^{-4}	0.716	[2.125, 2.174]	[-44.859, -41.601]
5000	2.150	-43.300	1.7×10^{-4}	0.768	[2.124, 2.176]	[-44.987, -41.579]
10000	2.150	-43.294	1.7×10^{-4}	0.755	[2.124, 2.176]	[-44.992, -41.590]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.294, with the MSE to be 0.755.

(5) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.36. Table 4.35 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

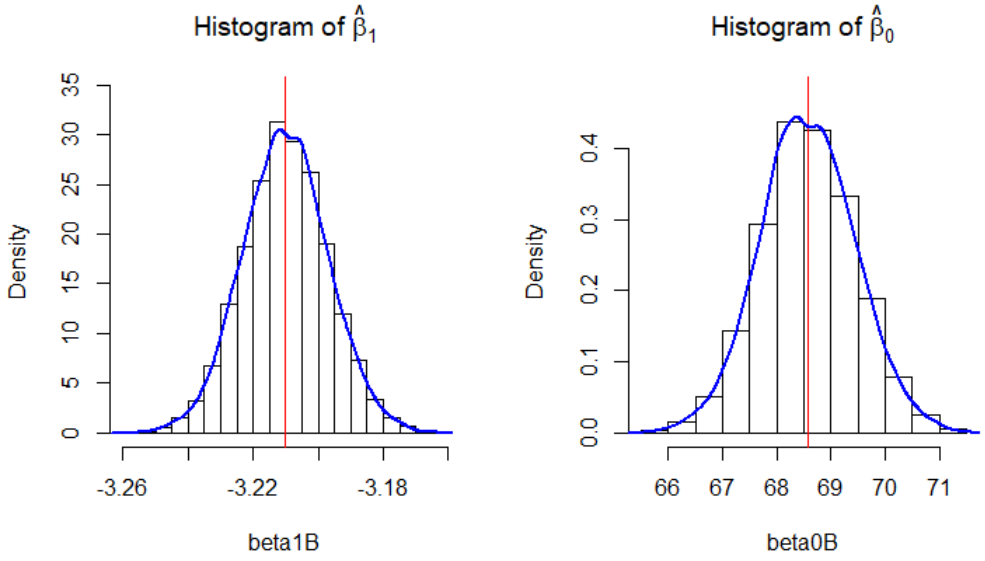

Figure 4.36: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $(a, b) = (10, 12.45)$, $\beta_1 = -3.21$, $\beta_0 = 68.57$

Table 4.35: *Summary of Simulation by Method I* : $\beta_1 = -3.21$, $\beta_0 = 68.57$ $\sigma_e = 10$, $(a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	68.547	1.8×10^{-4}	0.775	[-3.237, -3.185]	[66.833, 70.235]
5000	-3.210	68.543	1.7×10^{-4}	0.767	[-3.235, -3.184]	[66.860, 70.237]
10000	-3.210	68.586	1.7×10^{-4}	0.788	[-3.236, -3.184]	[66.855, 70.337]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.586, with the MSE to be 0.788.

(6) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.37. Table 4.36 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

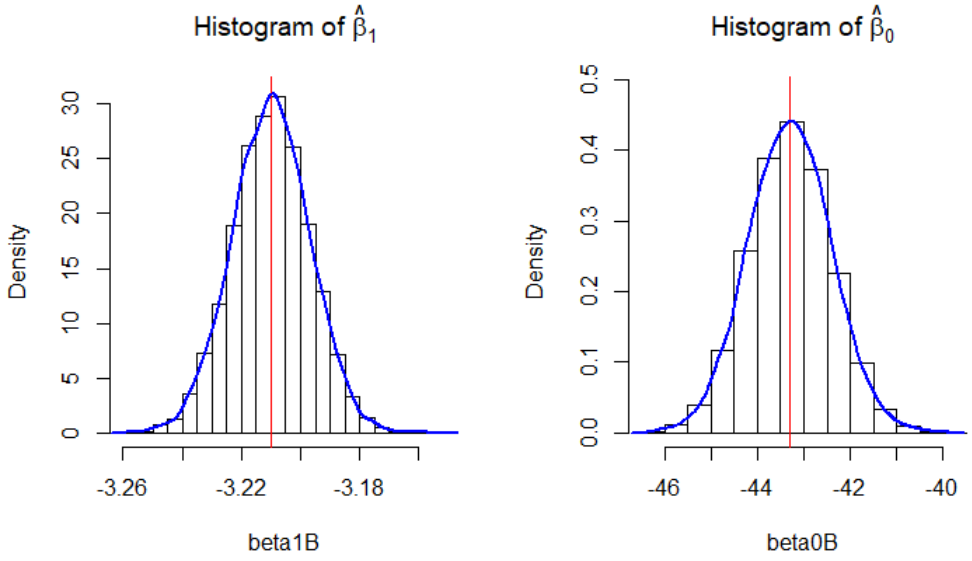

Figure 4.37: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $(a, b) = (10, 12.45)$, $\beta_1 = -3.21$, $\beta_0 = -43.29$

Table 4.36: *Summary of Simulation by Method I* : $\beta_1 = -3.21$, $\beta_0 = -43.29$ $\sigma_e = 10$, $(a, b) = (10, 12.45)$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	-43.291	1.7×10^{-4}	0.783	[-3.236, -3.184]	[-45.093, -41.590]
5000	-3.210	-43.292	1.7×10^{-4}	0.784	[-3.236, -3.185]	[-45.060, -41.578]
10000	-3.210	-43.290	1.7×10^{-4}	0.764	[-3.235, -3.185]	[-44.992, -41.584]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.290, with the MSE to be 0.764.

To compare the performances of the proposed method for different settings of the standard deviation of the error term, σ_e, and the lower and upper bounds of the uniform distribution from which the interval ranges $X^{(r)}$ are generated, (a, b), we create the six tables as follows to summarize the averages of $\hat{\beta}_1$ and $\hat{\beta}_0$ as well as $MSE(\hat{\beta}_1)$ and $MSE(\hat{\beta}_0)$. Each table corresponds to a setting of the slope and the intercept parameter values (β_1, β_0), and with 10000 repetitions.

I. When $(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_0) = (0.64,\ 68.57)$, the results are summarized in Table 4.37.

Table 4.37: $\hat{\beta}_1$, $\hat{\beta}_0$, $MSE(\hat{\beta}_1)$, $MSE(\hat{\beta}_0)$ for $(\beta_1,\ \beta_0) = (0.64,\ 68.57)$

σ_e	(a, b)	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$
3	(6.5, 9.25)	0.640	68.565	2×10^{-5}	0.070
	(10, 12.45)	0.640	68.571	2×10^{-5}	0.070
7	(6.5, 9.25)	0.640	68.566	8×10^{-5}	0.370
	(10, 12.45)	0.640	68.580	8×10^{-5}	0.371
10	(6.5, 9.25)	0.640	68.565	2×10^{-4}	0.771
	(10, 12.45)	0.640	68.580	1.7×10^{-4}	0.767

II. When $(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_0) = (0.64,\ -43.29)$, the results are summarized in Table 4.38.

Table 4.38: $\hat{\beta}_1$, $\hat{\beta}_0$, $MSE(\hat{\beta}_1)$, $MSE(\hat{\beta}_0)$ for $(\beta_1,\ \beta_0) = (0.64,\ -43.29)$

σ_e	(a, b)	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$
3	(6.5, 9.25)	0.640	-43.294	2×10^{-5}	0.071
	(10, 12.45)	0.640	-43.292	2×10^{-5}	0.070
7	(6.5, 9.25)	0.640	-43.303	8×10^{-5}	0.382
	(10, 12.45)	0.640	-43.292	9×10^{-5}	0.384
10	(6.5, 9.25)	0.640	-43.296	2×10^{-4}	0.790
	(10, 12.45)	0.640	-43.296	1.7×10^{-4}	0.752

III. When $(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_0) = (2.15,\ 68.57)$, the results are summarized in Table 4.39.

Table 4.39: $\hat{\beta}_1$, $\hat{\beta}_0$, $MSE(\hat{\beta}_1)$, $MSE(\hat{\beta}_0)$ for $(\beta_1,\ \beta_0) = (2.15,\ 68.57)$

σ_e	(a, b)	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$
3	(6.5, 9.25)	2.150	68.573	2×10^{-5}	0.068
	(10, 12.45)	2.150	68.571	2×10^{-5}	0.068
7	(6.5, 9.25)	2.150	68.569	8×10^{-5}	0.375
	(10, 12.45)	2.150	68.561	8×10^{-5}	0.380
10	(6.5, 9.25)	2.150	68.586	2×10^{-4}	0.756
	(10, 12.45)	2.150	68.569	1.7×10^{-4}	0.768

IV. When $(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_0) = (2.15,\ -43.29)$, the results are summarized in Table 4.40.

Table 4.40: $\hat{\beta}_1$, $\hat{\beta}_0$, $MSE(\hat{\beta}_1)$, $MSE(\hat{\beta}_0)$ for $(\beta_1,\ \beta_0) = (2.15,\ -43.29)$

σ_e	(a, b)	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$
3	(6.5, 9.25)	2.150	-43.288	2×10^{-5}	0.071
	(10, 12.45)	2.150	-43.296	2×10^{-5}	0.069
7	(6.5, 9.25)	2.150	-43.288	8×10^{-5}	0.375
	(10, 12.45)	2.150	-43.281	8×10^{-5}	0.377
10	(6.5, 9.25)	2.150	-43.289	2×10^{-4}	0.762
	(10, 12.45)	2.150	-43.294	1.7×10^{-4}	0.755

V. When $(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_0) = (-3.21, 68.57)$, the results are summarized in Table 4.41.

Table 4.41: $\hat{\beta}_1$, $\hat{\beta}_0$, $MSE(\hat{\beta}_1)$, $MSE(\hat{\beta}_0)$ for $(\beta_1,\ \beta_0) = (-3.21,\ 68.57)$

σ_e	(a, b)	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$
3	(6.5, 9.25)	-3.210	68.571	2×10^{-5}	0.069
	(10, 12.45)	-3.210	68.569	2×10^{-5}	0.071
7	(6.5, 9.25)	-3.210	68.570	8×10^{-5}	0.376
	(10, 12.45)	-3.210	68.567	9×10^{-5}	0.389
10	(6.5, 9.25)	-3.210	68.581	2×10^{-4}	0.762
	(10, 12.45)	-3.210	68.586	1.7×10^{-4}	0.788

VI. When $(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_0) = (-3.21,\ -43.29)$, the results are summarized in Table 4.42.

Table 4.42: $\hat{\beta}_1$, $\hat{\beta}_0$, $MSE(\hat{\beta}_1)$, $MSE(\hat{\beta}_0)$ for $(\beta_1,\ \beta_0) = (-3.21,\ -43.29)$

σ_e	(a, b)	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$
3	(6.5, 9.25)	-3.210	-43.289	2×10^{-5}	0.069
	(10, 12.45)	-3.210	-43.290	2×10^{-5}	0.069
7	(6.5, 9.25)	-3.210	-43.278	8×10^{-5}	0.379
	(10, 12.45)	-3.210	-43.289	9×10^{-5}	0.387
10	(6.5, 9.25)	-3.210	-43.289	2×10^{-4}	0.747
	(10, 12.45)	-3.210	-43.290	1.7×10^{-4}	0.764

From Table 4.37 to Table 4.42, we can observe that

1) The average value of $\hat{\beta}_1$ is always equal to the true value of β_1, and the average value of $\hat{\beta}_0$ is consistently very close to the true value of β_0, with each of the settings for β_1, β_0, σ_e and (a, b). This indicates that the proposed method gives unbiased estimators for both the slope and the intercept parameters.

2) Under the same setting of (β_1, β_0) and (a, b), the MSEs of both $\hat{\beta}_1$ and $\hat{\beta}_0$ become larger as the value of σ_e is set larger.

3) Under the same setting of (β_1, β_0) and σ_e, the larger values set for (a, b) do not result in big differences in $MSE(\hat{\beta}_1)$ or $MSE(\hat{\beta}_0)$.

4) Under the same settings for σ_e and (a, b), the MSEs of $\hat{\beta}_1$ are almost the same to each other, no matter of whether the slope or the intercept parameters are set to be positive or negative; so do the MSEs of $\hat{\beta}_0$.

In conclusion, the proposed approach performs well in estimating the regression coefficients on the data sets simulated by the first method.

Simulation: Method II

Now we study the performance of the proposed approach by simulations conducted by the second method. The settings for simulations are given as follows:

1. μ: the values of the means of the normal distributions from which the interval means $X^{(c)}$ are generated, -35, -25, -15, -5, 5, 15, 25, 35, 45, 55, 65, 75, 85, 95, 105, 115, 125;

2. σ: the standard deviation of the normal distributions from which the interval means $X^{(c)}$ are generated; this is set to be equal to 7;

3. n_0: the numbers of observations with the interval means at each of the seventeen values in Step **1**. The n_0 is randomly sampled from the discrete uniform distribution $Uniform(5, 9)$ with support $k \in \{5, 6, 7, 8, 9\}$;

4. σ_e: the standard deviation of the error term to be generated, 3, 7, and 10;

5. (a, b, μ_0): the lower bound, the upper bound and the value of the mean of the truncated normal distribution from which the interval ranges $X^{(r)}$ are generated, (9.43, 13.69, 11.12);

6. σ_0: the standard deviation of the truncated normal distribution from which the interval ranges $X^{(r)}$ are generated, 5.25 and 8.07;

7. β_1: the true slope parameters, 0.64, 2.15, -3.21;

8. β_0: the true intercept parameters, 68.57 and -43.29;

9. m: the number of values drawn from the uniform distribution $U(x_{Li}, x_{Ui})$, for $i = 1, ..., n$, and n is the total number of observations. Set m to be 3000;

10. B repetition times of drawing samples: 2000, 5000, and 10000.

Simulations based on the settings delineated here are conducted, and we apply the proposed method to give point estimators and confidence intervals for the slope parameter β_1 and the intercept parameter β_0, respectively.

To illustrate, we take the setting $\sigma_e = 10$, $\sigma_0 = 5.25$, $\beta_1 = 2.15$ and $\beta_0 = 68.57$ as an example. Figure 4.38 presents the scatter plot with the fitted regression line given by the average values of $\hat{\beta}_1$ and $\hat{\beta}_0$; and Figure 4.39 shows the histograms of the observations along the point estimators $\hat{\beta}_1$ and $\hat{\beta}_0$. Table 4.43 summarizes the MSEs as well as 95% confidence intervals. The results are based on 10000 repetitions of sampling.

EXAMPLE II: When we set simulations for parameter values $\sigma_e = 10$, $\sigma_0 = 5.25$, $\beta_1 = 2.15$, $\beta_0 = 68.57$, we obtain the data shown in Figure 4.38. The histogram plots of the resulting estimates for the slope parameter ($\hat{\beta}_1$) and for the intercept parameter ($\hat{\beta}_0$) are shown in Figure 4.39.

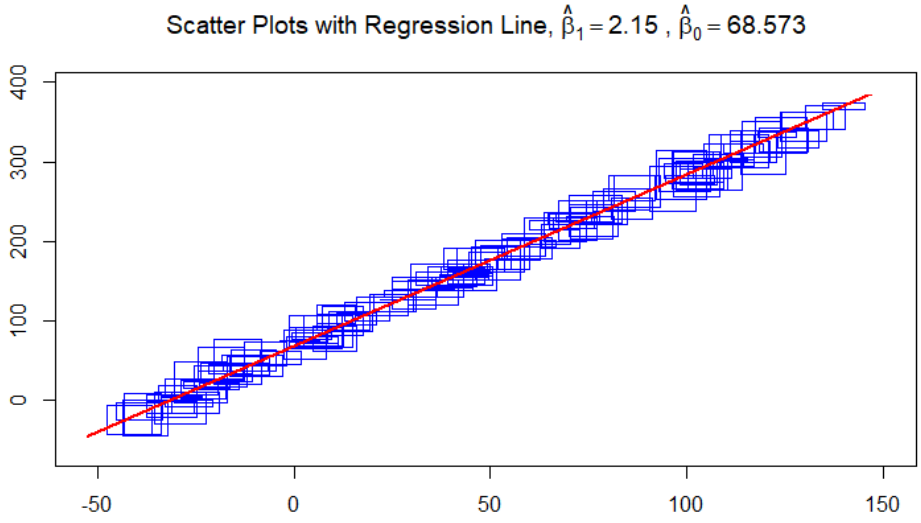

Figure 4.38: Scatter plot - $\sigma_e = 10$, $\sigma_0 = 5.25$, $\beta_1 = 2.15$, $\beta_0 = 68.57$

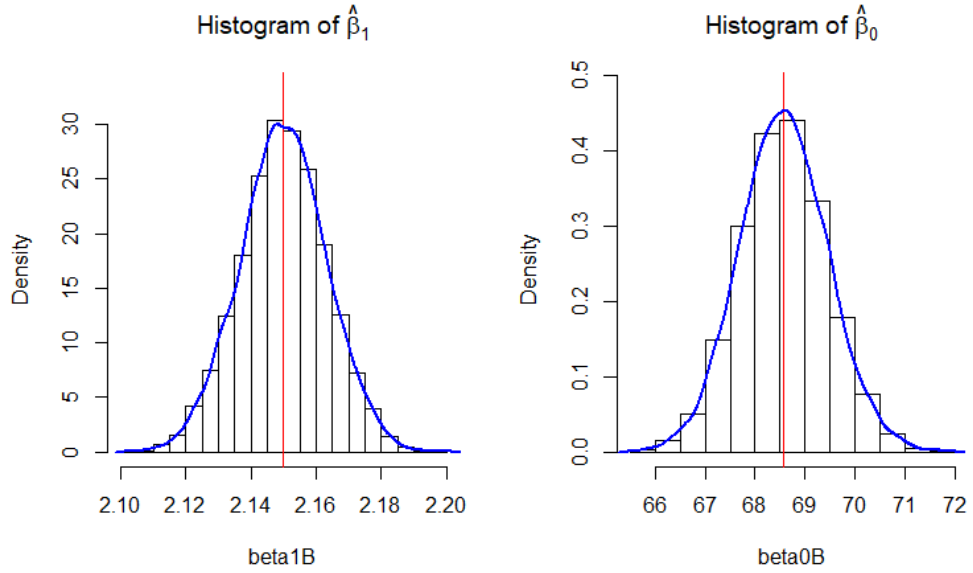

Figure 4.39: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 10$, $\sigma_0 = 5.25$, $\beta_1 = 2.15$, $\beta_0 = 68.57$

In Figure 4.39, the values of $\hat{\beta}_1$ and $\hat{\beta}_0$ obtained by different repetitions are indicated on the $x-$axis by beta1B and beta0B, respectively. The two vertical red lines on Figure

4.39 display the positions of the true parameter values, i.e., $\beta_1 = 2.15$ and $\beta_0 = 68.57$. The following table summarizes the simulation results for the setting in EXAMPLE II for the different numbers of repetitions, $B = (2000, 5000, 10000)$.

Table 4.43: *Summary of Simulation by Method II* : $\beta_1 = 2.15,\ \beta_0 = 68.57$
$\sigma_e = 10,\ \sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	68.576	2×10^{-4}	0.780	[2.125, 2.175]	[66.858, 70.371]
5000	2.150	68.578	2×10^{-4}	0.755	[2.125 , 2.176]	[66.881, 70.263]
10000	2.150	68.560	2×10^{-4}	0.763	[2.124, 2.176]	[66.844, 70.276]

From Table 4.43, similar to what we have observed in the simulation results in Section 4.2.1, under simulation method I, the average of the point estimators for both β_1 and β_0 are also equal to or quite close to the true values, with small values of MSEs, indicating the proposed approach gives accurate estimations for the regression coefficients, especially for β_1 with different repetitions from 2000 to 10000; by the histograms in Figure 4.39, we also see that both of $\hat{\beta}_1$ and $\hat{\beta}_0$ are distributed with shapes of normality, which verifies the normal property shown in (3.72) and (3.77). The 95% confidence intervals of both $\hat{\beta}_1$ and $\hat{\beta}_0$ given in Table 4.43 cover the true values of β_1 and β_0 and have almost the same length from the lower bound to the true value, and from the upper bound to the true value.

We now will use different sets of values for (β_1, β_0). There are six sets of values for the error standard deviation σ_e, and the standard deviation of the truncated normal distributions from which the interval ranges $X^{(r)}$ are generated, σ_0. For each of these sets, there are six different pairs of values for the regression parameters β_1 and β_0. The simulation results are provided along the same lines above as illustrated in Figure 4.39 and Table 4.43 for EXAMPLE II. These are briefly described as follows. Then, for each (β_1, β_0) pairing,

comparisons of these results are discussed and presented in Tables 4.79 - 4.84.

Method 2 - <u>Set 1</u>: $\sigma_e = 3$, $\sigma_0 = 5.25$

(1) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.40. Table 4.44 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

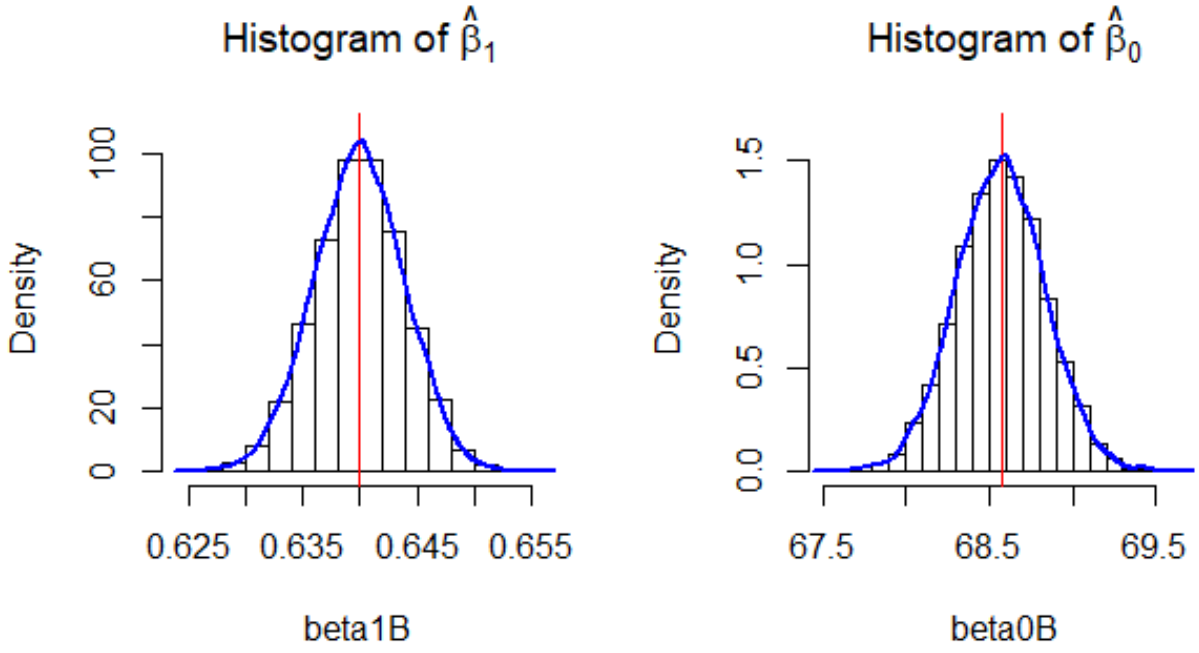

Figure 4.40: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 3$, $\sigma_0 = 5.25$, $\beta_1 = 0.64$, $\beta_0 = 68.57$

Table 4.44: *Summary of Simulation by Method II* : $\beta_1 = 0.64,\ \beta_0 = 68.57$ $\sigma_e = 3,\ \sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	68.586	2×10^{-5}	0.068	[0.632, 0.647]	[68.084, 69.091]
5000	0.640	68.571	2×10^{-5}	0.067	[0.632 , 0.648]	[68.066, 69.072]
10000	0.640	68.570	2×10^{-5}	0.070	[0.632, 0.647]	[68.047, 69.089]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.570, with the MSE to be 0.070.

(2) When $\boldsymbol{\beta}_1 = 0.64,\ \boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.41. Table 4.45 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

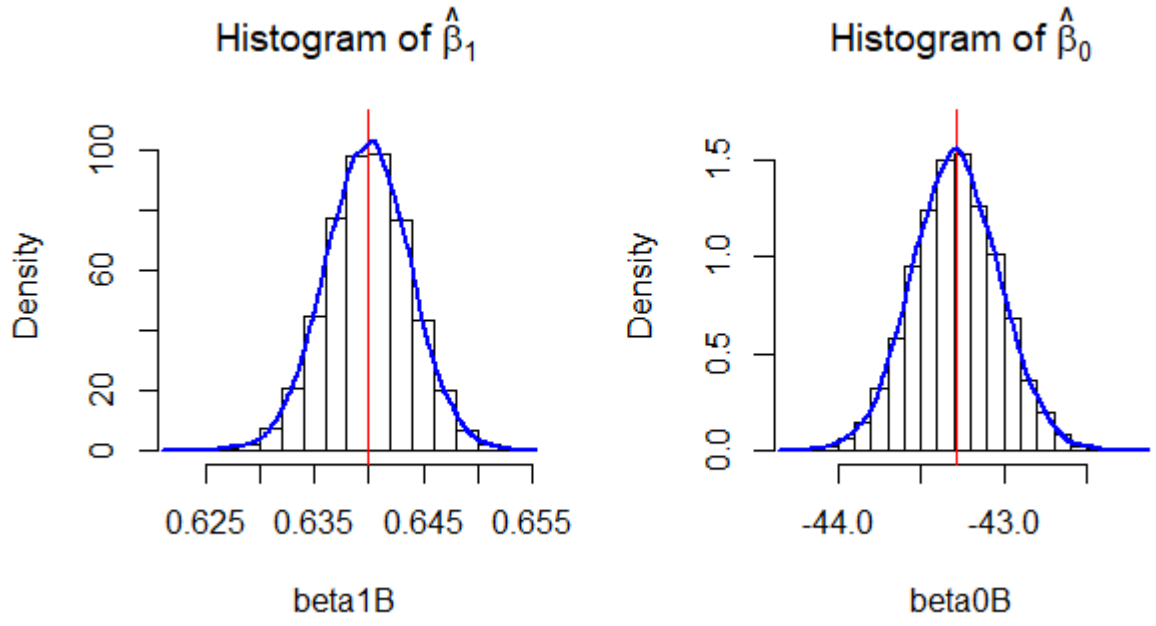

Figure 4.41: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 3$, $\sigma_0 = 5.25$, $\beta_1 = 0.64$, $\beta_0 = -43.29$

Table 4.45: *Summary of Simulation by Method II* : $\beta_1 = 0.64$, $\beta_0 = -43.29$
$\sigma_e = 3$, $\sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	-43.289	2×10^{-5}	0.070	[0.633, 0.648]	[-43.797, -42.784]
5000	0.640	-43.288	2×10^{-5}	0.069	[0.632, 0.648]	[-43.812, -42.775]
10000	0.640	-43.287	2×10^{-5}	0.068	[0.632, 0.648]	[-43.793, -42.771]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.287, with the MSE to be 0.068.

(3) When $\boldsymbol{\beta}_1 = 2.15,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.42. Table 4.46 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

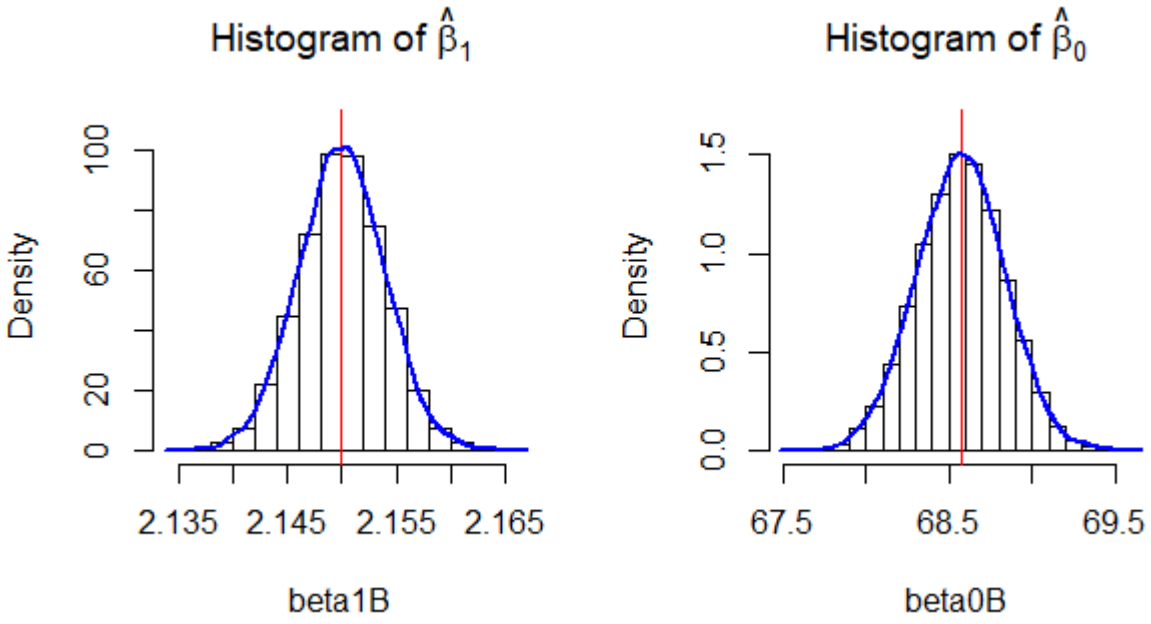

Figure 4.42: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 3,\ \sigma_0 = 5.25,\ \beta_1 = 2.15,\ \beta_0 = 68.57$

Table 4.46: *Summary of Simulation by Method II* : $\beta_1 = 2.15,\ \beta_0 = 68.57$
$\sigma_e = 3,\ \sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	68.572	2×10^{-5}	0.070	[2.142, 2.158]	[68.053, 69.088]
5000	2.150	68.572	2×10^{-5}	0.070	[2.142, 2.158]	[68.061, 69.078]
10000	2.150	68.572	2×10^{-5}	0.070	[2.142, 2.158]	[68.051, 69.084]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.572, with the MSE to be 0.070.

(4) When $\boldsymbol{\beta}_1 = 2.15$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.43. Table 4.47 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

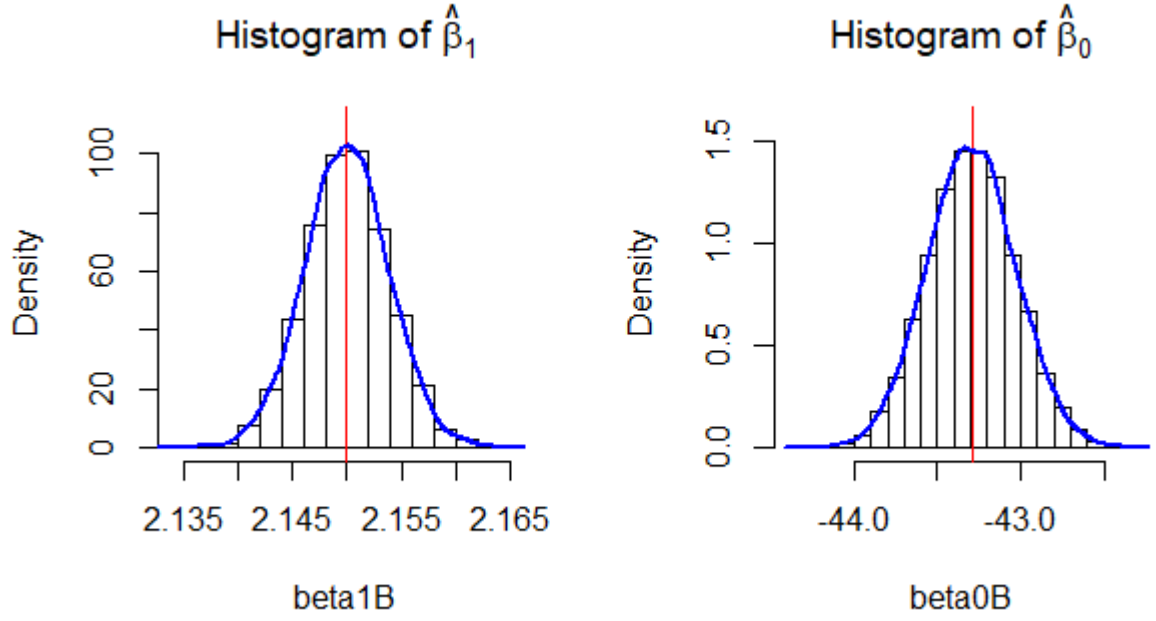

Figure 4.43: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3$, $\sigma_0 = 5.25$, $\beta_1 = 2.15$, $\beta_0 = -43.29$

Table 4.47: *Summary of Simulation by Method II* : $\beta_1 = 2.15,\ \beta_0 = -43.29$ $\sigma_e = 3,\ \sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	-43.295	2×10^{-5}	0.071	[2.142, 2.158]	[-43.801, -42.761]
5000	2.150	-43.292	2×10^{-5}	0.069	[2.142, 2.158]	[-43.805, -42.769]
10000	2.150	-43.293	2×10^{-5}	0.070	[2.142, 2.158]	[-43.809, -42.777]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.293, with the MSE to be 0.070.

(5) When $\boldsymbol{\beta}_1 = -3.21,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.44. Table 4.48 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

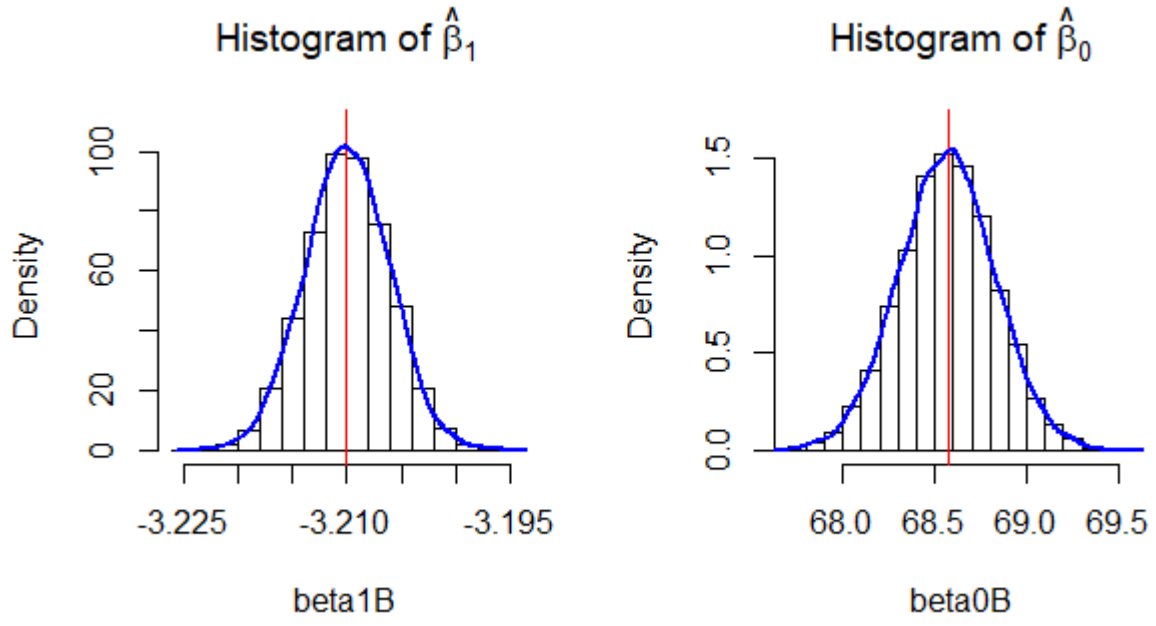

Figure 4.44: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 3$, $\sigma_0 = 5.25$, $\beta_1 = -3.21$, $\beta_0 = 68.57$

Table 4.48: *Summary of Simulation by Method II* : $\beta_1 = -3.21$, $\beta_0 = 68.57$
$\sigma_e = 3$, $\sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	68.565	2×10^{-5}	0.068	[-3.217, -3.202]	[68.053, 69.076]
5000	-3.210	68.571	2×10^{-5}	0.070	[-3.218, -3.202]	[68.047, 69.079]
10000	-3.210	68.567	2×10^{-5}	0.068	[-3.218, -3.202]	[68.055, 69.076]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.567, with the MSE to be 0.068.

(6) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.45. Table 4.49 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

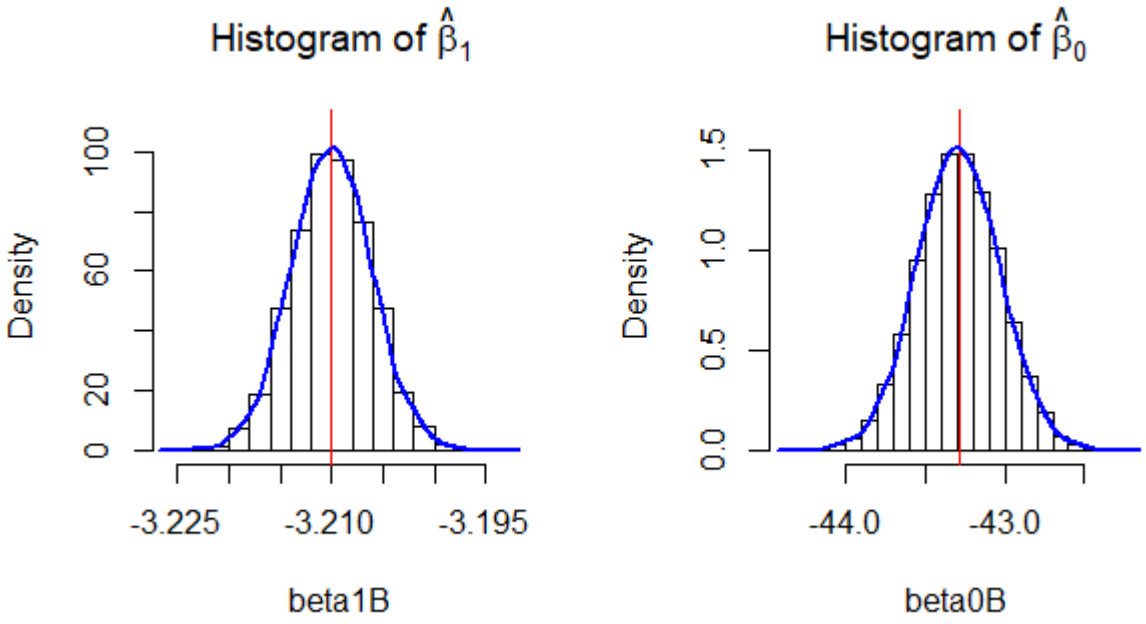

Figure 4.45: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3$, $\sigma_0 = 5.25$, $\beta_1 = -3.21$, $\beta_0 = -43.29$

Table 4.49: *Summary of Simulation by Method II* : $\beta_1 = -3.21$, $\beta_0 = -43.29$ $\sigma_e = 3$, $\sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	-43.294	2×10^{-5}	0.069	[-3.217, -3.203]	[-43.796, -42.777]
5000	-3.210	-43.296	2×10^{-5}	0.070	[-3.218, -3.202]	[-43.825, -42.774]
10000	-3.210	-43.291	2×10^{-5}	0.068	[-3.218, -3.202]	[-43.800, -42.780]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.291, with the MSE to be 0.068.

Method 2 - Set 2: $\sigma_e = 7$, $\sigma_0 = 5.25$

(1) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.46. Table 4.50 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

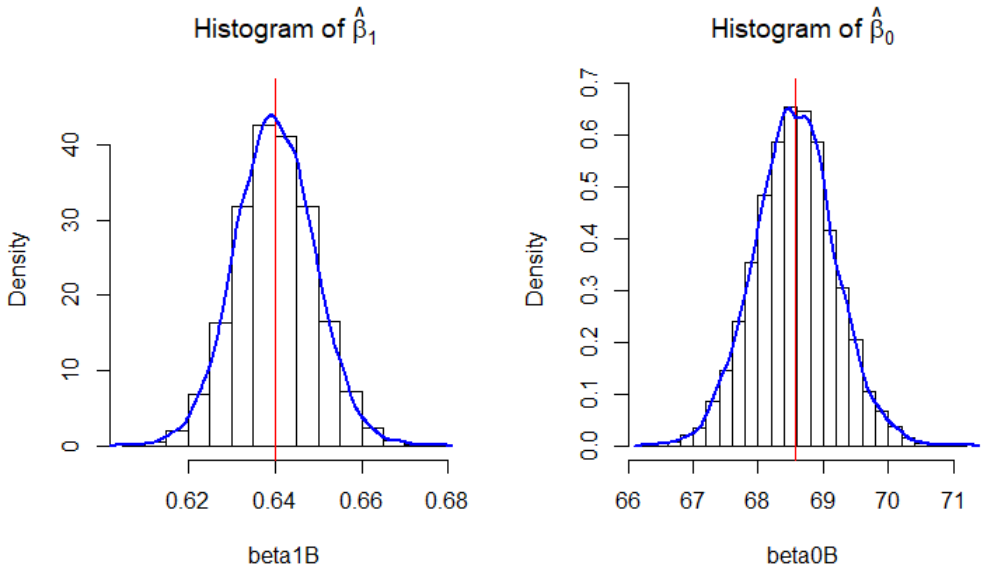

Figure 4.46: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $\sigma_0 = 5.25$, $\beta_1 = 0.64$, $\beta_0 = 68.57$

Table 4.50: *Summary of Simulation by Method II* : $\beta_1 = 0.64,\ \beta_0 = 68.57$ $\sigma_e = 7,\ \sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	68.569	8×10^{-5}	0.379	[0.623, 0.658]	[67.373, 69.795]
5000	0.640	68.585	8×10^{-5}	0.381	[0.622, 0.658]	[67.368, 69.791]
10000	0.640	68.564	8×10^{-5}	0.378	[0.622, 0.658]	[67.355, 69.799]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.564, with the MSE to be 0.378.

(2) When $\boldsymbol{\beta}_1 = 0.64,\ \boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.47. Table 4.51 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

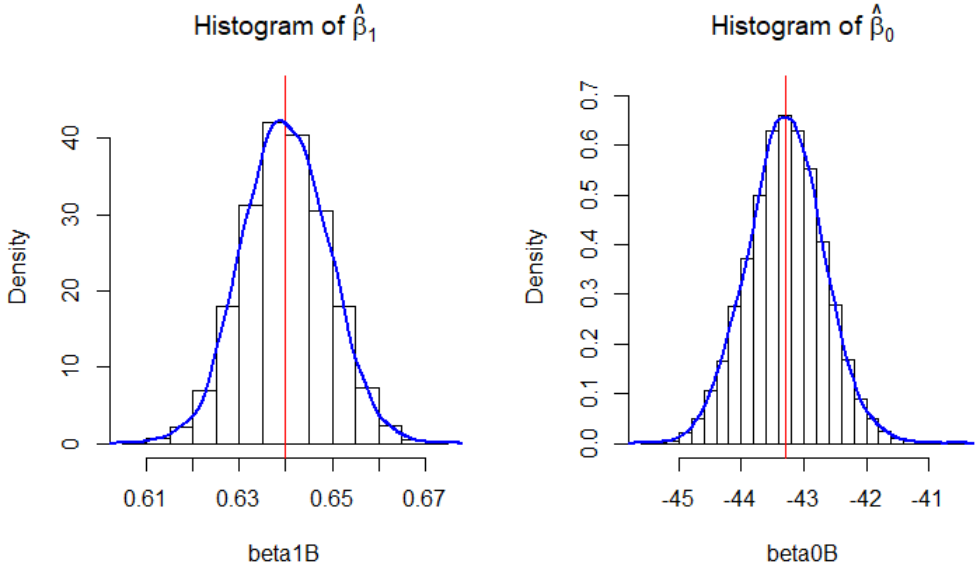

Figure 4.47: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $\sigma_0 = 5.25$, $\beta_1 = 0.64$, $\beta_0 = -43.29$

Table 4.51: *Summary of Simulation by Method II* : $\beta_1 = 0.64$, $\beta_0 = -43.29$ $\sigma_e = 7$, $\sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	-43.307	9×10^{-5}	0.375	[0.622, 0.658]	[-44.514, -42.123]
5000	0.640	-43.296	8×10^{-5}	0.374	[0.622, 0.658]	[-44.499, -42.090]
10000	0.640	-43.290	8×10^{-5}	0.370	[0.622, 0.658]	[-44.485, -42.097]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.290, with the MSE to be 0.370.

(3) When $\boldsymbol{\beta}_1 = 2.15,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.48. Table 4.52 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

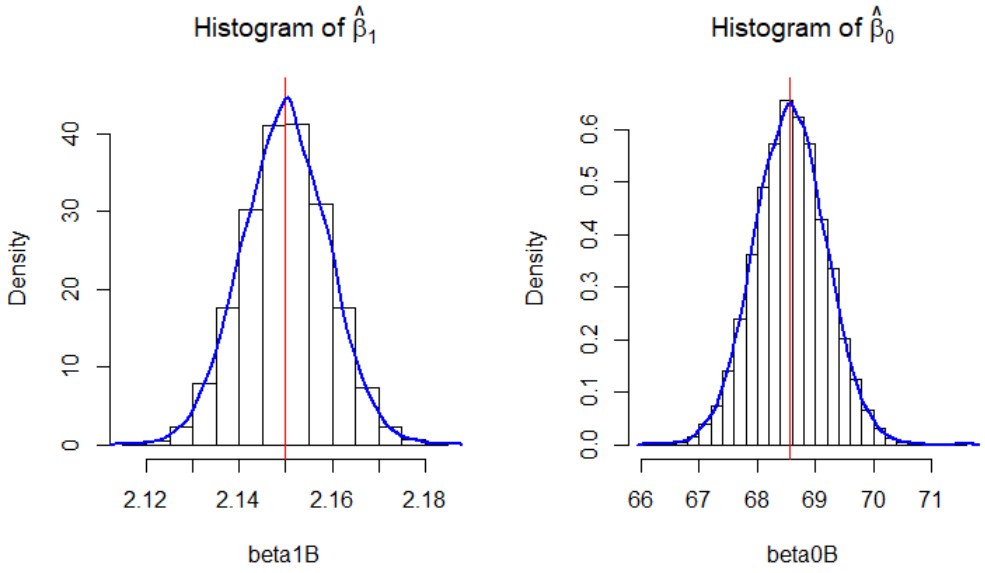

Figure 4.48: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $\sigma_0 = 5.25$, $\beta_1 = 2.15$, $\beta_0 = 68.57$

Table 4.52: *Summary of Simulation by Method II* : $\beta_1 = 2.15,\ \beta_0 = 68.57$ $\sigma_e = 7$, $\sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	68.555	8×10^{-5}	0.373	[2.132, 2.168]	[67.375, 69.721]
5000	2.150	68.585	8×10^{-5}	0.365	[2.131, 2.167]	[67.398, 69.766]
10000	2.150	68.571	9×10^{-5}	0.378	[2.132, 2.168]	[67.375, 69.781]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 9×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.571, with the MSE to be 0.378.

(4) When $\boldsymbol{\beta}_1 = 2.15$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.49. Table 4.53 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

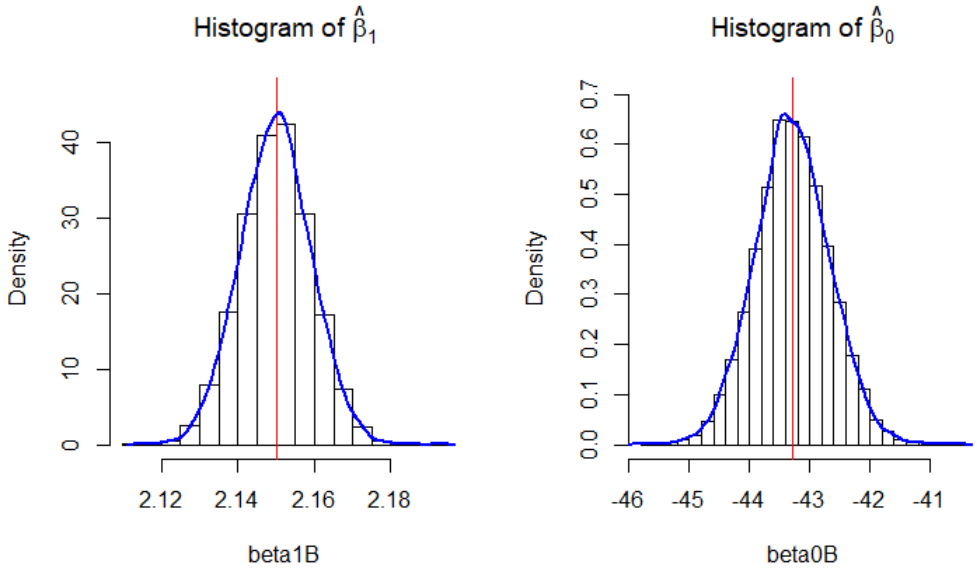

Figure 4.49: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $\sigma_0 = 5.25$, $\beta_1 = 2.15$, $\beta_0 = -43.29$

Table 4.53: *Summary of Simulation by Method II* : $\beta_1 = 2.15$, $\beta_0 = -43.29$ $\sigma_e = 7$, $\sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	-43.316	9×10^{-5}	0.367	[2.132, 2.169]	[-44.533, -42.102]
5000	2.150	-43.289	8×10^{-5}	0.377	[2.132, 2.168]	[-44.521, -42.096]
10000	2.150	-43.287	8×10^{-5}	0.382	[2.132, 2.168]	[-44.482, -42.067]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.287, with the MSE to be 0.382.

(5) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.50. Table 4.54 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

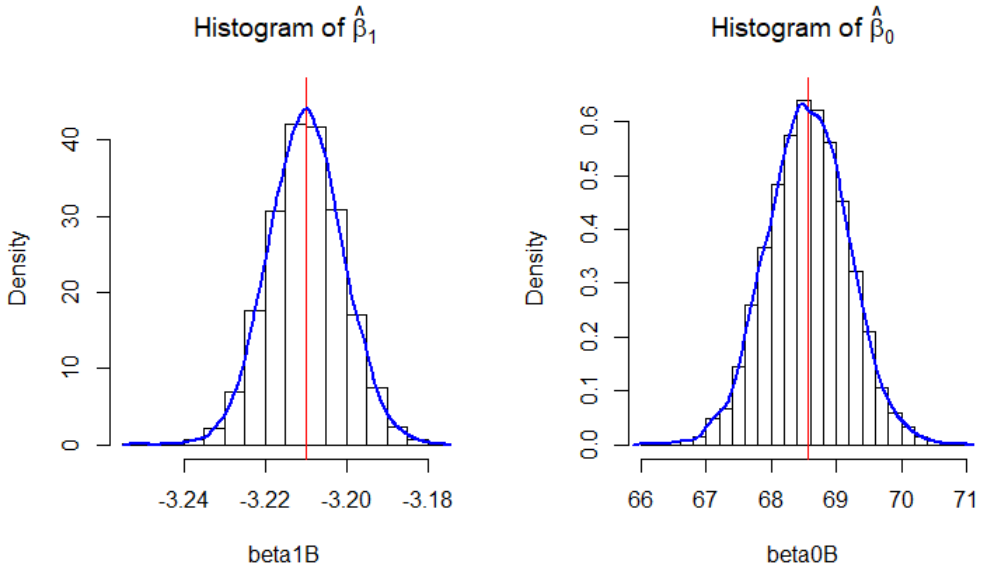

Figure 4.50: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $\sigma_0 = 5.25$, $\beta_1 = -3.21$, $\beta_0 = 68.57$

Table 4.54: *Summary of Simulation by Method II* : $\beta_1 = -3.21$, $\beta_0 = 68.57$ $\sigma_e = 7$, $\sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	68.574	8×10^{-5}	0.370	[-3.227, -3.192]	[67.412, 69.737]
5000	-3.210	68.574	8×10^{-5}	0.378	[-3.228, -3.192]	[67.369, 69.783]
10000	-3.210	68.564	8×10^{-5}	0.382	[-3.228, -3.192]	[67.355, 69.776]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.564, with the MSE to be 0.382.

(6) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.51. Table 4.55 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

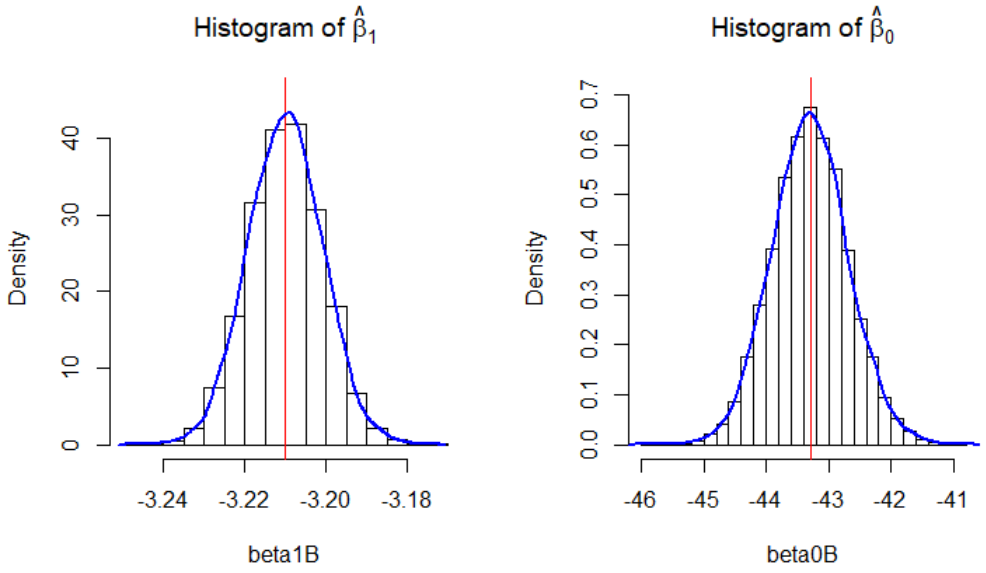

Figure 4.51: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 7$, $\sigma_0 = 5.25$, $\beta_1 = -3.21$, $\beta_0 = -43.29$

Table 4.55: *Summary of Simulation by Method II* : $\beta_1 = -3.21$, $\beta_0 = -43.29$
$\sigma_e = 7$, $\sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	-43.287	9×10^{-5}	0.375	[-3.228, -3.191]	[-44.505, -42.110]
5000	-3.210	-43.298	8×10^{-5}	0.382	[-3.228, -3.192]	[-44.478, -42.066]
10000	-3.210	-43.295	8×10^{-5}	0.372	[-3.228, -3.192]	[-44.467, -42.085]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.295, with the MSE to be 0.372.

Method 2 - Set 3: $\sigma_e = 10$, $\sigma_0 = 5.25$

(1) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.52. Table 4.56 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

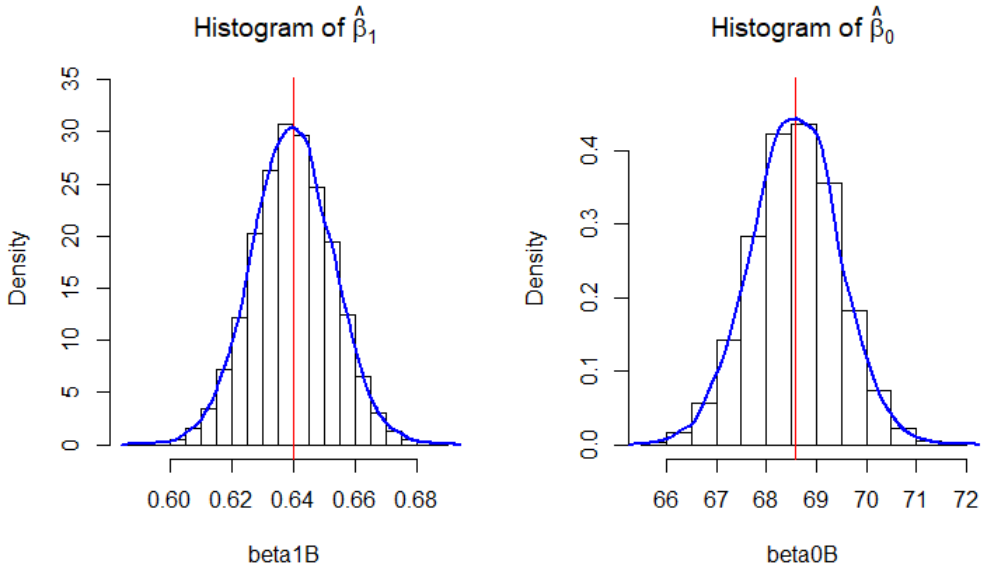

Figure 4.52: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $\sigma_0 = 5.25$, $\beta_1 = 0.64$, $\beta_0 = 68.57$

Table 4.56: *Summary of Simulation by Method II* : $\beta_1 = 0.64$, $\beta_0 = 68.57$ $\sigma_e = 10$, $\sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	68.563	1.6×10^{-4}	0.737	[0.615, 0.664]	[66.882, 70.172]
5000	0.640	68.555	1.7×10^{-4}	0.786	[0.614, 0.665]	[66.795, 70.301]
10000	0.640	68.578	1.7×10^{-4}	0.760	[0.614, 0.665]	[68.857, 70.273]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 1.7×10^{-4};

the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.578, with the MSE to be 0.760.

(2) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.53. Table 4.57 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

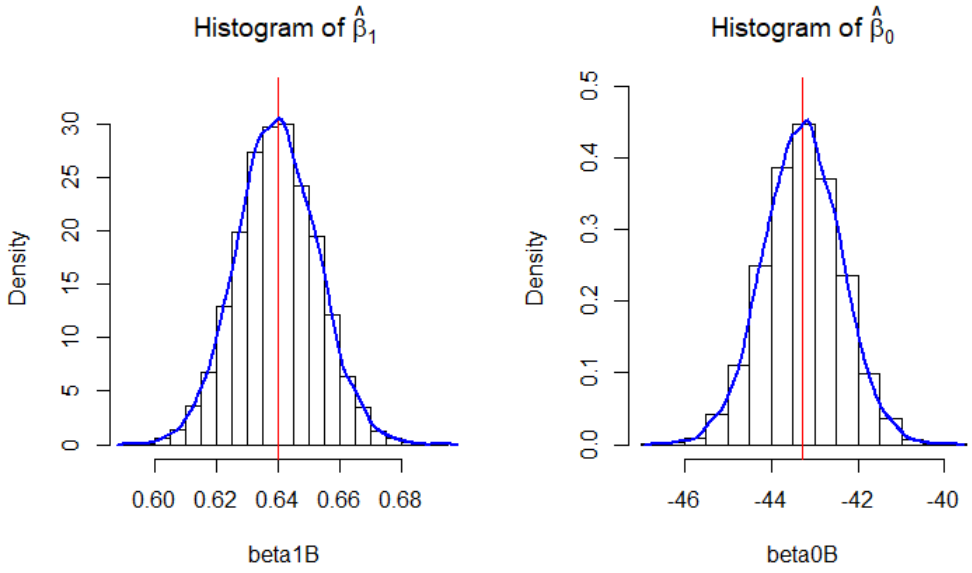

Figure 4.53: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $\sigma_0 = 5.25$, $\beta_1 = 0.64$, $\beta_0 = -43.29$

Table 4.57: *Summary of Simulation by Method II* : $\beta_1 = 0.64$, $\beta_0 = -43.29$ $\sigma_e = 10$, $\sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.639	-43.235	1.7×10^{-4}	0.759	[0.613, 0.665]	[-44.852, -41.509]
5000	0.640	-43.278	1.8×10^{-4}	0.777	[0.615, 0.667]	[-44.957, -41.554]
10000	0.640	-43.280	1.7×10^{-4}	0.783	[0.614, 0.666]	[-45.026, -41.549]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.280, with the MSE to be 0.783.

(3) When $\boldsymbol{\beta}_1 = 2.15$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.54. Table 4.58 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

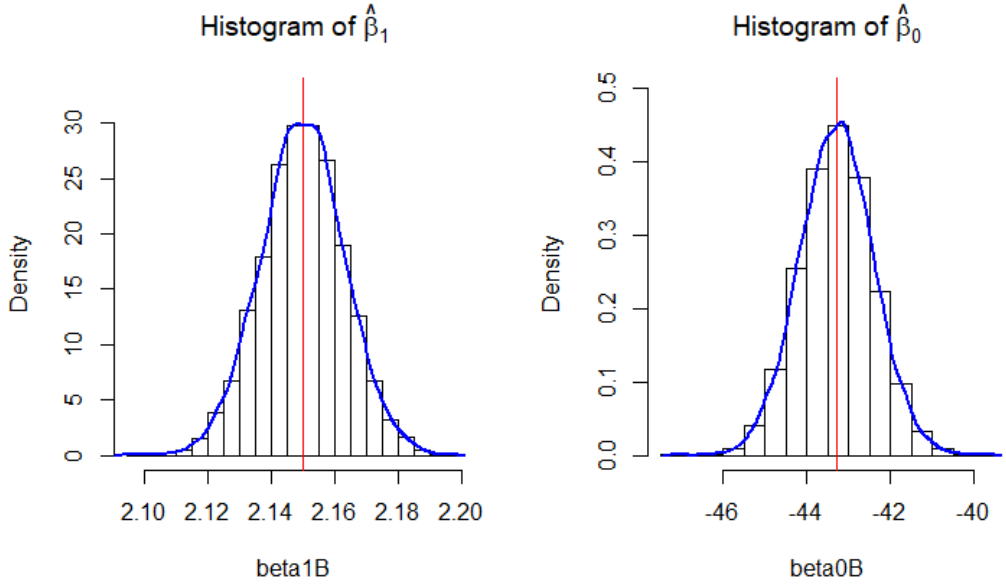

Figure 4.54: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $\sigma_0 = 5.25$, $\beta_1 = 2.15$, $\beta_0 = -43.29$

Table 4.58: *Summary of Simulation by Method II* : $\beta_1 = 2.15,\ \beta_0 = -43.29$ $\sigma_e = 10,\ \sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	-43.303	1.8×10^{-4}	0.760	[2.124, 2.176]	[-44.942, -41.533]
5000	2.150	-43.279	1.7×10^{-4}	0.759	[2.125, 2.175]	[-45.034, -41.584]
10000	2.150	-43.290	1.7×10^{-4}	0.771	[2.124, 2.176]	[-45.018, -41.567]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be $\times 10^{-5}$; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.290, with the MSE to be 0.771.

(4) When $\boldsymbol{\beta}_1 = -3.21,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.55. Table 4.59 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

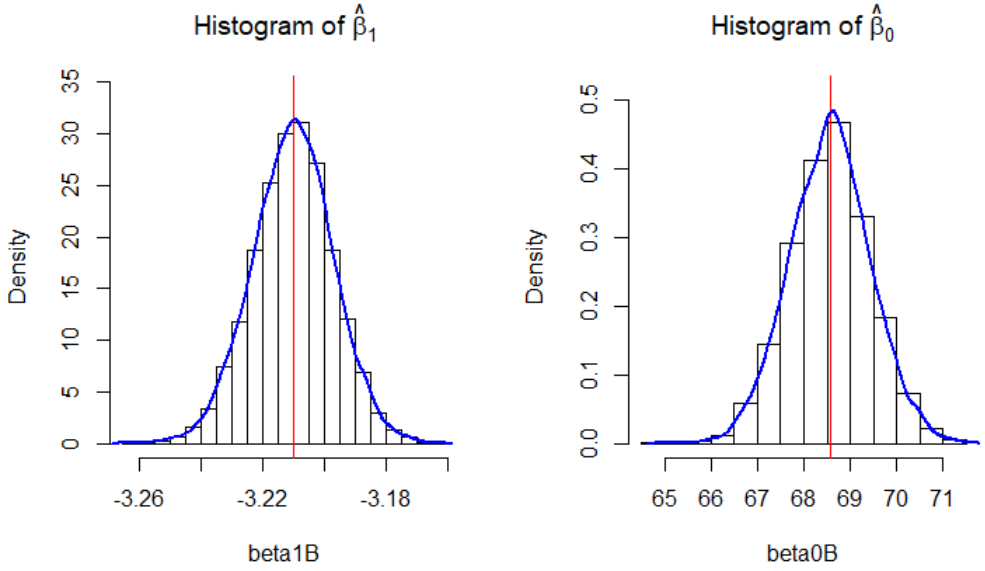

Figure 4.55: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $\sigma_0 = 5.25$, $\beta_1 = -3.21$, $\beta_0 = 68.57$

Table 4.59: *Summary of Simulation by Method II* : $\beta_1 = -3.21$, $\beta_0 = 68.57$ $\sigma_e = 10$, $\sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	68.545	1.6×10^{-4}	0.766	[-3.235, -3.184]	[66.846, 70.277]
5000	-3.210	68.582	1.7×10^{-4}	0.783	[-3.235, -3.185]	[66.861, 70.289]
10000	-3.210	68.572	1.7×10^{-4}	0.748	[-3.236, -3.185]	[66.861, 70.289]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.572, with the MSE to be 0.748.

(5) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.56. Table 4.60 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

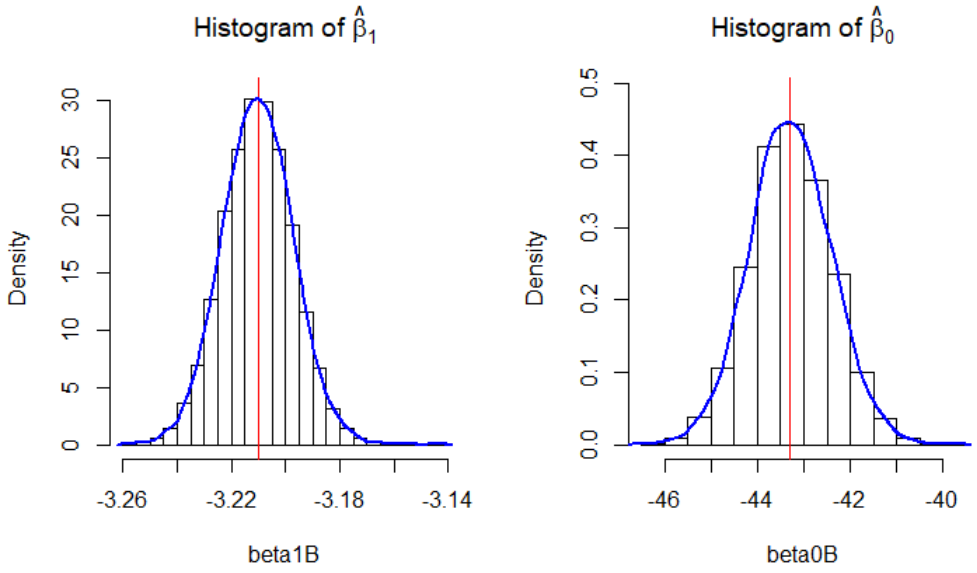

Figure 4.56: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 10$, $\sigma_0 = 5.25$, $\beta_1 = -3.21$, $\beta_0 = -43.29$

Table 4.60: *Summary of Simulation by Method II* : $\beta_1 = -3.21$, $\beta_0 = -43.29$
$\sigma_e = 10$, $\sigma_0 = 5.25$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	-43.309	1.7×10^{-4}	0.728	[-3.236, -3.184]	[-44.987, -41.625]
5000	-3.210	-43.290	1.7×10^{-4}	0.781	[-3.235, -3.184]	[-45.050, -41.556]
10000	-3.210	-43.277	1.7×10^{-4}	0.767	[-3.236, -3.184]	[-44.988, -41.540]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.277, with the MSE to be 0.767.

Method 2 - Set 4: $\sigma_e = 3$, $\sigma_0 = 8.07$

(1) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.57. Table 4.61 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

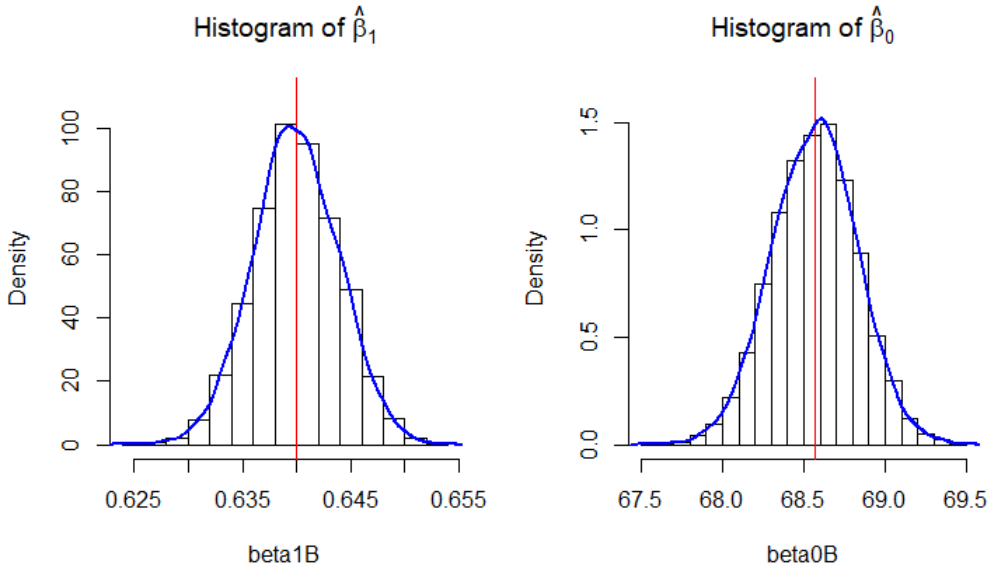

Figure 4.57: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3$, $\sigma_0 = 8.07$, $\beta_1 = 0.64$, $\beta_0 = 68.57$

Table 4.61: *Summary of Simulation by Method II* : $\beta_1 = 0.64$, $\beta_0 = 68.57$ $\sigma_e = 3$, $\sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	68.575	2×10^{-5}	0.069	[0.632, 0.648]	[68.063, 69.088]
5000	0.640	68.563	2×10^{-5}	0.069	[0.632, 0.648]	[68.041, 69.090]
10000	0.640	68.568	2×10^{-5}	0.069	[0.632, 0.648]	[68.051, 69.075]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.568, with the MSE to be 0.069.

(2) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.58. Table 4.62 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

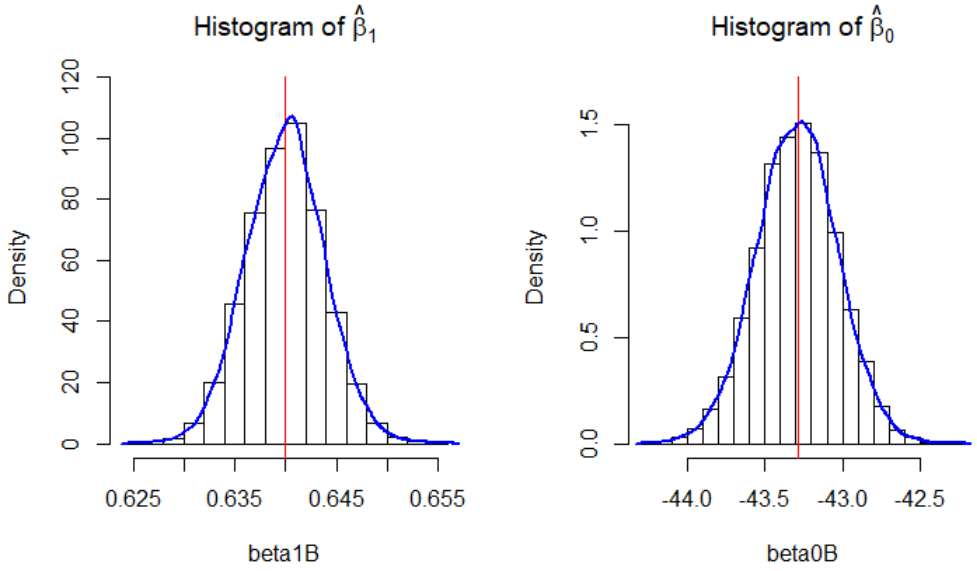

Figure 4.58: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3$, $\sigma_0 = 8.07$, $\beta_1 = 0.64$, $\beta_0 = -43.29$

Table 4.62: *Summary of Simulation by Method II* : $\beta_1 = 0.64,\ \beta_0 = -43.29$ $\sigma_e = 3,\ \sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	-43.288	2×10^{-5}	0.072	[0.632, 0.648]	[-43.809, -42.751]
5000	0.640	-43.291	2×10^{-5}	0.070	[0.632, 0.648]	[-43.794, -42.763]
10000	0.640	-43.292	2×10^{-5}	0.067	[0.633, 0.648]	[-43.806, -42.788]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.292, with the MSE to be 0.067.

(3) When $\boldsymbol{\beta}_1 = 2.15,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.59. Table 4.63 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

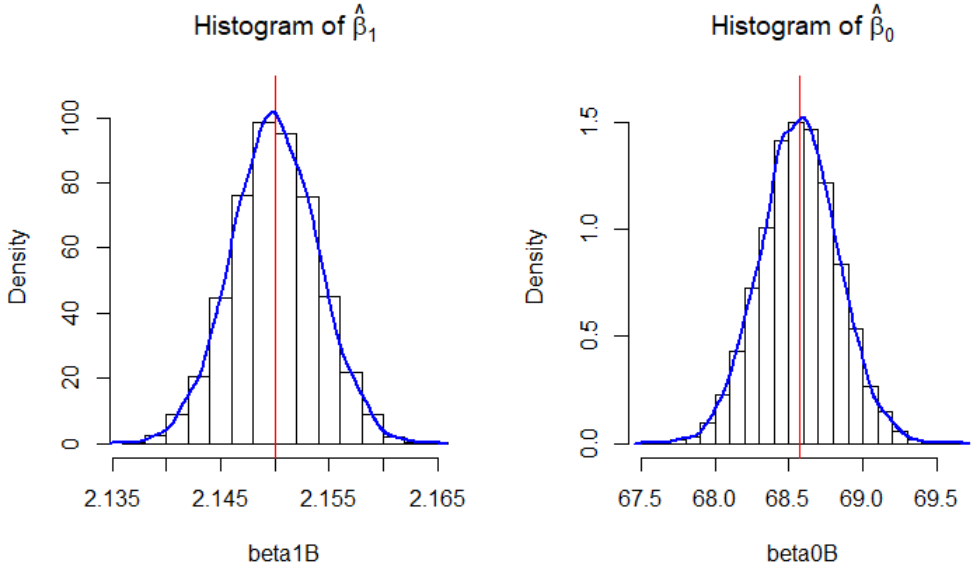

Figure 4.59: FHistograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 3$, $\sigma_0 = 8.07$, $\beta_1 = 2.15$, $\beta_0 = 68.57$

Table 4.63: *Summary of Simulation by Method II* : $\beta_1 = 2.15$, $\beta_0 = 68.57$
$\sigma_e = 3$, $\sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	68.571	2×10^{-5}	0.067	[2.143, 2.158]	[68.071, 69.074]
5000	2.150	68.566	2×10^{-5}	0.070	[2.142, 2.158]	[68.037, 69.074]
10000	2.150	68.570	2×10^{-5}	0.068	[2.142, 2.158]	[68.054, 69.086]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.570, with the MSE to be 0.068.

(4) When $\boldsymbol{\beta}_1 = 2.15$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.60. Table 4.64 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

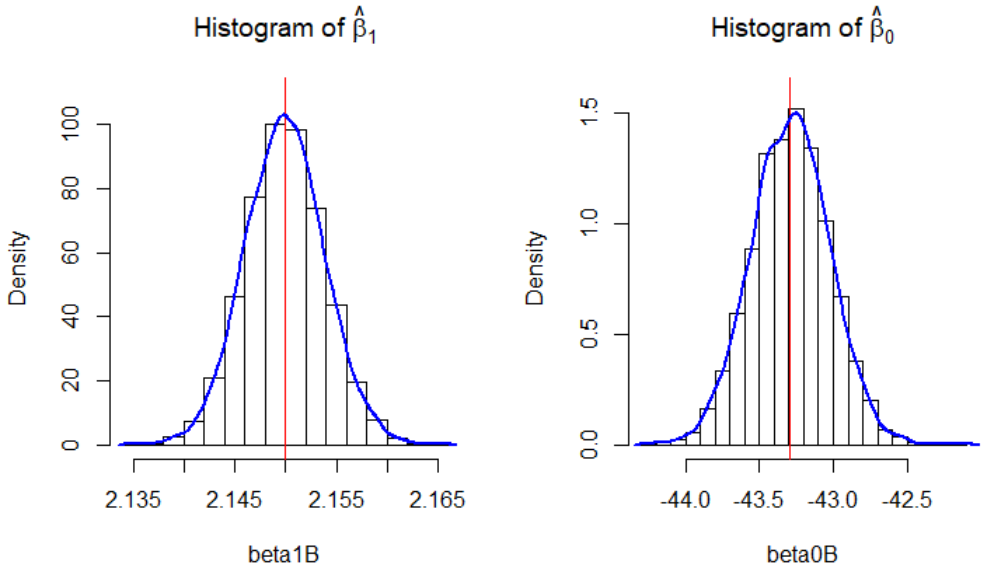

Figure 4.60: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3$, $\sigma_0 = 8.07$, $\beta_1 = 2.15$, $\beta_0 = -43.29$

Table 4.64: *Summary of Simulation by Method II* : $\beta_1 = 2.15$, $\beta_0 = -43.29$ $\sigma_e = 3$, $\sigma_0 = 8.07$

$B.$	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	-43.287	2×10^{-5}	0.071	[2.142, 2.158]	[-43.807, -42.769]
5000	2.150	-43.290	2×10^{-5}	0.072	[2.142, 2.158]	[-43.810, -42.766]
10000	2.150	-43.288	2×10^{-5}	0.069	[2.142, 2.158]	[-43.801, -42.779]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.288, with the MSE to be 0.069.

(5) $\boldsymbol{\beta}_1 = -3.21, \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.61. Table 4.65 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

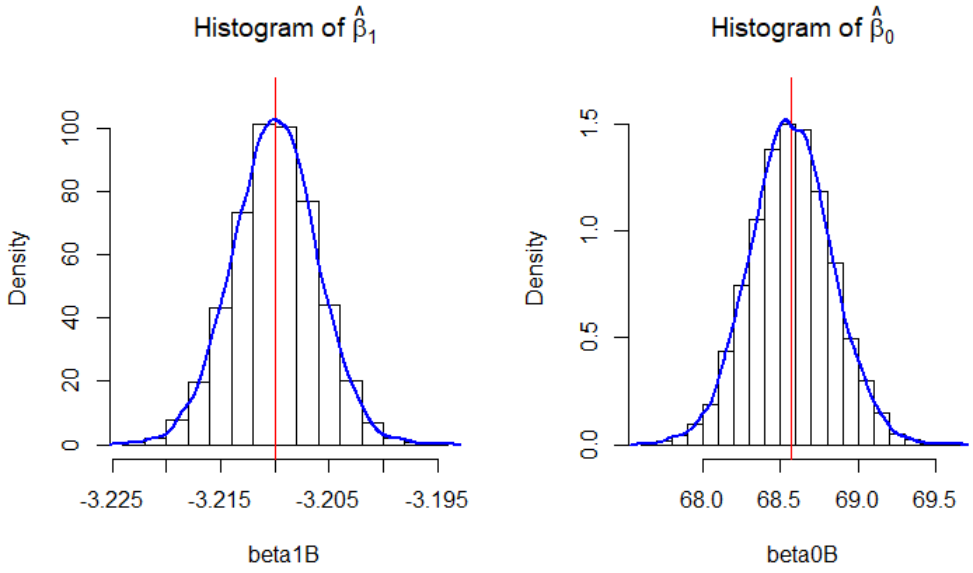

Figure 4.61: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3$, $\sigma_0 = 8.07$, $\beta_1 = -3.21$, $\beta_0 = 68.57$

Table 4.65: *Summary of Simulation by Method II* : $\beta_1 = -3.21$, $\beta_0 = 68.57$ $\sigma_e = 3$, $\sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	68.571	2×10^{-5}	0.073	[-3.218, -3.202]	[68.035, 69.094]
5000	-3.210	68.576	2×10^{-5}	0.070	[-3.218, -3.202]	[68.071, 69.105]
10000	-3.210	68.569	2×10^{-5}	0.069	[-3.218, -3.202]	[68.054, 69.085]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.569, with the MSE to be 0.069.

(6) $\boldsymbol{\beta}_1 = -3.21, \boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.62. Table 4.66 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

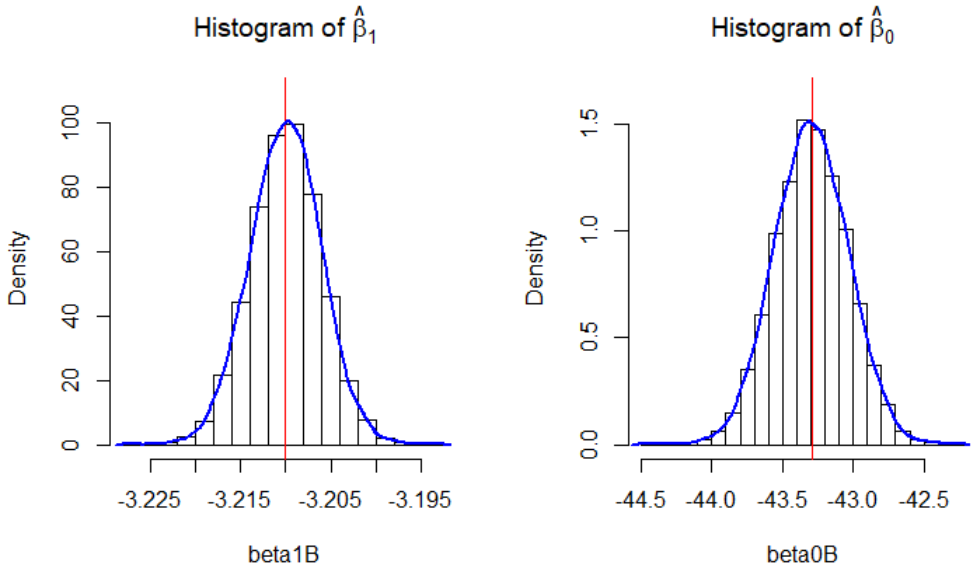

Figure 4.62: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 3$, $\sigma_0 = 8.07$, $\beta_1 = -3.21$, $\beta_0 = -43.29$

Table 4.66: *Summary of Simulation by Method II* : $\beta_1 = -3.21$, $\beta_0 = -43.29$ $\sigma_e = 3$, $\sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	-43.301	2×10^{-5}	0.066	[-3.218, -3.202]	[-43.821, -42.797]
5000	-3.210	-43.292	2×10^{-5}	0.069	[-3.218, -3.202]	[-43.814, -42.772]
10000	-3.210	-43.294	2×10^{-5}	0.068	[-3.218, -3.202]	[-43.800, -42.784]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 2×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.294, with the MSE to be 0.069.

Method 2 - Set 5: $\sigma_e = 7$, $\sigma_0 = 8.07$

(1) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.63. Table 4.67 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

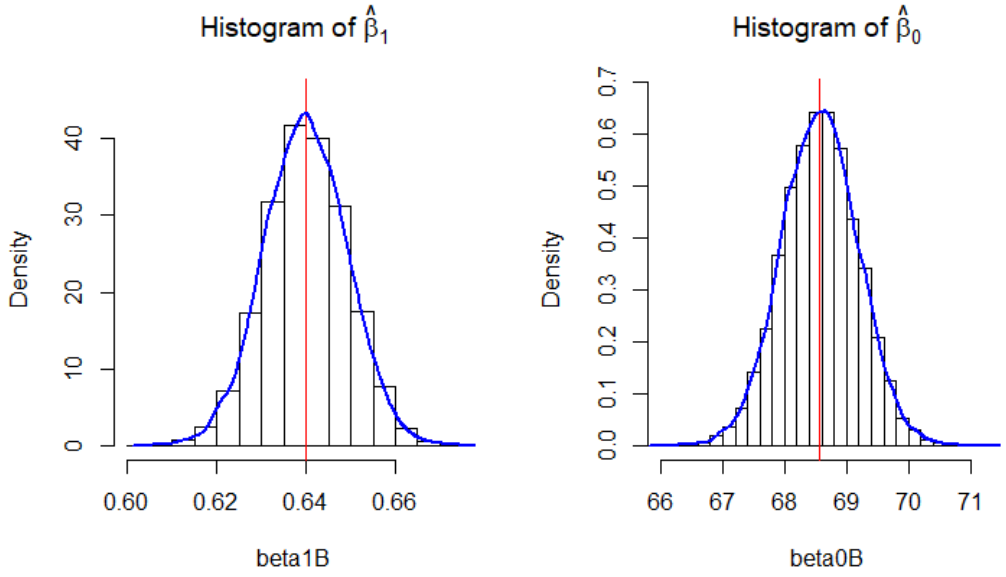

Figure 4.63: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $\sigma_0 = 8.07$, $\beta_1 = 0.64$, $\beta_0 = 68.57$

Table 4.67: *Summary of Simulation by Method II* : $\beta_1 = 0.64$, $\beta_0 = 68.57$ $\sigma_e = 7$, $\sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	68.582	9×10^{-5}	0.391	[0.621, 0.658]	[67.366, 69.741]
5000	0.640	68.572	9×10^{-5}	0.373	[0.622, 0.657]	[67.385, 69.773]
10000	0.640	68.574	9×10^{-5}	0.373	[0.622, 0.658]	[67.378, 69.767]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 9×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.574, with the MSE to be 0.373.

(2) When $\boldsymbol{\beta}_1 = 0.64, \boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.64. Table 4.68 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

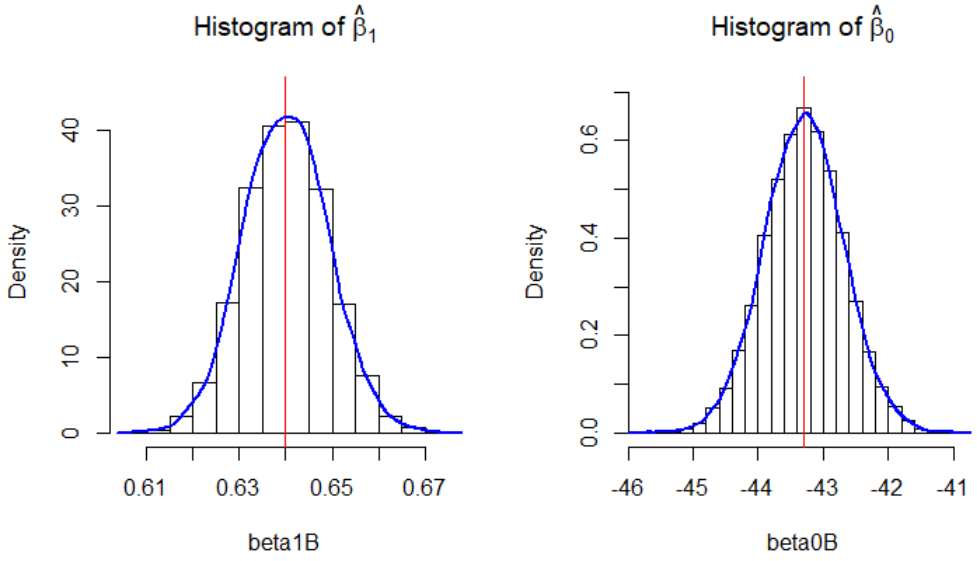

Figure 4.64: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $\sigma_0 = 8.07$, $\beta_1 = 0.64$, $\beta_0 = -43.29$

Table 4.68: *Summary of Simulation by Method II* : $\beta_1 = 0.64,\ \beta_0 = -43.29$ $\sigma_e = 7,\ \sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	-43.300	8×10^{-5}	0.378	[0.622, 0.658]	[-44.500, -42.127]
5000	0.640	-43.300	8×10^{-5}	0.379	[0.622, 0.659]	[-44.529, -42.088]
10000	0.640	-43.294	8×10^{-5}	0.369	[0.622, 0.658]	[-44.484, -42.091]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.294, with the MSE to be 0.369.

(3) When $\boldsymbol{\beta}_1 = 2.15,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.65. Table 4.69 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

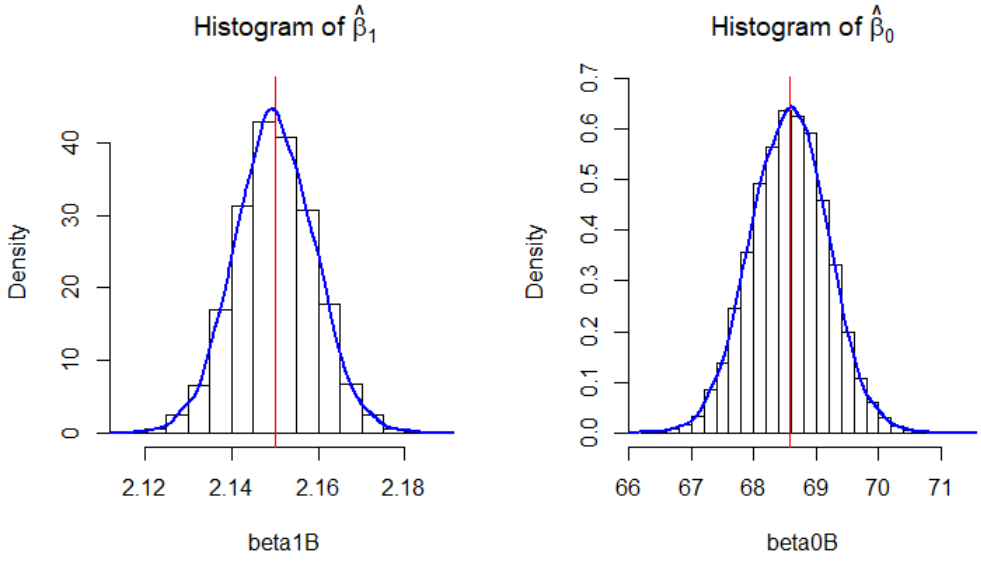

Figure 4.65: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 7$, $\sigma_0 = 8.07$, $\beta_1 = 2.15$, $\beta_0 = 68.57$

Table 4.69: *Summary of Simulation by Method II* : $\beta_1 = 2.15$, $\beta_0 = 68.57$
$\sigma_e = 7$, $\sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	68.574	9×10^{-5}	0.375	[2.131, 2.167]	[67.406, 69.768]
5000	2.150	68.584	8×10^{-5}	0.372	[2.132, 2.168]	[67.383, 69.783]
10000	2.150	68.569	8×10^{-5}	0.376	[2.132, 2.168]	[67.355, 69.755]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.569, with the MSE to be 0.376.

(4) When $\boldsymbol{\beta}_1 = 2.15$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.66. Table 4.70 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

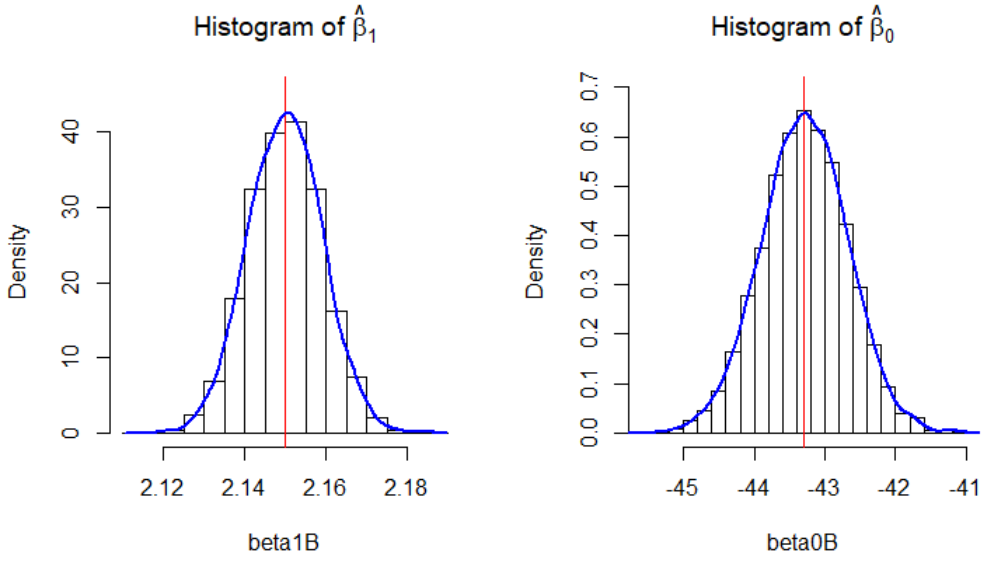

Figure 4.66: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $\sigma_0 = 8.07$, $\beta_1 = 2.15$, $\beta_0 = -43.29$

Table 4.70: *Summary of Simulation by Method II* : $\beta_1 = 2.15$, $\beta_0 = -43.29$ $\sigma_e = 7$, $\sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	-43.295	8×10^{-5}	0.357	[2.133, 2.168]	[-44.485, -42.145]
5000	2.150	-43.298	8×10^{-5}	0.371	[2.133, 2.168]	[-44.512, -42.118]
10000	2.150	-43.283	8×10^{-5}	0.371	[2.132, 2.168]	[-44.496, -42.108]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.283, with the MSE to be 0.371.

(5) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.67. Table 4.71 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

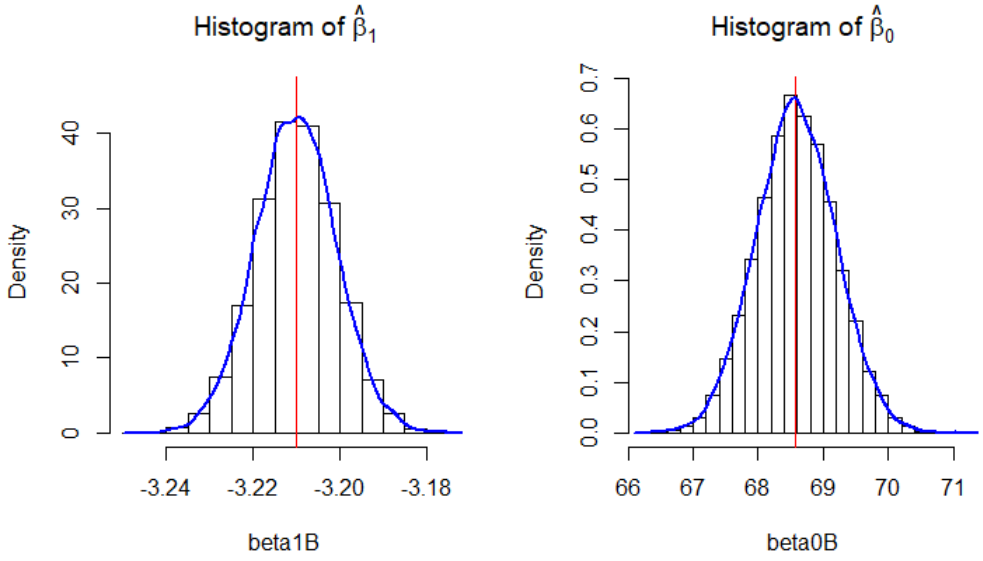

Figure 4.67: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 7$, $\sigma_0 = 8.07$, $\beta_1 = -3.21$, $\beta_0 = 68.57$

Table 4.71: *Summary of Simulation by Method II* : $\beta_1 = -3.21$, $\beta_0 = 68.57$ $\sigma_e = 7$, $\sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	68.570	9×10^{-5}	0.385	[-3.228, -3.192]	[67.326, 69.778]
5000	-3.210	68.564	8×10^{-5}	0.381	[-3.228, -3.191]	[67.359, 69.783]
10000	-3.210	68.583	9×10^{-5}	0.374	[-3.228, -3.192]	[67.395, 69.791]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 9×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.583, with the MSE to be 0.374.

(6) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.68. Table 4.72 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

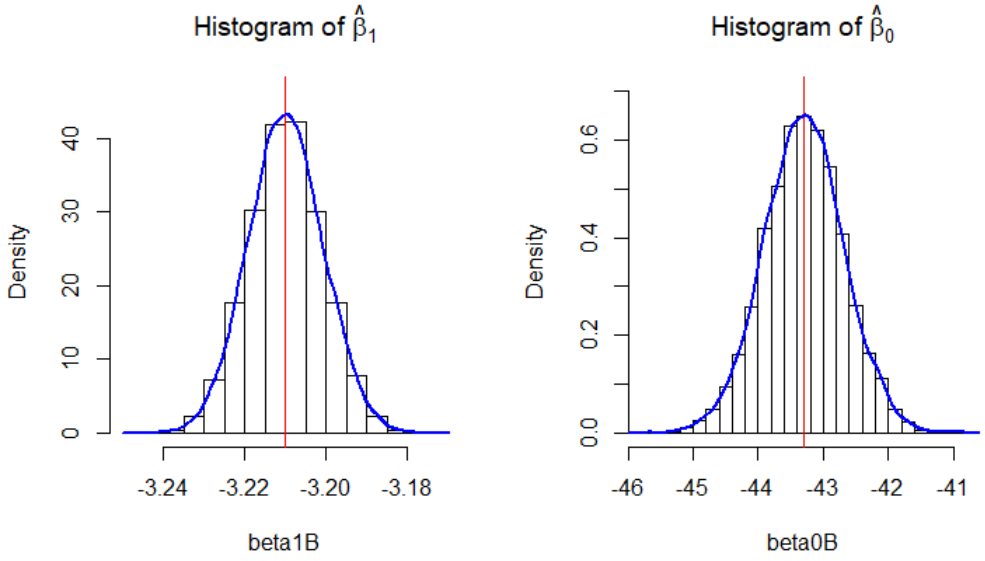

Figure 4.68: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 7$, $\sigma_0 = 8.07$, $\beta_1 = -3.21$, $\beta_0 = -43.29$

Table 4.72: *Summary of Simulation by Method II* : $\beta_1 = -3.21$, $\beta_0 = -43.29$
$\sigma_e = 7$, $\sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	-43.282	8×10^{-5}	0.375	[-3.228, -3.192]	[-44.486, -42.126]
5000	-3.210	-43.302	8×10^{-5}	0.382	[-3.227, -3.192]	[-44.552, -42.100]
10000	-3.210	-43.296	8×10^{-5}	0.374	[-3.228, -3.192]	[-44.511, -42.096]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 8×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.296, with the MSE to be 0.374.

Method 2 - Set 6: $\sigma_e = 10$, $\sigma_0 = 8.07$

(1) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.69. Table 4.73 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

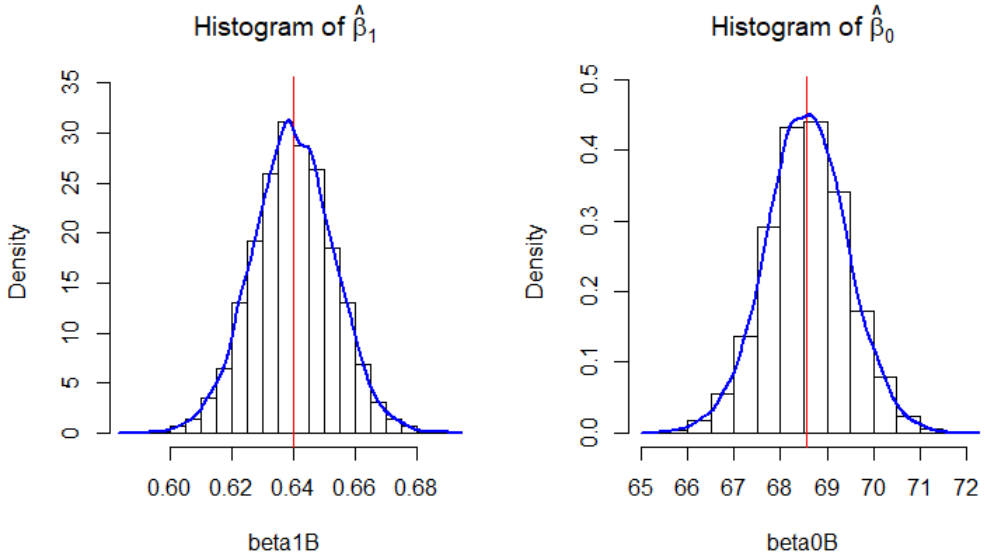

Figure 4.69: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 10$, $\sigma_0 = 8.07$, $\beta_1 = 0.64$, $\beta_0 = 68.57$

Table 4.73: *Summary of Simulation by Method II* : $\beta_1 = 0.64$, $\beta_0 = 68.57$
$\sigma_e = 10$, $\sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	68.560	1.7×10^{-4}	0.755	[0.615, 0.666]	[66.895, 70.263]
5000	0.640	68.565	1.7×10^{-4}	0.767	[0.614, 0.666]	[66.866, 70.295]
10000	0.640	68.569	1.7×10^{-4}	0.766	[0.614, 0.665]	[66.810, 70.282]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.569, with the MSE to be 0.766.

(2) When $\boldsymbol{\beta}_1 = 0.64$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.70. Table 4.74 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

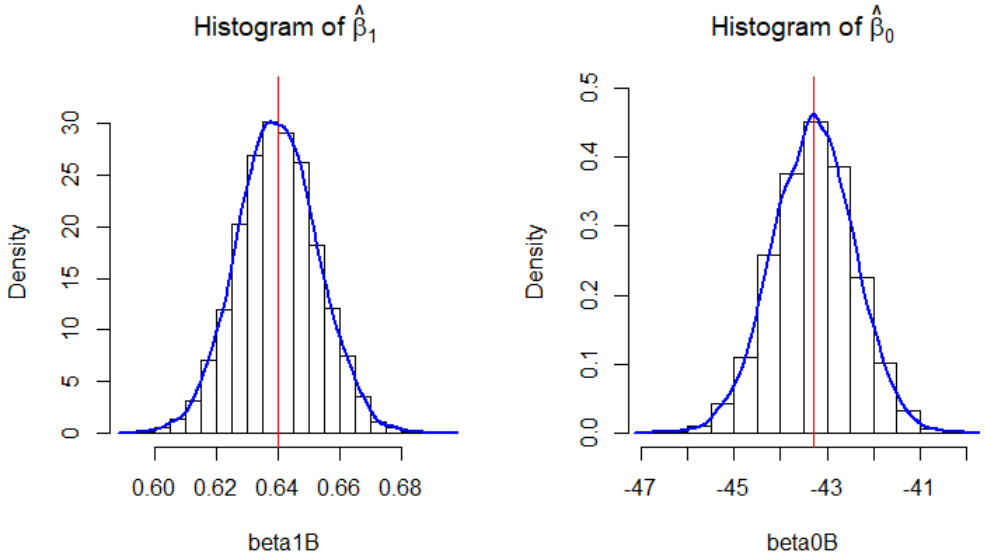

Figure 4.70: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $\sigma_0 = 8.07$, $\beta_1 = 0.64$, $\beta_0 = -43.29$

Table 4.74: *Summary of Simulation by Method II* : $\beta_1 = 0.64,\ \beta_0 = -43.29$ $\sigma_e = 10,\ \sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	0.640	-43.262	1.8×10^{-4}	0.792	[0.613, 0.666]	[-45.028, -41.525]
5000	0.640	-43.288	1.8×10^{-4}	0.781	[0.614, 0.666]	[-45.049, -41.598]
10000	0.640	-43.290	1.7×10^{-4}	0.767	[0.615, 0.665]	[-45.025, -41.593]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 0.640, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.290, with the MSE to be 0.767.

(3) When $\boldsymbol{\beta}_1 = 2.15,\ \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.71. Table 4.75 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

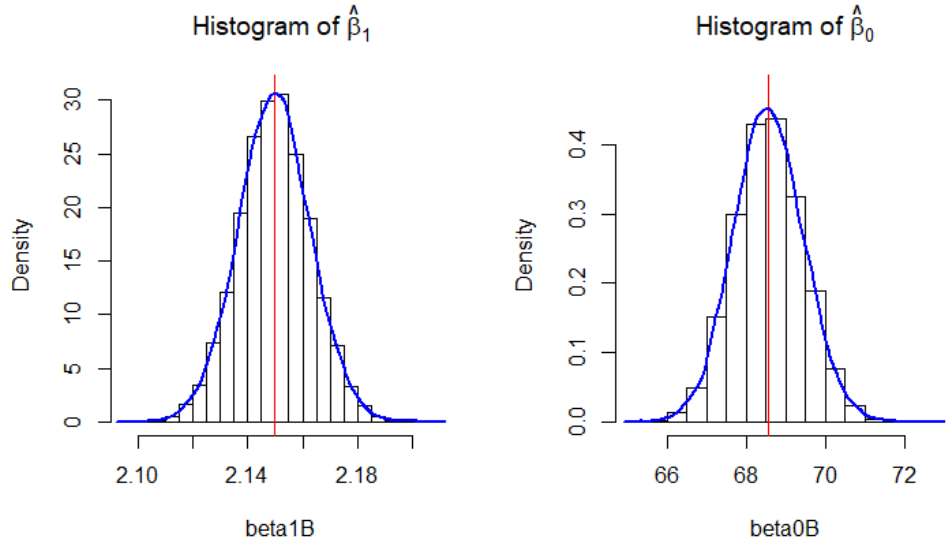

Figure 4.71: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10,\ \sigma_0 = 8.07,\ \beta_1 = 2.15,\ \beta_0 = 68.57$

Table 4.75: *Summary of Simulation by Method II* : $\beta_1 = 2.15,\ \beta_0 = 68.57$ $\sigma_e = 10,\ \sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	68.561	1.7×10^{-4}	0.736	[2.125, 2.175]	[66.964, 70.273]
5000	2.150	68.564	1.7×10^{-4}	0.751	[2.125, 2.175]	[66.866, 70.296]
10000	2.150	68.569	1.7×10^{-4}	0.764	[2.124, 2.176]	[66.882, 70.300]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.569, with the MSE to be 0.764.

(4) When $\boldsymbol{\beta}_1 = 2.15,\ \boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.72. Table 4.76 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

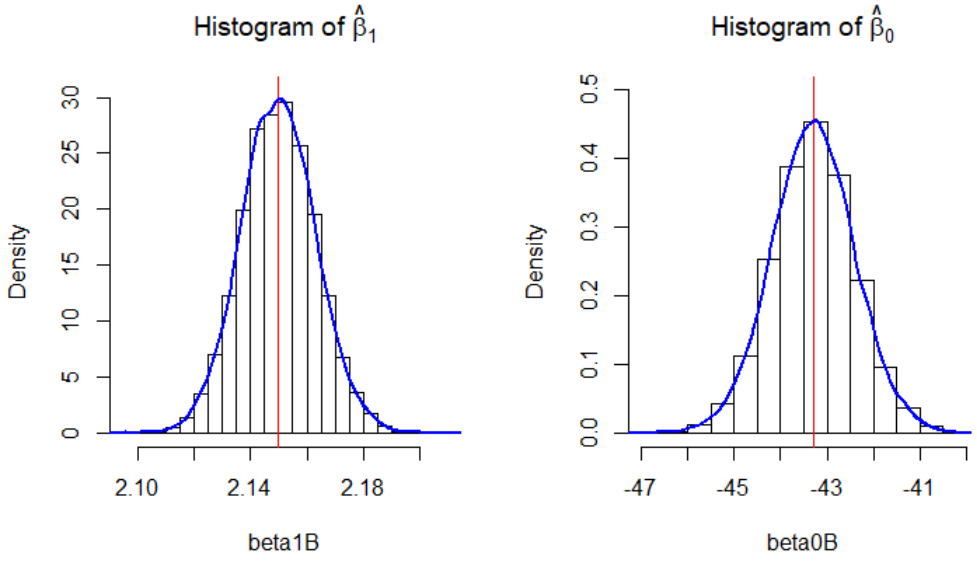

Figure 4.72: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions,
$\sigma_e = 10$, $\sigma_0 = 8.07$, $\beta_1 = 2.15$, $\beta_0 = -43.29$

Table 4.76: *Summary of Simulation by Method II* : $\beta_1 = 2.15$, $\beta_0 = -43.29$
$\sigma_e = 10$, $\sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	2.150	-43.265	1.6×10^{-4}	0.738	[2.124, 2.174]	[-44.948, -41.567]
5000	2.150	-43.263	1.7×10^{-4}	0.722	[2.125, 2.176]	[-45.092, -41.615]
10000	2.150	-43.293	1.7×10^{-4}	0.783	[2.124, 2.176]	[-45.041, -41.534]

The average value of $\hat{\beta}_1$ for 10000 repetitions is 2.150, with the MSE to be 1.7×10^{-5}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.293, with the MSE to be 0.783.

(5) When $\boldsymbol{\beta}_1 = -3.21, \boldsymbol{\beta}_0 = 68.57$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.73. Table 4.77 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

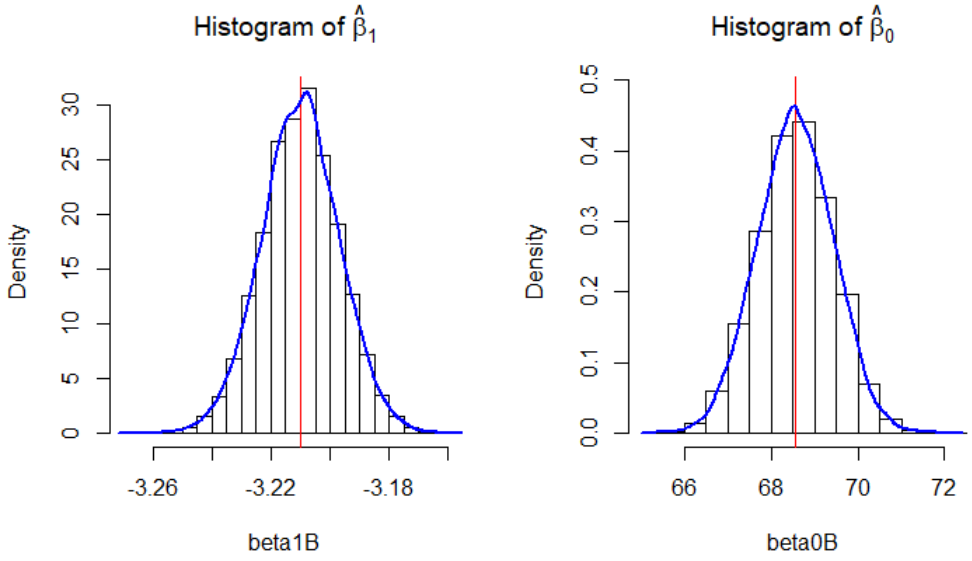

Figure 4.73: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions, $\sigma_e = 10$, $\sigma_0 = 8.07$, $\beta_1 = -3.21$, $\beta_0 = 68.57$

Table 4.77: *Summary of Simulation by Method II* : $\beta_1 = -3.21$, $\beta_0 = 68.57$ $\sigma_e = 10$, $\sigma_0 = 8.07$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	68.568	1.7×10^{-5}	0.742	[-3.236, -3.184]	[66.895, 70.226]
5000	-3.210	68.576	1.7×10^{-4}	0.782	[-3.235, -3.184]	[66.818, 70.258]
10000	-3.210	68.564	1.7×10^{-4}	0.756	[-3.236, -3.184]	[66.859, 70.223]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is 68.564, with the MSE to be 0.756.

(6) When $\boldsymbol{\beta}_1 = -3.21$, $\boldsymbol{\beta}_0 = -43.29$, we obtain the histograms of the estimated $\hat{\beta}_1$ and $\hat{\beta}_0$ values shown in Figure 4.74. Table 4.78 provides the overall estimates for each parameter, the resulting MSEs, and 95% confidence intervals for $\hat{\beta}_1$ and $\hat{\beta}_0$, based on each of the $B = 2000, 5000, 10000$ numbers of repetitions.

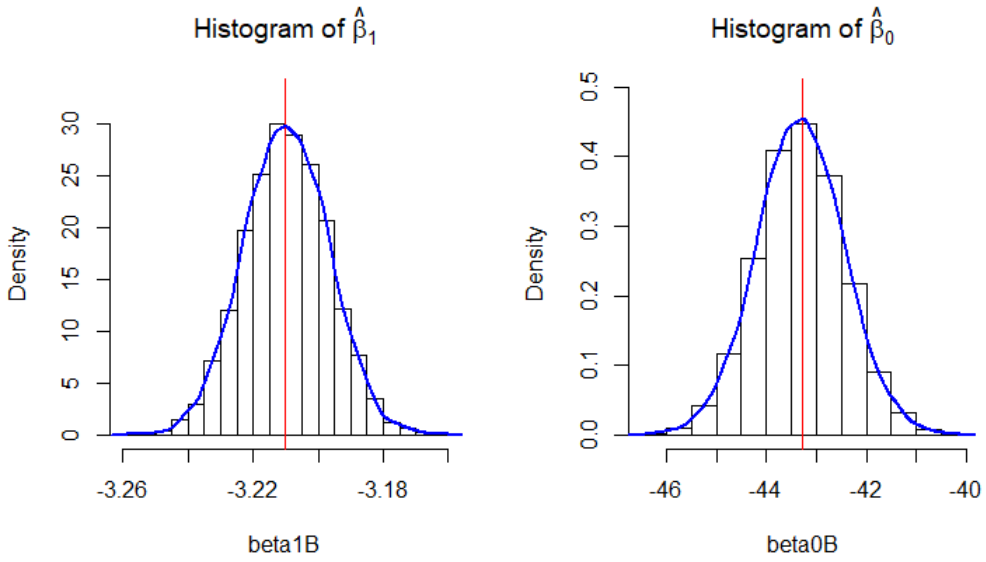

Figure 4.74: Histograms of $\hat{\beta}_1$ and $\hat{\beta}_0$ with 10000 repetitions
$\sigma_e = 10$, $\sigma_0 = 8.07$, $\beta_1 = -3.21$, $\beta_0 = -43.29$

Table 4.78: *Summary of Simulation by Method II* : $\beta_1 = -3.21$, $\beta_0 = -43.29$

B	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$	95% $C.I.$ of $\hat{\beta}_1$	95% $C.I.$ of $\hat{\beta}_0$
2000	-3.210	-43.298	1.7×10^{-4}	0.751	[-3.236, -3.185]	[-44.960, -41.593]
5000	-3.210	-43.287	1.7×10^{-4}	0.774	[-3.236, -3.185]	[-44.975, -41.550]
10000	-3.210	-43.311	1.7×10^{-4}	0.763	[-3.235, -3.185]	[-45.036, -41.577]

The average value of $\hat{\beta}_1$ for 10000 repetitions is -3.210, with the MSE to be 1.7×10^{-4}; the average value of $\hat{\beta}_0$ for 10000 repetitions is -43.311, with the MSE to be 0.763.

As in Section 4.2.1, to compare the performances of the proposed method under the second simulation approach, for different settings of the standard deviation of the error term, σ_e, and the standard deviation of the truncated normal distribution from which the interval ranges $X^{(r)}$ are generated, σ_0, we create the following six tables to summarize the averages of $\hat{\beta}_1$ and $\hat{\beta}_0$ as well as $MSE(\hat{\beta}_1)$ and $MSE(\hat{\beta}_0)$. Each table corresponds to a setting of the slope and the intercept parameter values (β_1, β_0), and with 10000 repetitions.

I. When $(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_0) = (0.64,\ 68.57)$, the results are summarized in Table 4.79.

Table 4.79: $\hat{\beta}_1$, $\hat{\beta}_0$, $MSE(\hat{\beta}_1)$, $MSE(\hat{\beta}_0)$ for $(\beta_1,\ \beta_0) = (0.64,\ 68.57)$

σ_e	σ_0	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$
3	5.25	0.640	68.570	2×10^{-5}	0.070
	8.07	0.640	68.568	2×10^{-5}	0.069
7	5.25	0.640	68.564	8×10^{-5}	0.378
	8.07	0.640	68.574	9×10^{-5}	0.373
10	5.25	0.640	68.569	1.7×10^{-4}	0.760
	8.07	0.640	68.569	1.7×10^{-4}	0.766

II. When $(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_0) = (0.64,\ -43.29)$, the results are summarized in Table 4.80.

Table 4.80: $\hat{\beta}_1$, $\hat{\beta}_0$, $MSE(\hat{\beta}_1)$, $MSE(\hat{\beta}_0)$ for $(\beta_1,\ \beta_0) = (0.64,\ -43.29)$

σ_e	σ_0	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$
3	5.25	0.640	-43.287	2×10^{-5}	0.068
	8.07	0.640	-43.292	2×10^{-5}	0.067
7	5.25	0.640	-43.290	8×10^{-5}	0.370
	8.07	0.640	-43.294	8×10^{-5}	0.369
10	5.25	0.640	-43.280	1.7×10^{-4}	0.783
	8.07	0.640	-43.290	1.7×10^{-4}	0.767

III. When $(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_0) = (2.15,\ 68.57)$, the results are summarized in Table 4.81.

Table 4.81: $\hat{\beta}_1$, $\hat{\beta}_0$, $MSE(\hat{\beta}_1)$, $MSE(\hat{\beta}_0)$ for $(\beta_1,\ \beta_0) = (2.15,\ 68.57)$

σ_e	σ_0	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$
3	5.25	2.150	68.572	2×10^{-5}	0.070
	8.07	2.150	68.570	2×10^{-5}	0.068
7	5.25	2.150	68.571	9×10^{-5}	0.378
	8.07	2.150	68.569	8×10^{-5}	0.376
10	5.25	2.150	68.560	2×10^{-4}	0.763
	8.07	2.150	68.569	1.7×10^{-4}	0.764

IV. When $(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_0) = (2.15,\ -43.29)$, the results are summarized in Table 4.82.

Table 4.82: $\hat{\beta}_1$, $\hat{\beta}_0$, $MSE(\hat{\beta}_1)$, $MSE(\hat{\beta}_0)$ for $(\beta_1,\ \beta_0) = (2.15,\ -43.29)$

σ_e	σ_0	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$
3	5.25	2.150	-43.293	2×10^{-5}	0.070
	8.07	2.150	-43.288	2×10^{-5}	0.069
7	5.25	2.150	-43.287	8×10^{-5}	0.382
	8.07	2.150	-43.283	8×10^{-5}	0.371
10	5.25	2.150	-43.290	1.7×10^{-4}	0.771
	8.07	2.150	-43.293	1.7×10^{-4}	0.783

V. When $(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_0) = (-3.21,\ 68.57)$, the results are summarized in Table 4.83.

Table 4.83: $\hat{\beta}_1$, $\hat{\beta}_0$, $MSE(\hat{\beta}_1)$, $MSE(\hat{\beta}_0)$ for $(\beta_1,\ \beta_0) = (-3.21,\ 68.57)$

σ_e	$sigma_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$
3	5.25	-3.210	68.567	2×10^{-5}	0.068
	8.07	-3.210	68.569	2×10^{-5}	0.069
7	5.25	-3.210	68.564	8×10^{-5}	0.382
	8.07	-3.210	68.583	9×10^{-5}	0.374
10	5.25	-3.210	68.572	1.7×10^{-4}	0.748
	8.07	-3.210	68.564	1.7×10^{-4}	0.756

VI. When $(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_0) = (-3.21,\ -43.29)$, the results are summarized in Table 4.84.

Table 4.84: $\hat{\beta}_1$, $\hat{\beta}_0$, $MSE(\hat{\beta}_1)$, $MSE(\hat{\beta}_0)$ for $(\beta_1,\ \beta_0) = (-3.21,\ -43.29)$

σ_e	σ_0	$\hat{\beta}_1$	$\hat{\beta}_0$	$MSE(\hat{\beta}_1)$	$MSE(\hat{\beta}_0)$
3	5.25	-3.210	-43.291	2×10^{-5}	0.068
	8.07	-3.210	-43.294	2×10^{-5}	0.068
7	5.25	-3.210	-43.295	8×10^{-5}	0.372
	8.07	-3.210	-43.296	8×10^{-5}	0.374
10	5.25	-3.210	-43.277	1.7×10^{-4}	0.767
	8.07	-3.210	-43.311	1.7×10^{-4}	0.763

From Table 4.79 to Table 4.84, we can observe that

1) The average value of $\hat{\beta}_1$ is always equal to the true value of β_1, and the average value of $\hat{\beta}_0$ is consistently very close to the true value of β_0, with each of the settings for β_1, β_0, σ_e and σ_0. This indicates that the proposed method gives unbiased estimators for both the slope and the intercept parameters.

2) Under the same setting of (β_1, β_0) and σ_0, the MSEs of both $\hat{\beta}_1$ and $\hat{\beta}_0$ become larger as the value of σ_e is set larger.

3) Under the same setting of (β_1, β_0) and σ_e, the larger values set for σ_0 do not result in big differences on $MSE(\hat{\beta}_1)$ or $MSE(\hat{\beta}_0)$.

4) Under the same settings for σ_e and σ_0, the MSEs of $\hat{\beta}_1$ are almost the same to each other, no matter of whether the slope or the intercept parameters are set to be positive or negative; so do the MSEs of $\hat{\beta}_0$.

Comparing the simulation results in Tables 4.37 to 4.42 and the ones in Tables 4.79 to 4.84, we observe that for the same pairing of (β_1, β_0) and the same value of σ_e, the estimated average of $\hat{\beta}_1$, $\hat{\beta}_0$, and $MSE(\hat{\beta}_1)$, $MSE(\hat{\beta}_0)$ are all quite close to each other. Recalling Section 4.1, we know that the first simulation method guarantees the linear relationship between the expected interval response Y and the interval explanatory variable X, while it has the drawback that the range of the simulated y_i is always no less than the range of $x_i\beta_1$, $i = 1, ..., n$ and n is the sample size, which may not be true in the real world. For the second method, since the way it generates the interval explanatory variable satisfies the basic assumption that the distribution within each interval is uniform, and the lower and upper bound of each response interval are given by the minimum and maximum values of the m values of $\beta_0 + x_i\beta_1$, $i = 1, ..., n$, respectively, this method is closer to how interval data arise and avoids the drawback of the first method, though it is hard to guarantee the response follows a uniform distribution internally. Therefore, the two simulation methods can be considered complementary, in terms of their advantages and disadvantages.

From the above simulation results, we can see that by both of the two simulation approaches, the proposed method preforms very well in estimating the slope parameter β_1 and the intercept parameter β_0, with very small mean square errors. The properties of unbiasedness as well as normality for the coefficient estimators are also verified.

4.3 APPENDIX

Simulation Method I

R Function for the Simulation Studies Under Method I, and **R** code for the simulations

```
# Method I


sim_m1 <- function(B, beta1, beta0, X_mean, sig, n_sam, a, b, sigma_e){


  beta1B <- NULL
  beta0B <- NULL
  var_beta1B <- NULL
  var_beta0B <- NULL


  for (i in 1:B){
    # 1st, generate interval-valued X
    X_i <- NULL
    n_i <- NULL


    for (l in 1:length(X_mean)){
      n_i <- c(n_i, sample(n_sam, 1))
      X_i <- c(X_i, rnorm(n_i[l], X_mean[l], sig))
    }
    len_Xi <- length(X_i)


    r <- runif(len_Xi, a, b)
```

```
# 3. generate lower and upper points of X
X_Li <- X_i - 1/2 * r
X_Ui <- X_i + 1/2 * r

# 2nd, generate error terms based on assumption
e_Li <- NULL
e_Ui <- NULL

e_Li <- rnorm(len_Xi, 0, sigma_e)
e_Ui <- rnorm(len_Xi, 0, sigma_e)

if (beta1 >= 0){

Y_Li <- X_Li * beta1 + beta0 + e_Li
Y_Ui <- X_Ui * beta1 + beta0 + e_Ui

# Next, calculate beta1B, beta0B and variances for each time
# call function "est_ord"

beta1B <- c(beta1B, est_ord(X_Li, X_Ui, Y_Li, Y_Ui, sigma_e)[1])
beta0B <- c(beta0B, est_ord(X_Li, X_Ui, Y_Li, Y_Ui, sigma_e)[2])
var_beta1B <- c(var_beta1B, est_ord(X_Li, X_Ui, Y_Li,
                                    Y_Ui, sigma_e)[3])
var_beta0B <- c(var_beta0B, est_ord(X_Li, X_Ui, Y_Li,
                                    Y_Ui, sigma_e)[4])
}else{
```

```
    # Calculate beta1B, beta0B and variances for each time
    # call function "est_ord_n"

    Y_Li <- X_Ui * beta1 + beta0 + e_Li
    Y_Ui <- X_Li * beta1 + beta0 + e_Ui

    beta1B <- c(beta1B, est_ord_n(X_Li, X_Ui, Y_Li, Y_Ui, sigma_e)[1])
    beta0B <- c(beta0B, est_ord_n(X_Li, X_Ui, Y_Li, Y_Ui, sigma_e)[2])
    var_beta1B <- c(var_beta1B, est_ord_n(X_Li, X_Ui, Y_Li,
                                          Y_Ui, sigma_e)[3])
    var_beta0B <- c(var_beta0B, est_ord_n(X_Li, X_Ui, Y_Li,
                                          Y_Ui, sigma_e)[4])

    }

}

# Obtain all the outputs
beta1_hat <- mean(beta1B)
beta0_hat <- mean(beta0B)
var_beta1 <- mean(var_beta1B)
var_beta0 <- mean(var_beta0B)
MSE_beta1 <- mean((beta1B - beta1)^2)
MSE_beta0 <- mean((beta0B - beta0)^2)
```

```
# empirical Confidence Interval
cl_beta1B <- quantile(beta1B, probs = c(.025, .975))[1]
cl_beta0B <- quantile(beta0B, probs = c(.025, .975))[1]

cu_beta1B <- quantile(beta1B, probs = c(.025, .975))[2]
cu_beta0B <- quantile(beta0B, probs = c(.025, .975))[2]

# histograms h1 and h0
par(mfrow = c(1, 2))
y1hist <- hist(beta1B, plot = FALSE)
y0hist <- hist(beta0B, plot = FALSE)
highestDensity1 <- max(y1hist$density)
highestDensity2 <- max(y0hist$density)
h1 <- hist(beta1B, breaks = 20, freq = FALSE,
           ylim = c(0, highestDensity1 * 1.1),
           main = expression(paste("Histogram of ", hat(beta)[1])))
abline(v = beta1, col = "red")
lines(density(beta1B), lwd = 2, col = "blue")

h0 <- hist(beta0B, breaks = 20, freq = FALSE,
           ylim = c(0, highestDensity2 * 1.1),
           main = expression(paste("Histogram of ", hat(beta)[0])))
abline(v = beta0, col = "red")
lines(density(beta0B), lwd = 2, col = "blue")

res1 <- paste("avg beta1: ", round(beta1_hat, 3), "avg beta0: "
```

```
                , round(beta0_hat, 3), "avg var(beta1): "
                , round(var_beta1, 5), "avg var(beta0)"
                , round(var_beta0, 3), "MSE(beta1): "
                , round(MSE_beta1, 5), "MSE(beta0): "
                , round(MSE_beta0, 3), "C.I. for beta_1: ["
                , round(cl_beta1B, 3), ",", round(cu_beta1B, 3),"]",
                "C.I. for beta_0: [", round(cl_beta0B, 3), ",",
                round(cu_beta0B, 3),"]")

  res <- list(res1, h1, h0)
  return(res)
}

set.seed(1025)
# Generate (X_L, X_U) with the following settings:
sigma_e <- 10
a <- 6.5
b <- 9.25

# Generate the interval means

mu <- c(-35, -25, -15, -5, 5, 15, 25, 35, 45, 55, 65, 75, 85, 95,
        105, 115, 125)
N0 <- c(5, 6, 7, 8, 9)
sigma_e <- c(3, 7, 10)
a1 <- c(6.5, 10)
```

```
b1 <- c(9.25, 12.45)
beta1 <- c(.64, 2.15, -3.21)
beta0 <- c(68.57, -43.29)

B0 <- 10000

rec11 <- matrix(list(), 3, 2) # save all the results by different values of
                              # beta1 and beta0
for (i in 1:2){
  for (j in 1:3){
    rec11[[j,i]] <- sim_m1(B1, beta1[j], beta0[i], X_mean = mu, sig = 7,
                           n_sam = N0, a = 10, b = 12.45, sigma_e = 3)
  }
}

rec11[[1, 1]] # beta1 = .64, beta0 = 68.57
rec11[[1, 2]] # beta1 = .64, beta0 = -43.29
rec11[[2, 1]] # beta1 = 2.15, beta0 = 68.57
rec11[[2, 2]] # beta1 = 2.15 beta0 = -43.29
rec11[[3, 1]] # beta1 = -3.21, beta0 = 68.57
rec11[[3, 2]] # beta1 = -3.21, beta0 = -43.29

rec21 <- matrix(list(), 3, 2) # save all the results by different values of
# beta1 and beta0
for (i in 1:2){
  for (j in 1:3){
```

```
    rec21[[j,i]] <- sim_m1(B1, beta1[j], beta0[i], X_mean = mu, sig = 7,
                           n_sam = N0, a =  10, b = 12.45, sigma_e = 7)
  }
}

rec21[[1, 1]] # beta1 = .64, beta0 = 68.57
rec21[[1, 2]] # beta1 = .64, beta0 = -43.29
rec21[[2, 1]] # beta1 = 2.15, beta0 = 68.57
rec21[[2, 2]] # beta1 = 2.15 beta0 = -43.29
rec21[[3, 1]] # beta1 = -3.21, beta0 = 68.57
rec21[[3, 2]] # beta1 = -3.21, beta0 = -43.29



rec31 <- matrix(list(), 3, 2) # save all the results by different values of
# beta1 and beta0
for (i in 1:2){
  for (j in 1:3){
    rec31[[j,i]] <- sim_m1(B1, beta1[j], beta0[i], X_mean = mu, sig = 7,
                           n_sam = N0, a =  10, b = 12.45, sigma_e = 10)
  }
}

rec31[[1, 1]] # beta1 = .64, beta0 = 68.57
rec31[[1, 2]] # beta1 = .64, beta0 = -43.29
rec31[[2, 1]] # beta1 = 2.15, beta0 = 68.57
rec31[[2, 2]] # beta1 = 2.15 beta0 = -43.29
```

```
rec31[[3, 1]] # beta1 = -3.21, beta0 = 68.57
rec31[[3, 2]] # beta1 = -3.21, beta0 = -43.29



# Scatter plot as example
X_mean <- NULL
n_lev <- NULL

for (l in 1:length(mu)){

  n_lev <- c(n_lev, sample(N0, 1)) # consider if 1 need to be changed
  X_mean <- c(X_mean, rnorm(n_lev[l], mu[l], sig))
}

n_lev
len_X <- length(X_mean) # 124 points of X

# 2. generate interval ranges
r <- runif(len_X, a, b)

# 3. generate lower and upper points of X
X_L <- X_mean - 1/2 * r
X_U <- X_mean + 1/2 * r

# 4. generate the lower and upper points of the error term
e_L <- NULL
```

```
e_U <- NULL


e_L <- rnorm(len_X, 0, sigma_e)
e_U <- rnorm(len_X, 0, sigma_e)


# 5. generate lower and upper points of Y
Y_L <- beta0 + X_L * beta1 + e_L
Y_U <- beta0 + X_U * beta1 + e_U


# plot with regression line
par(mfrow = c(1, 1))
plot(c(min(X_L)-5,max(X_U)+5), c(min(Y_L)-20, max(Y_U)+20),
     type = "n", xlab = "", ylab = "",
     main = expression(paste("Scatter Plots with Regression Line, ",
     hat(beta)[1] == 2.150, " , ", hat(beta)[0] == 68.574)))
rect(X_L, Y_L, X_U, Y_U, border = "blue")
clip(min(X_L)-5, max(X_U)+5, min(Y_L)-10, max(Y_U)+10)
abline(a = 67.574, b = 2.150, lwd = "2", col = "red")
```

Simulation Method II

R Function for the Simulation Studies Under Method II, and **R** code for the simulations

```
# Simulation: Method II
library(truncnorm)
```

```
# Input:
# B: times of repeats
# beta1: true value of \beta1
# beta0: true value of \beta0
# X_mean: mu for X interval means
# sig: standard deviation for X interval means
# n_sam: vector for number of X around each mu to be sampled
# a, b, mu0: parameter of truncated normal distribution to
# generate interval range
# sigma0: sd of truncated normal distribution
# sigma_e: standard deviation of the error term
# m0: number of values generated form each X interval

# Output:
# beta1_hat: average point estimate of \beta1
# beta0_hat: average point estimate of \beta0
# var_beta1: average variance of beta1B
# var_beta0: average variance of beta0B
# MSE_beta1: mean square error of beta1B
# MSE_beta0: mean square error of beta0B
# h1: histogram of beta1_B
# h0: histogram of beta0_B

sim_m2 <- function(B, beta1, beta0, X_mean, sig, n_sam, a, b, mu0,
                   sigma0, sigma_e, m0 = 3000){
```

```
beta1B <- NULL
beta0B <- NULL
var_beta1B <- NULL
var_beta0B <- NULL

for (i in 1:B){
  # 1st, generate interval-valued X
  X_i <- NULL
  n_i <- NULL

  for (l in 1:length(X_mean)){
    n_i <- c(n_i, sample(n_sam, 1))
    X_i <- c(X_i, rnorm(n_i[l], X_mean[l], sig))
  }
  len_Xi <- length(X_i)

  r2 <- rtruncnorm(len_Xi, a, b, mu0, sigma0)

  # generate lower and upper bounds of X
  X_Li <- X_i - 1/2 * r2
  X_Ui <- X_i + 1/2 * r2

  # generate lower and upper points of Y

  # generate the lower and upper points of the error term
  # draw m random samples from U(X_Li, X_Ui)
```

```
Y_Li0 <- NULL
Y_Ui0 <- NULL

for (j in 1:len_Xi){

  X_jl <- runif(m0, X_Li[j], X_Ui[j])

  Y_Li0 <- c(Y_Li0, beta0 + min(X_jl * beta1))
  Y_Ui0 <- c(Y_Ui0, beta0 + max(X_jl * beta1))
}

# 2nd, generate error terms based on assumption
e_Li <- NULL
e_Ui <- NULL

e_Li <- rnorm(len_Xi, 0, sigma_e)
e_Ui <- rnorm(len_Xi, 0, sigma_e)

Y_Li <- Y_Li0 + e_Li
Y_Ui <- Y_Ui0 + e_Ui

# Next, calculate beta1B, beta0B and variances for each time
# call function "est_ord"

if (beta1 >= 0){
```

```
  beta1B <- c(beta1B, est_ord(X_Li, X_Ui, Y_Li, Y_Ui, sigma_e)[1])
  beta0B <- c(beta0B, est_ord(X_Li, X_Ui, Y_Li, Y_Ui, sigma_e)[2])
  var_beta1B <- c(var_beta1B, est_ord(X_Li, X_Ui, Y_Li,
                                      Y_Ui, sigma_e)[3])
  var_beta0B <- c(var_beta0B, est_ord(X_Li, X_Ui, Y_Li,
                                      Y_Ui, sigma_e)[4])
}else{
  beta1B <- c(beta1B, est_ord_n(X_Li, X_Ui, Y_Li, Y_Ui, sigma_e)[1])
  beta0B <- c(beta0B, est_ord_n(X_Li, X_Ui, Y_Li, Y_Ui, sigma_e)[2])
  var_beta1B <- c(var_beta1B, est_ord_n(X_Li, X_Ui, Y_Li,
                                      Y_Ui, sigma_e)[3])
  var_beta0B <- c(var_beta0B, est_ord_n(X_Li, X_Ui, Y_Li,
                                      Y_Ui, sigma_e)[4])
  }
}
# Obtain all the outputs
beta1_hat <- mean(beta1B)
beta0_hat <- mean(beta0B)
var_beta1 <- mean(var_beta1B)
var_beta0 <- mean(var_beta0B)
MSE_beta1 <- mean((beta1B - beta1)^2)
MSE_beta0 <- mean((beta0B - beta0)^2)

# empirical Confidence Interval
cl_beta1B <- quantile(beta1B, probs = c(.025, .975))[1]
```

```
cl_beta0B <- quantile(beta0B, probs = c(.025, .975))[1]

cu_beta1B <- quantile(beta1B, probs = c(.025, .975))[2]
cu_beta0B <- quantile(beta0B, probs = c(.025, .975))[2]

# histograms h1 and h0
par(mfrow = c(1, 2))
y1hist <- hist(beta1B, plot = FALSE)
y0hist <- hist(beta0B, plot = FALSE)
highestDensity1 <- max(y1hist$density)
highestDensity2 <- max(y0hist$density)
h1 <- hist(beta1B, breaks = 20, freq = FALSE,
           ylim = c(0, highestDensity1 * 1.1),
           main = expression(paste("Histogram of ", hat(beta)[1])))
abline(v = beta1, col = "red")
lines(density(beta1B), lwd = 2, col = "blue")

h0 <- hist(beta0B, breaks = 20, freq = FALSE,
           ylim = c(0, highestDensity2 * 1.1),
           main = expression(paste("Histogram of ", hat(beta)[0])))
abline(v = beta0, col = "red")
lines(density(beta0B), lwd = 2, col = "blue")

res2 <- paste("avg beta1: ", round(beta1_hat, 3), "avg beta0: "
              , round(beta0_hat, 3), "avg var(beta1): "
              , round(var_beta1, 5), "avg var(beta0)"
```

```
                , round(var_beta0, 3), "MSE(beta1): "
                , round(MSE_beta1, 5), "MSE(beta0): "
                , round(MSE_beta0, 3), "C.I. for beta_1: ["
                , round(cl_beta1B, 3), ",", round(cu_beta1B, 3),"]",
                "C.I. for beta_0: [", round(cl_beta0B, 3), ",",
                round(cu_beta0B, 3),"]")

  res <- list(res2, h1, h0)
  return(res)
}

# Simulation results of all different settings by Method II

sigma_e <- c(3, 7, 10)
a <- 9.43
b <- 13.69
mu0 <- 11.12
sigma0 <- c(5.25, 8.07)
beta1 <- c(.64, 2.15, -3.21)
beta0 <- c(68.57, -43.29)

B1 = 10000

rec21 <- matrix(list(), 3, 2) # save all the results by different values of
# beta1 and beta0
for (i in 1:2){
```

```
  for (j in 1:3){
    rec21[[j,i]] <- sim_m2(B1, beta1[j], beta0[i], X_mean = mu, sig = 7,
                           n_sam = N0, a = 9.43, b = 13.69, mu0 = 11.12,
                           sigma0 = 5.25, sigma_e = 10)
  }
}

rec21[[1, 1]] # beta1 = .64, beta0 = 68.57
rec21[[1, 2]] # beta1 = .64, beta0 = -43.29
rec21[[2, 1]] # beta1 = 2.15, beta0 = 68.57
rec21[[2, 2]] # beta1 = 2.15, beta0 = -43.29
rec21[[3, 1]] # beta1 = -3.21, beta0 = 68.57
rec21[[3, 2]] # beta1 = -3.21, beta0 = -43.29

rec22 <- matrix(list(), 3, 2) # save all the results by different values of
# beta1 and beta0
for (i in 1:2){
  for (j in 1:3){
    rec22[[j,i]] <- sim_m2(B1, beta1[j], beta0[i], X_mean = mu, sig = 7,
                           n_sam = N0, a = 9.43, b = 13.69, mu0 = 11.12,
                           sigma0 = 8.07, sigma_e = 10)
  }
}

rec22[[1, 1]] # beta1 = .64, beta0 = 68.57
rec22[[1, 2]] # beta1 = .64, beta0 = -43.29
```

```
rec22[[2, 1]] # beta1 = 2.15, beta0 = 68.57
rec22[[2, 2]] # beta1 = 2.15, beta0 = -43.29
rec22[[3, 1]] # beta1 = -3.21, beta0 = 68.57
rec22[[3, 2]] # beta1 = -3.21, beta0 = -43.29

# Scatter plot as example

set.seed(1025)

# 1. generate interval ranges
a <- 9.43
b <- 13.69
mu0 <- 11.21
sigma0 <- 5.25
r2 <- rtruncnorm(len_X, a, b, mu0, sigma0)

# 2. generate lower and upper points of X
X_L <- X_mean - 1/2 * r2
X_U <- X_mean + 1/2 * r2

m <- 3000

# 3. generate lower and upper points of Y

# generate the lower and upper points of the error term
# draw m random samples from U(X_Li, X_Ui)
```

```
sigma_e <- 10

Y_L0 <- NULL
Y_U0 <- NULL

for (i in 1:len_X){

  X_il <- runif(m, X_L[i], X_U[i])

  Y_L0 <- c(Y_L0, beta0 + min(X_il * beta1))
  Y_U0 <- c(Y_U0, beta0 + max(X_il * beta1))
}

e_L <- NULL
e_U <- NULL

e_L <- rnorm(len_X, 0, sigma_e)
e_U <- rnorm(len_X, 0, sigma_e)

  Y_L <- Y_L0 + e_L
  Y_U <- Y_U0 + e_U

# Scatter plot of X and Y
## set up the plot region:
plot(c(min(X_L)-5,max(X_U)+5), c(min(Y_L)-20, max(Y_U)+20),
```

```
     type = "n", xlab = "", ylab = "")
rect(X_L, Y_L, X_U, Y_U, border = "blue")

# plot with regression line
par(mfrow = c(1, 1))
plot(c(min(X_L)-5,max(X_U)+5), c(min(Y_L)-20, max(Y_U)+20),
     type = "n", xlab = "", ylab = "",
     main = expression(paste("Scatter Plots with Regression Line, ",
                    hat(beta)[1] == 2.150, " , ", hat(beta)[0] == 68.573)))
rect(X_L, Y_L, X_U, Y_U, border = "blue")
clip(min(X_L)-5, max(X_U)+5, min(Y_L)-10, max(Y_U)+10)
abline(a = 68.573, b = 2.150, lwd = "2", col = "red")
```

Chapter 5

REAL DATA APPLICATION

In Chapter 5, we apply the proposed method on two real data sets to conduct statistical inference and evaluate the model performances. The first data set contains information about price, velocity, acceleration and cylinder capacity for eight different car models. It is an example of inherently interval-valued data. The second data set records measurements for three features of 100 species of mushrooms. It is an example of interval-valued data arising by data aggregation.

5.1 EXAMPLE I: CARS data set

Data Description

First, we apply the proposed approach to the cars data set. The data set is referred from Billard and Diday (2006) [1]. In this data set, measurements of eight different car models are recorded. There are four interval-valued variables, namely Y = Price ($\times 10^{-3}$, in euros), X_1 = Maximum Velocity, X_2 = Acceleration Time to reach a given speed, and X_3 = Cylinder Capacity of the car. The data set is given in Table 5.1.

Table 5.1: Cars data set

Car	Y = Price	X_1 = Maximum Velocity	X_2 = Acceleration Time	X_3 = Cylinder Capacity
Aston Martin	[260.5, 460.0]	[298, 306]	[4.7, 5.0]	[5935, 5935]
Audi A6	[68.2, 140.3]	[216, 250]	[6.7, 9.7]	[1781, 4172]
Audi A8	[123.8, 171.4]	[232, 250]	[5.4, 10.1]	[2771, 4172]
BMW 7	[104.9, 276.8]	[228, 240]	[7.0, 8.6]	[2793, 5397]
Ferrari	[240.3, 391.7]	[295, 298]	[4.5, 5.2]	[3586, 5474]
Honda NSR	[205.2, 215.2]	[260, 270]	[5.7, 6.5]	[2977, 3179]
Mercedes C	[55.9, 115.2]	[210, 250]	[5.2, 11.0]	[1998, 3199]
Porsche	[147.7, 246.4]	[280, 305]	[4.2, 5.2]	[3387, 3600]

We first conduct an explanatory analysis to discover relations between the interval-valued variables in the data set. As the first step, by (2.3) and (2.8), we compute the symbolic variance-covariance matrix. The result is as follows:

$$V = \begin{pmatrix} 2591.93 & 901.80 & 6.43 & 34756.76 \\ 901.80 & 397.96 & -3.96 & 13013.41 \\ 6.43 & -3.96 & 0.71 & 30.77 \\ 34756.76 & 13013.41 & 30.77 & 606978.69 \end{pmatrix}. \tag{5.1}$$

The variance-covariance matrix in (5.1) is of the order (Y, X_1, X_2, X_3). By the matrix V, we can see that the variances of the four variables are very different, ranging from as small as 0.71 for X_2 Acceleration Time, to as large as 606978.69 for X_3 Cylinder Capacity, with the variance for X_1, Maximum Velocity, to be 397.96, and the variance for Y, Price, to be 2591.93. The huge difference between the variances is largely because of different measurement units

for each of the four variables.

In addition, by (5.1) and (2.9), we can also compute the symbolic correlation matrix as

$$R = \begin{pmatrix} 1.000 & 0.888 & 0.149 & 0.876 \\ 0.888 & 1.000 & -0.235 & 0.837 \\ 0.149 & -0.235 & 1.000 & 0.047 \\ 0.876 & 0.837 & 0.047 & 1.000 \end{pmatrix}. \tag{5.2}$$

From (5.2), the correlations among Maximum Velocity, Cylinder Capacity and Price are all greater than 0.8, indicating they are positively, highly correlated to each other. The coefficient of correlation between Price and Maximum Velocity is 0.888. The coefficient of correlation between Price and Cylinder Capacity is 0.876. Meanwhile, the correlation matrix shows the correlation coefficient between Price and Acceleration Time is 0.149, indicating very weak correlation between these two variables. The only negative correlation coefficient is the one between Maximum Velocity and Acceleration Time, which is -0.235, indicating weak negative correlation between these two variables.

By the variance-covariance matrix in (5.1), and the correlation matrix in (5.2), we consider building two simple regression models, both of which are with Y = Price as the response. For the first model, we use X_1 = Max Velocity as the explanatory variable; for the second model, we use X_3 = Cylinder Capacity as the explanatory variable.

Data Analysis

First, we use Y = Price as the response variable and X_1 = Max Velocity as the explanatory variable to build the simple regression model.

By the proposed approach illustrated in Sections 3.2.2, we build the model by the following steps.

1. We determine the sign of the slope parameter β_1 by checking the correlation between Price (Y) and Max Velocity (X_1). By the correlation matrix in (5.2), the coefficient of correlation is 0.888, which is greater than zero. By Section 3.4, we judge that $\beta_1 \geq 0$ and use the corresponding likelihood function in (3.50) to fit the model.

2. We give the point estimators for the intercept parameter β_0 and the slope parameter β_1. Through calculation by (3.54) and (3.55), we have $\hat{\beta}_0 = -509.115$, $\hat{\beta}_1 = 2.715$. Therefore, by substituting the values of $\hat{\beta}_0$ and $\hat{\beta}_1$ to (3.5), the regression model is given by

$$[Y_L,\ Y_U] = -509.115 + 2.715 \times [X_{1L},\ X_{1U}]. \tag{5.3}$$

3. We estimate the standard deviation of the error term by (3.78) and (3.79), which gives $\hat{\sigma}_e = 52.835$.

Next, by Section 3.2.3, Section 3.3 and Section 3.4, we give confidence intervals for the estimated regression coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$, predictions of the response, and the measurement of the model fit.

4. By (3.70) and (3.71), substituting the standard deviation of the error term by the value of $\hat{\sigma}_e$ obtained in Step 3, we have the estimated variances of $\hat{\beta}_0$ and $\hat{\beta}_1$ as: $Var(\hat{\beta}_0) = 0.178$ and $Var(\hat{\beta}_1) = 12396.195$, respectively.

5. By (3.80) and (3.81), with all the values of $\hat{\beta}_0$, $\hat{\beta}_1$, $Var(\hat{\beta}_0)$, and $Var(\hat{\beta}_1)$ obtained through Step 2 to 4, the 95% confidence intervals for β_0 and β_1 are as follows:

$$\beta_0 \in [-781.549,\ -236.68],\ \beta_1 \in [1.681,\ 3.748]. \tag{5.4}$$

6. The predicted response value, together with the 95% confidence intervals for the lower bound and the upper bound of the response of a new observation can be given by (3.83), (3.87) and (3.88). Suppose we have another car model, say, its Maximum Velocity is in the

interval [273, 285]. Then we can predict the price range for the vehicle using Equation (3.83):

$$\hat{Y}_L = -509.115 + 2.715 \times 273 = 232.08, \tag{5.5}$$

$$\hat{Y}_U = -509.115 + 2.715 \times 285 = 264.66. \tag{5.6}$$

By (3.87), the 95% confidence interval for $\hat{Y}_L$ is:

$$\hat{Y}_L \in [-160.13, 624.29]; \tag{5.7}$$

and by (3.88), the 95% confidence interval for $\hat{Y}_U$ is:

$$\hat{Y}_U \in [-136.56, 665.88]. \tag{5.8}$$

From (5.8) and (5.9), we can observe that the confidence intervals for $\hat{Y}_L$ and $\hat{Y}_U$ are too wide, with the lower bounds smaller than zero, which are not valid values for the response Price. Recalling from classical regression, the same problem that the prediction intervals exceed a reasonable range may also exist.

7. To measure the model fit, by Section 3.4, we calculate the predicted value for each of the observations in the data set by (5.3), and based on the predicted values we calculate the residuals for the lower and upper bounds of the response by (3.42) and (3.43), respectively. The predicted values of Price (Y), the residuals, as well as the real values of Price (Y) are displayed in Table 5.2:

Table 5.2: Predictions and Residuals for $Y = Price$, $X_1 = Max\ Velocity$, Cars data set

i	Response (Y)		Prediction ($\hat{Y}$)		Residual (r)	
	Y_{iL}	Y_{iU}	$\hat{Y}_{iL}$	$\hat{Y}_{iU}$	r_{iL}	r_{il}
1	260.5	460.0	300.0	321.68	-39.46	138.33
2	68.2	140.3	77.33	169.64	-9.13	-29.34
3	123.8	171.4	120.77	169.64	3.04	1.77
4	104.9	276.8	109.91	142.49	-5.01	134.32
5	240.3	391.7	291.81	299.96	-51.51	91.75
6	205.2	215.2	196.79	223.94	8.42	-8.74
7	55.9	115.2	61.04	169.64	-5.14	-54.44
8	147.7	246.4	251.09	318.96	-103.39	-72.56

From the data values in Table 5.2, by (3.91), the R-square value is

$$R^2 = \frac{Var(\hat{Y})}{Var(Y)} = 0.715. \tag{5.9}$$

By (5.9), $R^2 = 0.715$, which indicates that 71.5% of the total variance for the response is explained by the model.

Then, as the second simple regression model, we use $Y =$ Price as the response variable and $X_3 =$ Cylinder Capacity as the explanatory variable.

Similar to the model building procedure for the first model, in Step 1, since the correlation between $Y =$ Price and $X_3 =$ Cylinder Capacity is positive, we judge that $\beta_1 > 0$, and use the likelihood function in (3.50) to calculate maximum likelihood estimators (MLEs) of the slope and the intercept parameters.

In Step 2, we calculate the point estimators by (3.54) and (3.55), which gives $\hat{\beta}_0 = -66.941$, and $\hat{\beta}_1 = 0.071$. Therefore, by substituting the values of $\hat{\beta}_0$ and $\hat{\beta}_1$ to (3.5), the regression model is given by

$$[Y_L,\ Y_U] = -66.941 + 0.071 \times [X_{3L},\ X_{3U}]. \tag{5.10}$$

In step 3, by (3.78) and (3.79), the standard deviation of the error term is estimated as $\hat{\sigma}_e = 46.392$.

In step 4, we estimate the variances for $\hat{\beta}_0$ and $\hat{\beta}_1$, by (3.70) and (3.71), respectively. We have $Var(\hat{\beta}_0) = 1321.092$, $Var(\hat{\beta}_1) = 1 \times 10^{-4}$.

In step 5, with all the values of $\hat{\beta}_0$, $\hat{\beta}_1$, $Var(\hat{\beta}_0)$, and $Var(\hat{\beta}_1)$ obtained in the previous steps, we have the 95% confidence intervals for β_0 and β_1 to be: $\beta_0 \in [-155.878,\ 21.997]$, $\beta_1 \in [0.049,\ 0.093]$.

In step 6, suppose we have a new car model with the Cylinder Capacity to be [3615, 4279]. By (3.83), the predicted lower bound and upper bound of the price for this model are

$$\hat{Y}_L = -66.941 + 0.071 \times 3615 = 189.72, \tag{5.11}$$

$$\hat{Y}_U = -66.941 + 0.071 \times 4279 = 236.87. \tag{5.12}$$

By (3.87) and (3.88), the 95% confidence intervals for $\hat{Y}_L$ and $\hat{Y}_U$ are:

$$\hat{Y}_L \in [64.28, 315.16],\ \hat{Y}_U \in [99.49, 374.25]. \tag{5.13}$$

In step 7, to measure the model fit, we calculate the predicted price for each of the car models in the data set by (5.3), and calculate the residuals for the lower and upper bounds of the price by (3.42) and (3.43), respectively. Then the R-square value, calculated by (3.91), is 0.662.

5.2 EXAMPLE II: MUSHROOM data set

Data Description

As the second example, we apply the proposed approach to the mushroom data set to see its performance. The data set records measurements of three features of 100 species of mushrooms. All the measurements are interval-valued and they are extracted from the Fungi of California Species Index. The original data set can be accessed at http://www.myknoweb.com /CAF/species_index.html. There are three interval-valued variables, namely Y_1 = Pileus Cap Width, Y_2 = Stipe Length, Y_3 = Stipe Thickness. There are in total 274 observations recorded in the data set, which can be found in Xu's dissertation (2010, page 111) [8]. To calculate the variance and covariance matrix and build the simple regression model, we remove the 10 observations with missing values.

First, we compute the sample variance and covariance matrix by (2.3) and (2.8). The result is as follows:

$$V = \begin{pmatrix} 3.89 & -0.23 & 5.13 \\ -0.23 & 5.87 & 1.37 \\ 5.13 & 1.37 & 10.62 \end{pmatrix}. \tag{5.14}$$

The variance-covariance matrix in (5.14) is of the order (Y_1, Y_2, Y_3). From the matrix V, we can observe that the variance of Y_3, Stipe Thickness, has the largest variance, which is 10.62, the variance of Y_1, Pileus Cap Width, is the smallest, which is 3.89. For Y_2, Stipe Length, the variance is 5.87.

By (5.14) and (2.9), we can also compute the symbolic correlation matrix as

$$R = \begin{pmatrix} 1.000 & -0.048 & 0.798 \\ -0.048 & 1.000 & 0.173 \\ 0.798 & 0.173 & 1.000 \end{pmatrix}. \tag{5.15}$$

From (5.15), Pileus Cap Width (Y_1) and Stipe Thickness (Y_3) have the largest correlation coefficient, which is 0.798, and the coefficient of correlation between Stipe Length (Y_2) and Stipe Thickness (Y_3) is 0.173. The coefficient of correlation between Pileus Cap Width (Y_1) and Stipe Length(Y_2) is -0.048.

By the variance-covariance matrix in (5.14), and the correlation matrix in (5.15), we consider building a simple regression model with Y_1 = Pileus Cap Width as the response, and Y_3 = Stipe Thickness as the explanatory variable.

Data Analysis

By the proposed approach illustrated in Section 3.2.2, we build the simple regression model:

$$[Y_1L,\ Y_1U] = \beta_0 + \beta_1[Y_{3L},\ Y_{3U}] + \epsilon \tag{5.16}$$

where $[Y_{1L}, Y_{1U}]$ = Pileus Cap Width and $[Y_{3L}, Y_{3U}]$ = Stipe Thickness.

As the first step, we judge that $\beta_1 > 0$ since the correlation coefficient for Y_1 and Y_3 is positive. Therefore, we use the likelihood function in (3.50) to compute MLEs of the slope and the intercept parameters.

Then, for step 2, we calculate the point estimators by (3.54) and (3.55). We have $\hat{\beta}_0 = -1.692$, and $\hat{\beta}_1 = 0.575$. Therefore, the regression model is

$$[Y_{1L},\ Y_{1U}] = -1.692 + 0.575 \times [Y_{3L},\ Y_{3U}]. \tag{5.17}$$

In step 3, we estimate the standard deviation of the error term by (3.78) and (3.79), which gives $\hat{\sigma}_e = 2.716$.

Next, we estimate the variances for $\hat{\beta}_0$ and $\hat{\beta}_1$ by (3.70) and (3.71), respectively, in step 4. We have $Var(\hat{\beta}_0) = 0.066$, $Var(\hat{\beta}_1) = 3 \times 10^{-4}$.

For step 5, we give the 95% confidence intervals for β_0 and β_1 by (3.80) and (3.81), with values of $\hat{\beta}_0$, $\hat{\beta}_1$, $Var(\hat{\beta}_0)$ and $Var(\hat{\beta}_1)$ obtained above, which gives

$$\beta_0 \in [-2.197,\ -1.186], \beta_1 \in [0.542,\ 0.608]. \tag{5.18}$$

In step 6, we provide the predicted lower bound and upper bound of the Pileus Cap Width (Y_1) for a new species with the Stipe Thickness (Y_3) value $[10, 19]$ by (3.83). We have

$$\hat{Y}_{1L} = -1.692 + 0.575 \times 10 = 4.06, \tag{5.19}$$

$$\hat{Y}_{1U} = -1.692 + 0.575 \times 19 = 9.23. \tag{5.20}$$

By (3.87) and (3.88), the 95% confidence intervals for $\hat{Y}_{1L}$ and $\hat{Y}_{1U}$ are:

$$\hat{Y}_{1L} \in [3.45, 4.67],\ \hat{Y}_{1U} \in [8.41, 10.05]. \tag{5.21}$$

To measure the model fit in step 7, by (5.3), we compute the predicted Pileus Cap Width for each of the observations in the data set, and by (3.42) and (3.43), we compute the residuals for the lower and upper bounds of the Pileus Cap Width, respectively. Then, the R-square value, calculated by (3.91), is 0.759.

5.3 APPENDIX

R code for data analyses on Cars data set and Mushroom data set

```
# 1st Model: Price (Y) ~ Max Velocity (X1)
# scatter plot

plot(c(min(mv$at_l)-50,max(mv$at_u)+50), c(min(price$mv_l)-100,
      max(price$mv_u)+100), type = "n", xlab = "", ylab = "")
rect(mv$at_l, price$mv_l, mv$at_u, price$mv_u, border = "blue")

m1 <- est_ord1(mv$at_l, mv$at_u, price$mv_l, price$mv_u)

# prediction: X_1 = [273, 285]
-509.115 + 2.715*273
-509.115 + 2.715*285

# 95% CI for Y_L
var_beta0 <- 12396.195
var_beta1 <- 0.1784

lower_YL <- 232.08 - qt(.975, 8-2)*sqrt(var_beta0 + var_beta1 * 273^2)
upper_YL <- 232.08 + qt(.975, 8-2)*sqrt(var_beta0 + var_beta1 * 273^2)

# 95% CI for Y_U
lower_YU <- 264.66 - qt(.975, 8-2)*sqrt(var_beta0 + var_beta1 * 285^2)
upper_YU <- 264.66 + qt(.975, 8-2)*sqrt(var_beta0 + var_beta1 * 285^2)
```

```
# predicted Y
pred1_YL <- -509.115 + 2.715*mv$at_l
pred1_YU <- -509.115 + 2.715*mv$at_u

pred1_YL
pred1_YU

# Residuals
r1_L <- price$mv_l - pred1_YL
r1_U <- price$mv_u - pred1_YU



# 2nd Model: Price (Y) ~ Cylinder Capacity (X3)
# scatter plot

plot(c(min(cc$w_l)-100,max(cc$w_u)+100), c(min(price$mv_l)-100,
     max(price$mv_u)+100), type = "n", xlab = "", ylab = "")
rect(cc$w_l, price$mv_l, cc$w_u, price$mv_u, border = "blue")

m2 <- est_ord1(cc$w_l, cc$w_u, price$mv_l, price$mv_u)

# prediction: X_3 = [3615, 4279]
-66.941 + 0.071*3615
-66.941 + 0.071*4279
```

```
# 95% CI for Y_L
var_beta0 <- 1321.0917
var_beta1 <- 0.0001

lower_YL <- 189.72 - qt(.975, 8-2)*sqrt(var_beta0 + var_beta1 * 3615^2)
upper_YL <- 189.72 + qt(.975, 8-2)*sqrt(var_beta0 + var_beta1 * 3615^2)

# 95% CI for Y_U
lower_YU <- 236.87 - qt(.975, 8-2)*sqrt(var_beta0 + var_beta1 * 4279^2)
upper_YU <- 236.87 + qt(.975, 8-2)*sqrt(var_beta0 + var_beta1 * 4279^2)



# 2nd data set: mushroom
setwd("C:/Users/Colin Cai/Documents/RESEARCH/Desertation/1st paper/data")

# use this one
mushroom1 <- read.table("mushroomsALL.dat", header = FALSE)
head(mushroom1)
dim(mushroom1)

mushroom1 <- mushroom1[,3:8]
head(mushroom1)

# remove missing values
mushroom[which(mushroom[,3]=="."),]
mushroom[which(mushroom[,4]="."),]
```

```
mushroom <- mushroom1[-c(3, 31, 57, 59, 184, 194, 198, 202, 210, 218),]
dim(mushroom) # 264*6

# assign variable names
colnames(mushroom) <- c("Y1_l", "Y1_u", "Y2_l", "Y2_u", "Y3_l", "Y3_u")

# calculate var-cov matrix for mushroom
cw <- mushroom[,1:2] # cap width
sl <- mushroom[,3:4] # stipe length
st <- mushroom[,5:6] # stipe thickness

# convert factor to numeric
sl[,1] <- as.numeric(sl[,1])
st[,1] <- as.numeric(st[,1])
st[,2] <- as.numeric(st[,2])

sym_cov(cw, sl, st)

plot(c(min(st$Y3_l)-30,max(st$Y3_u)+30), c(min(cw$Y1_l)-30,
               max(cw$Y1_u)+30), type = "n", xlab = "", ylab = "")
rect(st$Y3_l, cw$Y1_l, st$Y3_u, cw$Y1_u, border = "blue")

plot(c(min(sl$Y2_l)-40,max(sl$Y2_u)+40), c(min(cw$Y1_l)-40,
               max(cw$Y1_u)+40), type = "n", xlab = "", ylab = "")
rect(sl$Y2_l, cw$Y1_l, sl$Y2_u, cw$Y1_u, border = "blue")
```

```
# fit model

m3 <- est_ord1(st$Y3_l, st$Y3_u, cw$Y1_l, cw$Y1_u)

# prediction: Y_3 = [10, 19]
-1.692 + 0.575*10
-1.692 + 0.575*19

# 95% CI for Y_L
var_beta0 <- 0.0659
var_beta1 <- 0.0003

lower_YL <- 4.058 - qt(.975, 264-2)*sqrt(var_beta0 + var_beta1 * 10^2)
upper_YL <- 4.058 + qt(.975, 264-2)*sqrt(var_beta0 + var_beta1 * 10^2)

# 95% CI for Y_U
lower_YU <- 9.233 - qt(.975, 264-2)*sqrt(var_beta0 + var_beta1 * 19^2)
upper_YU <- 9.233 + qt(.975, 264-2)*sqrt(var_beta0 + var_beta1 * 19^2)
```

Chapter 6

FUTURE WORK

In Chapter 3, we proposed a novel approach to conduct statistical inference with point estimation and confidence interval on interval-valued data regressions. This approach is for now applied to deal with the simple regression model in (3.5). We also used the assumption in (3.39) to restrict that the error for the lower bound of response and the error for the upper bound of response within the same observation are independent, which may not be always advisable for real cases. In Section 6.1, we discuss about how to generalize the proposed method to handle multiple regression models; in Section 6.2, we consider relaxing the independence assumption in (3.39) to take account of possible relations between the lower and upper bounds for a response within each observation when applying the proposed method.

6.1 Generalization to Multiple Regression

Determination of Likelihood Function

Recalling (3.1) to (3.3), we can express the multiple regression model as

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon \tag{6.1}$$

where $Y = [Y_L, Y_U]$, $X_j = [X_{jL}, X_{jU}]$, for $j = 1, ..., p$, are all interval-valued variables. The Y is the response variable and X_j, $j = 1, ..., p$, are the predictors. In order to obtain the maximum likelihood estimators (MLEs) of the regression coefficients, β_j, $j = 0, ..., p$, we first need to generate the likelihood function for the error term. With the assumption that the errors are independent across observations, and the lower and upper bounds of the error are independent for each of the observations in (3.39), by (3.47), the likelihood function of the random error is

$$\begin{aligned} L(\epsilon_{L1}, \epsilon_{U1}, ..., \epsilon_{Ln}, \epsilon_{Un}) &= \prod_{i=1}^{n} g(\epsilon_{Li}, \epsilon_{Ui}) \\ &= (2\pi\sigma^2)^{-n} exp(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\epsilon_{Li}^2 + \epsilon_{Ui}^2)) \end{aligned} \tag{6.2}$$

where $g(\epsilon_{Li}, \epsilon_{Ui})$ is obtained in (3.46), for $i = 1, ..., n$, and n is the number of observations.

By (3.40), (3.41) and Figure 3.13, we know that the forms of ϵ_{Li} and ϵ_{Ui} are determined by whether the effect of each predictor is positive or negative. Suppose among the p predictors, k of them have positive effects to the response, denoted by $X_1, ..., X_k$; the rest of the $p - k$ predictors have negative effects to the response, denoted by $X_{k+1}, ..., X_p$. Then the random

error $\epsilon_i = [\epsilon_{Li}, \epsilon_{Ui}]$ can be expressed as

$$\epsilon_{Li} = Y_{Li} - \beta_0 - \beta_1 X_{1Li} - \beta_2 X_{2Li} - \cdots - \beta_k X_{kLi} - \beta_{k+1} X_{k+1,Ui} - \cdots - \beta_p X_{pUi}, \tag{6.3}$$

$$\epsilon_{Ui} = Y_{Ui} - \beta_0 - \beta_1 X_{1Ui} - \beta_2 X_{2Ui} - \cdots - \beta_k X_{kUi} - \beta_{k+1} X_{k+1,Li} - \cdots - \beta_p X_{pLi}, \tag{6.4}$$

for $i = 1, ..., n$. Replacing ϵ_{Li} and ϵ_{Ui} in (6.2) by (6.3) and (6.4), we can express the likelihood function of the random error as function of the $p + 1$ predictors.

Before computing the MLEs of β_j, for $j = 0, ..., p$, it is of importance to detect which of the predictors have positive effects or negative effects. Expanding upon the idea in Section 3.5, the value of the correlation between the jth predictor X_j and the response variable Y, for $j = 1, ..., p$, can be used as an indicator of positive effect or negative effect that the predictor has on the response. Therefore, we first calculate the correlation, $r_j = Corr(Y, X_j)$, $j = 1, ..., p$, by (2.8) and (2.9), and judge $\beta_j \geq 0$ if $r_j \geq 0$, and $\beta_j < 0$ if $r_j < 0$, for $j = 1, ..., p$. Then, the likelihood function with respect to the p predictors can be generated by (6.2), (6.3), and (6.4).

Point Estimation and Confidence Interval

Similar to what we did in Section 3.2.2 and Section 3.2.3, to give the point estimators and confidence intervals for the regression coefficients, β_j, $j = 1, ..., p$, firstly, we take the first derivative of the log likelihood function with respect to $\underline{\boldsymbol{\beta}} = (\beta_0, \beta_1, ..., \beta_p)^T$, and set it at $\underline{\boldsymbol{\beta}} = \underline{\hat{\boldsymbol{\beta}}}$ to be $\underline{\mathbf{0}} = (0, 0, ..., 0)_{p\times 1}$. Then, we solve the equations and obtain the MLEs of each of the $p + 1$ predictors, β_j, $j = 0, 1, ..., p$. Secondly, the expectation, the variance, as well as the distribution of the MLE $\hat{\beta}_j$, $j = 0, 1, ..., p$, can be obtained based on the assumption in (3.39). As the third step, we compute confidence intervals for the $p + 1$ predictors, β_j, $j = 0, 1, ..., p$, by the distributions obtained in the second step.

6.2 Measurement of Correlations Between the Lower and the Upper Bounds of Error Term

By the assumption given in (3.39), we restrict that the error for the lower bound of response and the error for the upper bound of response within the same observation are independent to each other. In the following two subsections, we try to relax this assumption to make it more flexible in order to fit more general cases encountered in the real world.

Correlation Measured by Additive Factor

The first relaxation for the assumption in (3.39) is to consider the difference between ϵ_{Li} and ϵ_{Ui} as being described as a constant, for $i = 1, ..., n$, where n is the number of observations. We call the constant as the "Additive Factor" in the error:

$$\epsilon_{Ui} = \epsilon_{Li} + c_0, \ \epsilon_{Li} \overset{iid}{\sim} N(0, \sigma^2), \ i = 1, ..., n. \tag{6.5}$$

In (6.5), c_0 is a constant, which can be pre-determined by two different ways. The first is to estimate it. Specifically, with the framework of simple regression models in (3.5), as the initial step, we assume c_0 is zero and compute the MLEs of the slope and the intercept parameters, β_1 and β_0, as proposed in Section 3.2.2. Then, we compute the predicted values of the response and the residuals by (3.42) and (3.43). The c_0 can be estimated by taking the average value of $r_{Ui} - r_{Li}$, where r_{Ui} and r_{Li} refer to the lower bound and the upper bound of the residual for the ith observation, for $i = 1, ..., n$. The second way to give the value of c_0 is by connecting it with the background of the data set. Sometimes the difference between the upper bound and the lower bound of error within the same observation is stable around a fixed value. With the relaxed assumption in (6.5), this kind of data sets can be fit better by a linear regression model.

With c_0 pre-determined, by (6.5), we have $\epsilon_{Ui} \overset{iid}{\sim} N(c_0, \sigma^2), i = 1, ..., n$. Then, similar to the procedures in Section 3.2.2 and 3.2.3, we can express the likelihood function of the error term with respect to β_1 and β_0, and then give the point estimators as well as confidence intervals for the slope and the intercept parameters under the assumption in (6.5).

Correlation Measured by Multiplication Factor

The second relaxation for the assumption in (3.39) is to consider the ratio of ϵ_{Li} and ϵ_{Ui} as a constant, for $i = 1, ..., n$, where n is the number of observations. We call the constant, ω, as the "Multiplication Factor" in the error:

$$\frac{\epsilon_{Ui}}{\epsilon_{Li}} = \omega, \ \epsilon_{Li} \overset{iid}{\sim} N(0, \sigma^2), \ i = 1, ..., n. \tag{6.6}$$

As in Section 6.2, the constant ω can also be given in two different ways. The first is by taking the average of the ratios $\frac{r_{Ui}}{r_{Li}}$, where r_{Ui} and r_{Li} refer to the lower bound and the upper bound of the residual for the ith observation, $i = 1, ..., n$, after assuming $\omega = 1$ as the initial step to calculate $\hat{\beta}_1$ and $\hat{\beta}_0$ by the proposed approach in Section 3.2.2 and Section 3.2.3, and then to obtain the corresponding predicted responses and residuals. The second way to set the value of ω is by studying the features of the data set, and detect the multiple factor accordingly.

With ω pre-specified, by (6.6), we have $\epsilon_{Ui} \overset{iid}{\sim} N(0, \omega^2\sigma^2), i = 1, ..., n$. Then, similar to the procedures in Section 3.2.2 and 3.2.3, we can express the likelihood function of the error term with respect to β_1 and β_0, and then give the point estimators as well as confidence intervals for the slope and the intercept parameters under the assumption in (6.6).

REFERENCES

[1] Billard, L. and Diday, E. (2006). *Symbolic Data Analysis: Conceptual Statistics and Data Mining*, John Wiley, Chichester.

[2] Billard, L. and Diday, E. (2000). Regression analysis for interval-valued data, *Data Analysis, Classification, and Related Methods*, Springer, 369-374.

[3] de Carvalho, F. A. T., Lima Neto, E.A. and Tenorio, C.P. (2004). A new method to fit a linear regression model for interval-valued data, *Annual Conference on Artificial Intelligence*, Springer, 295-306.

[4] Neto, L., de Carvalho, F. A. T., and Tenorio, C. P. (2004). Univariate and multivariate linear regression methods to predict interval-valued features, *Australasian Joint Conference on Artificial Intelligence*, Springer, 526-537.

[5] Neto, L. and de Carvalho, F. A. T. (2008). Centre and Range method for fitting a linear regression model to symbolic interval data, *Computational Statistics & Data Analysis*, 52(3):1500-1515.

[6] Neto, L., de Carvalho, F. A. T. and Freire, E. S. (2005). Applying constrained linear regression models to predict interval-valued data, *Annual Conference on Artificial Intelligence*, Springer, 92-106.

[7] Neto, L., de Carvalho, F. A. T. (2010). Constrained linear regression models for symbolic interval-valued variables, *Computational Statistics & Data Analysis*, 54(2):333-347.

[8] Xu, W. (2010). *Symbolic Data Analysis: Interval-valued Data Regression*, Doctoral Dissertation, University of Georgia.

[9] Bertrand, P. and Goupil, F. (2000). Descriptive statistics for symbolic data. In Bock, H. and Diday, E., editors, *Analysis of Symbolic Data: Exploratory Methods for Extracting Statistical Information from Complex Data*, Springer Berlin Heidelberg, 106-124.

[10] Billard, L. (2007). Dependencies and Variation Components of Symbolic Interval-Valued Data. In Brito, P., Cucumel, G., Bertrand, P. and de Carvalho, F. A. T., editors, *Selected Contributions in Data Analysis and Classification*, Springer Berlin Heidelberg, 3-12.

[11] Billard, L. (2008). Sample covariance functions for complex quantitative data. *World Congress International Association Statistical Computing*, Yokohama, Japan.

[12] Billard, L. (2011), Brief overview of symbolic data and analytic issues. *Statistical Analysis and Data Mining*, 4(2):149-156.

[13] Billard, L. and Diday, E. (2003). From the statistics of data to the statistics of knowledge: symbolic data analysis. *Journal of the American Statistical Association*, 98(462):470-487.

[14] Cazes, P., Chouakria, A., Diday, E. and Schektman, Y. (1997). Extension de l'analyse en composantes principales des donnes de type intervalle. *Revue de Statistique applique*, 45(3):5-24.

[15] Lauro, C. N. and Palumbo, F. (2000). Principal component analysis of interval data: a symbolic data analysis approach. *Computational statistics*, 15(1):73-87.

[16] Billard, L., Douzal-Chouakria, A. and Diday, E. (2008). Symbolic principal component for interval-valued observations. <hal-00361053>.

[17] Billard, L. and LeRademacher, J. (2012) Principal component analysis for interval data. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(6):535-540.

[18] Le-Rademacher, J. and Billard, L. (2013). Principal component histograms from interval-valued observations. *Computational Statistics*, 28(5):2117-2138.

[19] Le-Rademacher, J. and Billard, L. (2013) Principal component analysis for histogram-valued data. *Advances in Data Analysis and Classification*, 1-25.

[20] Palumbo, F. and Verde, R. (2000). Nonsymmetrical factorial discriminant analysis for symbolic objects. *Applied Stochastic Models in Business and Industry*, 15(4):419-427.

[21] Lauro, N. C., Verde, R. and Palumbo, F. (2000). Factorial discriminant analysis on symbolic objects. In: *Analysis of Symbolic Data Exploratory Methods for Extracting Statistical Information from Complex Data*, Springer, Berlin Heidelberg New York, 212-233.

[22] Hiremath, P. S. and Prabhakar, C. J. (2008). Symbolic factorial discriminant analysis for illumination invariant face recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, 22(03):371-387.

[23] Silva, A. P. D. and Brito, P. (2006). Linear discriminant analysis for interval data. *Computational Statistics*, 21(2):289-308.

[24] Silva, A. P. D. and Brito, P. (2015). Discriminant analysis of interval data: an assessment of parametric and distance-based approaches. *Journal of Classification* 32(3):516-541.

[25] Denoeux, T. and Masson, M. (2000). Multidimensional scaling of interval-valued dissimilarity data. *Pattern Recognition Letters* 21(1):83-92.

[26] Masson, M. and Denoeux, T. (2002). Multidimensional scaling of fuzzy dissimilarity data. *Fuzzy sets and systems* 128(3):339-352.

[27] Groenen, P. J., Winsberg, S., Rodriguez, O. and Diday, E. (2006). I-Scal: Multidimensional scaling of interval dissimilarities. *Computational Statistics & Data Analysis*, 51(1):360-378.

[28] Huang, J. J., Ong, C. S. and Tzeng, G. H. (2006). Interval multidimensional scaling for group decision using rough set concept. *Expert Systems with Applications*, 31(3):525-530.

[29] Terada, Y. and Yadohisa, H. (2011). Multidimensional scaling with the nested hypersphere model for percentile dissimilarities. *Procedia Computer Science*, 6:364-369.

[30] Ichino, M., Yaguchi, H. and Diday, E. (1996). Symbolic pattern classifiers based on the cartesian system model. *Ordinal and symbolic data analysis*, 92-102.

[31] Rasson, J. P. and Lissoir, S. (2000). Symbolic kernel discriminant analysis. *Computational Statistics*, 15(1):127-132.

[32] Prinel, E. and Lechevallier, Y. (2000). Symbolic discrimination rules. In: *Analysis of Symbolic Data: Exploratory Methods for Extracting Statistical Information From Complex Data* (eds. H.H. Bock and E. Diday), 244-265.

[33] Dinesh, M. S., Gowda, K. C. and Nagabhushan, P. (2005). Fuzzy-symbolic analysis for classification of symbolic data. In: *International Conference on Pattern Recognition and Machine Intelligence*, Springer Berlin Heidelberg, 338-343.

[34] Maia, A. L. S., de Carvalho, F. D. A. and Ludermir, T. B. (2008). Forecasting models for interval-valued time series. *Neurocomputing*, 71(16):3344-3352.

[35] Gowda, K. C. and Diday, E. (1991). Symbolic clustering using a new dissimilarity measure. *Pattern Recognition*, 24(6):567-578.

[36] Gowda, K. C. and Diday, E. (1992). Symbolic clustering using a new similarity measure. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(2):368-378.

[37] Ichino, M. and Yaguchi, H. (1994). Generalized Minkowski metrics for mixed feature-type data analysis. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(4):698-708.

[38] Chavent, M. (1998). A monothetic clustering method. *Pattern Recognition Letters* 19(11):989-996.

[39] Guru, D. S., Kiranagi, B. B. and Nagabhushan, P. (2004). Multivalued type proximity measure and concept of mutual similarity value useful for clustering symbolic patterns. *Pattern Recognition Letters*, 25(10):1203-1213.

[40] Guru, D. S. and Kiranagi, B. B. (2005). Multivalued type dissimilarity measure and concept of mutual dissimilarity value for clustering symbolic patterns. *Pattern Recognition*, 38(1):151-156.

[41] Kiranagi, B. B. and Guru, D. S. (2010). A New Symbolic Dissimilarity Measure for Multivalued Data Type and Novel Dissimilarity Approximation Techniques. *International Journal of Computer Applications*, 1(25):36-41.

[42] Bock, H. H. (2003). Clustering algorithms and kohonen maps for symbolic data (symbolic data analysis). *Journal of the Japanese Society of Computational Statistics*, 15(2):217-229.

[43] Chavent, M. and Lechevallier, Y. (2002). Dynamical clustering of interval data: optimization of an adequacy criterion based on Hausdorff distance. In: *Classification, Clustering, and Data Analysis*, Springer Berlin Heidelberg, 53-60.

[44] de Souza, R. M. and De Carvalho, F. D. A. (2004). Clustering of interval data based on cityblock distances. *Pattern Recognition Letters*, 25(3):353-365.

[45] De Carvalho, F. D. A., de Souza, R. M., Chavent, M. and Lechevallier, Y. (2006). Adaptive Hausdorff distances and dynamic clustering of symbolic interval data. *Pattern Recognition Letters*, 27(3):167-179.

[46] de Carvalho, F. D. A. (2007). Fuzzy c-means clustering methods for symbolic interval data. *Pattern Recognition Letters*, 28(4):423-437.

[47] de Souza, R. M. C. R., de Carvalho, F. D. A. T. and Pizzato, D. F. (2006). A partitioning method for mixed feature-type symbolic data using a squared Euclidean distance. In: *Annual Conference on Artificial Intelligence*, Springer Berlin Heidelberg, 260-273.

[48] Pimentel, B., Costa, A. and Souza, R. (2011). A partitioning method for symbolic interval data based on kernelized metric. In: *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*, ACM, 2189-2192.

[49] Lawson, C. L, and Hanson, R. J. (1995). *Solving Least Squares Problems*, Philadelphia, Pa.: Society for Industrial and Applied Mathematics.

[50] Casella, G. and Berger, R. (2002). *Statistical Inference, 2nd Edition*, Thomson Learning.