LINEAR REGRESSION PARAMETER ESTIMATION METHODS FOR THE WEIBULL DISTRIBUTION

ZHANG LIFANG

(B.Eng, BUAA (Beijing, China))

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF INDUSTRIAL & SYSTEEMS ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2008

Acknowledgements

First of all, I would like to express my sincerest gratitude to my supervisors, Professor Xie Min and Professor Tang Loon Ching, for their guidance, encouragement and useful comments throughout my PhD program. Their observations and comments helped me to establish the directions of my research so that I can easily move forward with investigations in depth.

I would like to thank my lab mates and friends, Hu Qingpei, Xu Zhiyong, Tang Yong, Peng Ji and others for their help in my research and life.

Thanks National University of Singapore, especially the Department of Industrial & Systems Engineering for offering research scholarship and the use of its resources and facilities.

Last but not least, I would like to thank my husband and my parents for their love and continuous support. I couldn't imagine if I can finish my study without them.

Table of Contents

Summary

Weibull distribution is one of the most widely used distributions in reliability data analysis. Many methods have been proposed for estimating the two Weibull parameters, among which Weibull probability plot (WPP), maximum likelihood estimation (MLE) and least squares estimation (LSE) are the methods frequently used nowadays.

LSE is the basic linear regression estimation method. It is frequently used with WPP to show a graphical presentation. Such a method is preferred by practitioners; however, it can perform very poorly for some data types. This thesis explores various refinements of the ordinary LSE (OLSE) method. First, it presents a thorough examination of the properties of the OLS estimators via both theoretical analyses and intensive Monte Carlo simulation experiments. Second, it provides suggestions on the procedure of the OLSE method including the selection of failure probability estimators and the regression direction. Third, it proposes simple bias correcting formulas for the OLSE of the shape parameter applied to both complete data and censored data. Fourth, sophisticated linear regression techniques including weighted least squares and robust regression are examined to replace the OLS technique for estimating the Weibull parameters. Finally, it provides application instructions for the linear regression estimation methods discussed in this study with numerical examples.

This thesis focuses on small samples, multiply censored samples, and samples with outliers. The proposed linear regression estimation methods are good for dealing with one or several of these data types. In addition, these methods are based on linear regression techniques and hence can be easily applied and understood.

List of Tables

List of Figures

Notations and Abbreviations

n	Sample size
i	Order number of observations from smallest to largest, $1 \le i \le n$
r	Number of failures in a sample
c	Censoring level, $c = (n-r)/n$
t	Time (failure time or censoring time)
$t_{(i)}$	The i^{th} smallest time of $t_1, t_2, \cdots, t_n$
$t_{f,(j)}$	The j^{th} smallest failure time in a censored sample, $1 \le j \le r$
$t_{c,(k)}$	The k^{th} smallest censoring time in a censored sample, $1 \le k \le (n-r)$
$F(t)$	CDF of the Weibull distribution
$R(t)$	Reliability, $R(t) = 1 - F(t)$
α	Scale parameter of the Weibull distribution
β	Shape parameter of the Weibull distribution
α_T, β_T	True parameter values for α and β (simulation experiment factors)
$\hat{\alpha}, \hat{\beta}$	Estimators of α and β
$\hat{\alpha}_{\alpha_T,\beta_T}$, $\hat{\beta}_{\alpha_T,\beta_T}$	Estimators of α and β with given values of α_T and β_T
I_j	The event number of the j^{th} failure in a censored sample
$m_{f,(j)}$	The modified failure order number of $t_{f,(j)}$, $j \le m_{f,(j)} \le n$
$\hat{F}_{(i)}$, $\hat{F}_{f,(j)}$	Estimators of $F(t)$, $\hat{F}_{(i)}$ for complete data and $\hat{F}_{f,(j)}$ for censored data
e	Error or residual
$E(\cdot)$	Mean or expected value
$S(\cdot)$	Standard deviation
$Var(\cdot)$	Variance

$B(\cdot)$	Bias function
$\hat{\beta}_U$	Unbiased $\hat{\beta}$
U	Bias correcting factor
U_{MR}	Bias correcting factor of the modified Ross' method
U_{MH}	Bias correcting factor of the modified Hirose's method
M	Iteration number of simulation experiment
R^2	Coefficient of determination
MS_{error}	Model statistic, mean square error
w_i	Weights for failure observations
w_{nor_i}	Normalized weights for failure observations
w_{app_i}	Approximated weights

ASM	Age Sensitive Method (estimation for $\hat{F}_{f,(j)}$)
BLIE	Best Linear Invariant Estimator
BLUE	Best Linear Unbiased Estimator
BUE	Best Unbiased Estimator
CDF	Cumulative Distribution Function
EASM	Exponential Age Sensitive Method (estimation for $\hat{F}_{f,(j)}$)
HJ	Herd-Johnson
JM	Modified Johnson
KM	Kaplan-Meier
LSE	Least Squares Estimation/Estimator

MFON	Modified Failure Order Number
MLE	Maximum Likelihood Estimation/Estimator
MME	Method of Moment
MMME	Modified Method of Moment
MSE	Mean Square Error
MTTF	Mean Time To Failure
NBLIE	Nearly Best Linear Invariant Estimator
NBLUE	Nearly Best Linear Unbiased Estimator
OLSE	Ordinary Least Squares Estimation/Estimator
PDF	Probability Density Function
WLSE	Weighted Least Squares Estimation/Estimator
WPP	Weibull Probability Plot
RR	Robust Regression
RRE	Robust Regression Estimation/Estimator
RRRM	Refined Rank Regression Method (estimation for $\hat{F}_{f,(j)}$)

Chapter 1

Introduction

The history of the Weibull distribution can be traced back to 1928, when two researchers, Fisher and Tippett, deduced the distribution in their study of the extreme value theory (Arora, 2000). In the late 1930s, a Swedish professor Waloddi Weibull derived the same distribution and his hallmark paper in 1951 made this distribution fashionable. In his hallmark paper (Weibull, 1951), Professor Weibull explained the reasoning of the Weibull distribution through the phenomena of the weakest link in the chain and he said

> The same method of reasoning may be applied to the large group of problems, where the occurrence of an event in any part of an object may be said to have occurred in the object as a whole, e.g., the phenomena of yield limits, statical or dynamical strengths, electrical insulation breakdowns, life of electric bulbs, or even death of man…

All these words have become accepted as truth. Today, the Weibull distribution has wide applications in various areas. These applications include using the distribution to model wind speed, rainfall, flood or earthquake frequency, age of disease onset, strength of materials, and so on. However, the most extensive use of the distribution is in life testing and reliability studies, where the Weibull distribution has been proven to be satisfactory in modeling the phenomena of fatigue and life of many devices such as ball bearings, electric bulbs, capacitors, transistors, motors and automotive radiators. Due to its wide application in reliability studies, reliability data analysis is frequently called Weibull analysis (Wang, 2004).

The general form of the Weibull distribution has three parameters: the scale parameter, the shape parameter and the location parameter. In reliability data analysis, the location parameter is frequently neglected. As pointed out in Dodson (2006), a non-zero location parameter should not be used unless there is a physical justification for a time period with a zero probability of failure. This thesis focuses on the parameter estimation methods for the two-parameter Weibull distribution. Unless otherwise indicated, the Weibull distribution in this thesis refers to the two-parameter Weibull distribution.

Reliability data can be obtained from life testing experiments or from the field. Unlike other data analyses, reliability data analysis is complicated because different types of data may need different approaches for processing (Liu, 1997). When it comes to the estimation of the Weibull parameters (assuming the data is Weibull distributed), no method can always outperform the others for all types of data in view of the properties of the estimators. Moreover, the commonly used estimation methods such as the maximum likelihood estimation (MLE) method and the least squares estimation (LSE) method have been discovered to be unsatisfactory under many circumstances. The main focus of this thesis is to investigate various linear regression estimation techniques including LSE for the estimation of Weibull parameters that aim at different types of life data including small data sets, censored data sets and data sets with outliers.

This chapter starts with an overview of the Weibull distribution and the physical meanings of its two parameters in the context of reliability in Section 1.1. The scope of the Weibull analysis is also briefly presented. Section 1.2 describes the common types of life data under different classification schemes. Then Section 1.3 presents an overview of the existing Weibull parameter estimation methods and their limitations with the focus on the commonly used ones. Finally, Section 1.4 and Section 1.5 present the scope and the contributions of this thesis, respectively.

1.1 The Weibull Distribution in Reliability Engineering

The cumulative distribution function (CDF) and the probability density function (PDF) of the Weibull distribution are expressed by

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right] \tag{1-1}$$

$$f(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}\exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right] \tag{1-2}$$

where the scale parameter α and the shape parameter β take on positive values.

In the context of reliability, $F(t)$ is the probability that a random unit drawn from the population fails by time t $(t > 0)$, or the fraction of all units in the population that fails by t (Tobias & Trindade, 1995). The complement of $F(t)$ is the reliability function, i.e., $R(t) = 1 - F(t)$. From Equation (1-1), the expression for the Weibull reliability function is

$$R(t) = \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right] \tag{1-3}$$

Other common reliability measures include mean time to failure (MTTF), percentile life t_p and failure rate (or hazard rate) $\lambda(t)$. Based on the Weibull CDF, the expressions for these measures are

$$MTTF = \alpha\Gamma\left(1+\frac{1}{\beta}\right) \tag{1-4}$$

$$t_p = \alpha\left[-\ln(1-p)\right]^{\frac{1}{\beta}} \tag{1-5}$$

$$\lambda(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} \tag{1-6}$$

where $\Gamma(\cdot)$ denotes the Gamma function.

All of the above measures are functions of the two Weibull parameters. In the following, the effects of the scale parameter and the shape parameter on the Weibull distribution are separately described.

1.1.1 The Scale Parameter

Figure 1-1 shows the PDF plot of the Weibull distribution with different values of α and a common value of β. As it can be observed, an increase or a decrease in α while β is kept unchanged has an effect of stretching out the distribution to the right or pushing in the distribution to the left and it has no effect on the shape of the distribution. In fact, a change in the scale parameter α is the same as a change of the abscissa scale. The parameter α has the same unit as t, such as hours, miles, cycles, etc.

From Equation (1-5), when $p = 0.632$, we obtain

$$t_{0.632} = \alpha \tag{1-7}$$

Hence α is the time at which 63.2% of the population failed. It is frequently called the characteristic life.

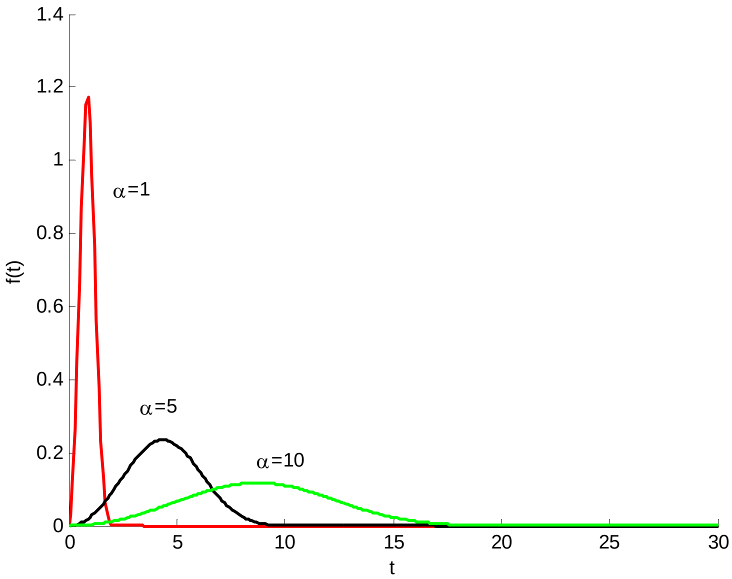


Figure 1-1: The effect of α on the Weibull PDF for a common β ($\beta = 3$).

1.1.2 The Shape Parameter

The shape parameter β is of great importance to the Weibull distribution because it determines the shape of the Weibull PDF and characterizes the failure rate trend. Figure 1-2 shows several typical examples of the Weibull PDF with different values of β and a common α. Figure 1-3 illustrates a variety of the failure rate curves with different values of β and a common α.

It can be observed from Figure 1-2 that when $0 < \beta < 1$, the PDF is exponentially decreasing. At $\beta = 1$, the Weibull distribution reduces to the exponential distribution. When $\beta > 1$, the PDF is unimodal and skewed to the right. When $3 \le \beta \le 4$, the PDF has a roughly bell-shape which is close to the normal distribution. Figure 1-3 shows the relationship between β and failure rate. As it can be observed, when $0 < \beta < 1$, the failure rate is exponentially decreasing (same as the PDF). At $\beta = 1$, the failure rate is constant and the failure rate $\lambda(t) = 1/\alpha$. When

$\beta > 1$, the failure rate is monotonically increasing. A special case is when $\beta = 2$ where the failure rate is linearly increasing. The distribution is called Rayleigh distribution. In other cases, the failure rate increases with different rates. Table 1-1 summarizes the typical characteristics of the Weibull PDF and failure rate with varying β.

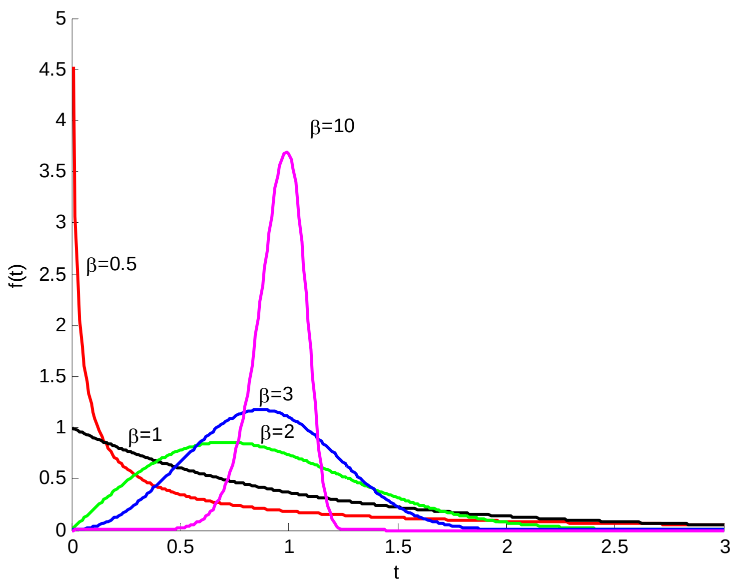


Figure 1-2: The effect of *β* on the Weibull PDF for a common *α* (*α* = 1).

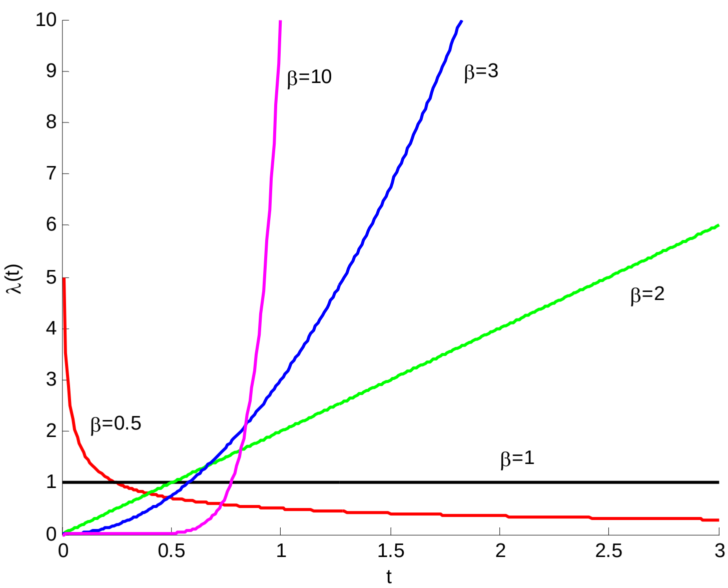


Figure 1-3: The effect of *β* on the failure rate for a common *α* (*α* = 1).

Table 1-1: Typical characteristics of the Weibull PDF and failure rate with varying shape parameter values.

Shape Parameter	PDF	Failure Rate
$0 < \beta < 1$	Exponentially decreasing from infinity	Exponentially decreasing
$\beta = 1$	Exponentially decreasing from $1/\alpha$	Constant
$\beta > 1$	Rises to peak and then decreases	Increasing
$\beta = 2$	A special case - Rayleigh distribution	Linearly increasing
$3 \leq \beta \leq 4$	"Normal" bell-shape appearance	Rapid increasing
$\beta > 10$	Similar to Type I extreme value distribution	Very rapidly increasing

The importance of the shape parameter to the Weibull distribution has been discussed by many researchers. Wu & Vollertsen (2002a, b) presented detailed analyses of the Weibull shape parameter in the context of the intrinsic breakdown of dielectric films. The shape parameter not only decides the characteristics of the Weibull PDF and failure rate, it also links the Weibull distribution to many other distributions. For example, the Weibull-to-exponential transformation is a commonly used method when the shape parameter can be obtained from material property or other sources (Xie et al., 2000). With this transformation, the simple statistical tests and analytical methods available for the exponential distribution can be applied to ease the data analysis for the Weibull distribution. Keats et al. (2000) presented the effect of the mis-specification of the shape parameter value on the estimation of the scale parameter, and Xie et al. (2000) extended the analysis to the effect of the mis-specification of the shape parameter on the estimation of reliability measures such as MTTF, percentiles and mission reliability. The authors found that it is true that the mis-specification will greatly affect the scale parameter because the two parameters are highly correlated; however, the effect on the MTTF, percentiles and mission reliability could be small.

1.1.3 The Bathtub Curve

The life cycles of mechanical and electronic units and systems are often described by the bathtub curve, see Figure 1-4. Based on the behavior of the failure rate, the life of a unit or system is divided into three periods: infant (or early failure) period, life (or intrinsic failure) period and wear-out (or aging) period. These periods are characterized by a decreasing, constant and increasing failure rate, respectively. Assuming the life distribution is Weibull, the value of the shape parameter can indicate which period the unit or system lies in. When $0 < \beta < 1$, it is in the infant period. When $\beta = 1$, it is in the life period, and when $\beta > 1$, it is in the wear-out period. The value of β also indicates the failure mechanism of a unit or system being early failures, random failures or wear-out failures. Table 1-2 summarizes the relationship of life periods, failure mechanisms and the values of β.

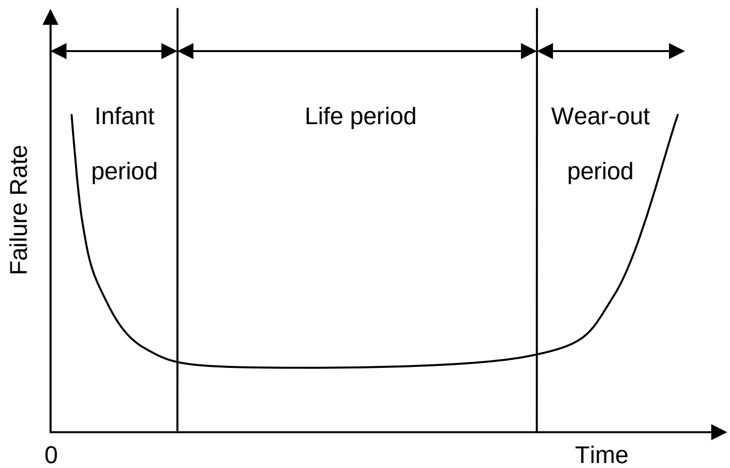


Figure 1-4: The bathtub curve.

Table 1-2: The relationship of life period, failure mechanism and *β*.

Shape Parameter	Life Period	Failure Mechanism
$0 < \beta < 1$	Infant period	Early failure
$\beta = 1$	Life period	Random failure
$\beta > 1$	Wear-out period	Wear-out failure

As can be seen from Figure 1-3, however, no matter what value of the shape parameter takes, the Weibull distribution has a monotonic failure rate. This monotonicity becomes a limitation as some products exhibit more than one stage of the bathtub curve. The turning point of the failure rate trend is considered a 'critical' time and is important (Bebbington et al., 2008). To overcome this, a group of new distributions have been proposed in the last decade, and these distributions are commonly named as modified/extended/generalized Weibull distributions. In recent years, great interests have been put to develop distributions with bathtub-shaped failure rate functions. A good example can be found in Xie et al. (2002). Murthy et al. (2004) summarized many of these new distributions and provided details for their backgrounds, statistical analysis methods, practical applications, etc. Bebbington et al. (2007) proposed a so-called flexible Weibull distribution which has only two parameters and is able to model a modified bathtub-shaped failure rate where the failure rate increases at the beginning and then follows a bathtub curve. Zhang & Xie (2007) proposed a three-parameter distribution called extended Weibull distribution. This distribution is very flexible in view of the failure rate function, which can be a modified bathtub-shaped curve with a first stage increasing, or initialing decreasing eventually decreasing but with increasing in the middle. Dimitrakopoulou et al. (2007) proposed another three-parameter distribution which can specially present an upside down bathtub-shaped failure rate. Pham & Lai (2007) summarized a few popular Weibull-related models and discussed the issues of parameter estimation and model validation.

1.1.4 Scope of the Weibull Analysis

Weibull analysis, or reliability data analysis, commonly involves the following activities (Abernethy, 2000):

- Plotting the data and interpreting the plot
- Failure forecasting and prediction
- Evaluating corrective action plans
- Maintenance planning
- Spare parts forecasting
- Warranty analysis
- Others

Parameter estimation of the two Weibull parameters often serves as the preliminary step of the Weibull analysis after samples are collected. Accurate parameter estimates may greatly affect the accuracy of the subsequent analyses.

1.2 Types of Life Data

The most common classification of life data is based on the life testing experiment scheme. If all the units are tested to failure, this sample is a *complete* or *uncensored sample*. Otherwise, if the experiment ends before all units fail, this sample is a *censored sample*. Censored units are called *censors* or *suspensions* and their failure times are only known to be beyond their present running times (i.e., the censoring times). If all units are started on the test together and all censors have a common running time, the data are *singly censored*. Such data are further classified into *time censored or Type I censored* if the test is stopped at a predetermined time, and *failure censored or Type II censored* if the test is stopped when a predetermined number of failures occur. If units begin their services at different times and thus when the test stops before all units are failed, the censoring times and the failure times are intermixed, the data are said to be *multiply censored*. Singly censored data can be treated as a special case from multiply censored data; however, they are often examined separately in the Weibull analysis. Besides, there are other types of censored data, e.g., *left censored data*, *doubly censored data*, *progressively Type II censored data*, etc., which are beyond the scope of this study. Figure 1-5 illustrates four common types of samples including a complete sample, a singly time censored sample (Type I censored), a singly failure censored sample (Type II censored) and a multiply censored sample.

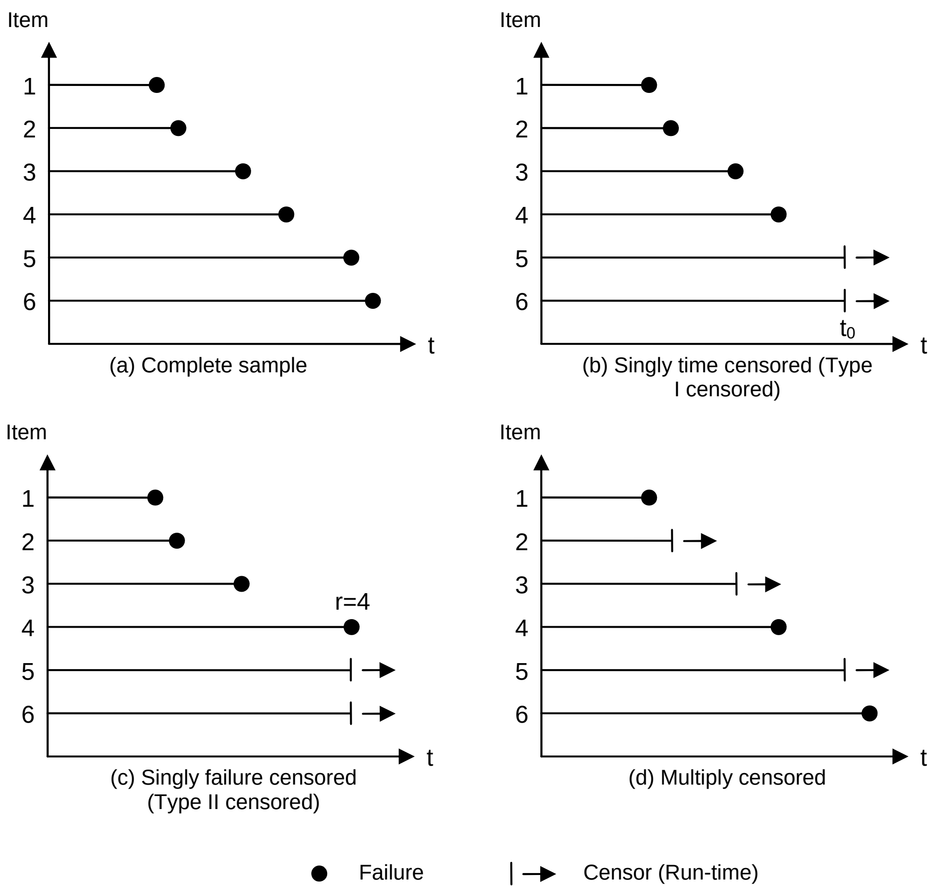


Figure 1-5: An illustration of different types of life data.

Besides the conventional classification which divides life data into complete data and censored data, life data can also be classified into different groups based on data source, sample size and the quality of the data. A summary of the classification is shown in Figure 1-6.

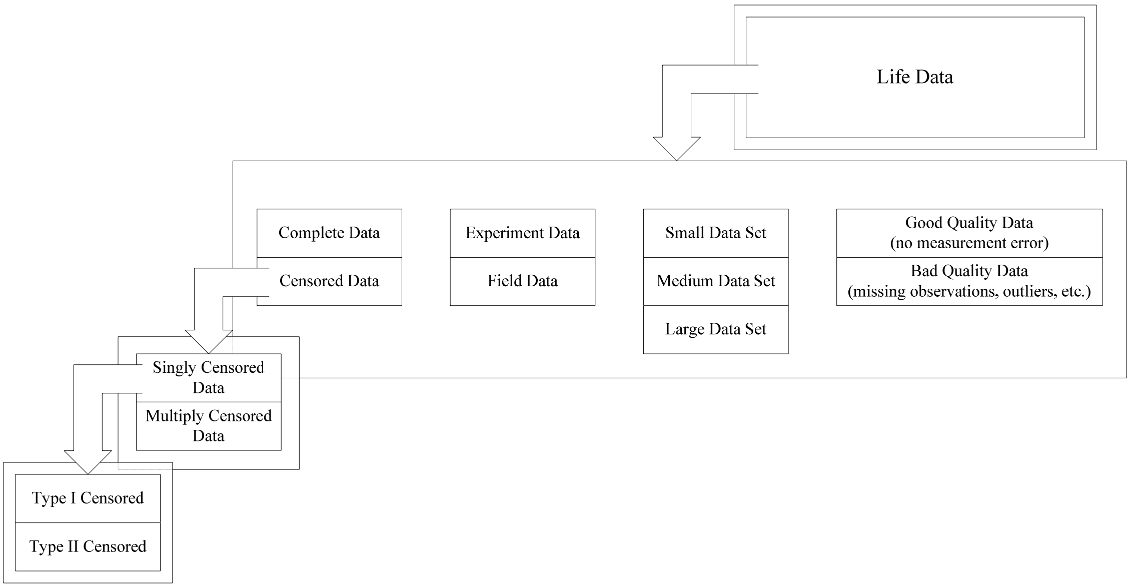


Figure 1-6: The classifications of life data based on testing schemes, data source, sample size and quality of observations.

In view of data source, life data are divided into experiment data and field data. Based on the number of observations or the sample size, a data set can be classified into a small, medium or large data set. Normally a data set with no more than 20 observations is considered as a small dataset (Abernethy, 2000). Besides, life data can be divided into good quality data and bad quality data. Good quality data ideally have no measurement errors in the observations (i.e., failure time), or the error is small enough to be neglected; while bad quality data involve outliers, influential points or missing observations, etc.

Figure 1-6 does not provide an exhaustive classification for life data. For example, there are other common data types such as group data and interval data which are not included. Recently, some methods were proposed to estimate Weibull parameters for interval data, see, e.g., Vittal & Phillips (2007).

Life data have some special characteristics. For example, small data sets and censored data sets are very common due to time and cost constrains. The increase of the number of highly reliable systems also leads to the difficulty of collecting failure data. These data conditions require specially designed data analysis techniques.

Given the perspectives of real applications, small data sets, multiply censored data and bad quality data with outliers or influential points, are the focuses of this research.

1.3 Overview of Weibull Parameter Estimation Methods

Since Weibull distribution became widely recognized in the 1950s, many methods have been proposed for estimating the parameters. Both graphical estimation methods and analytical estimation methods have been proposed. This section provides an overview of the existing parameter estimation methods for the Weibull distribution. It is impossible to list all the related work in the literature, thus the focus is given to those commonly used methods.

1.3.1 Graphical Estimation Methods

There are mainly two categories of graphical estimation methods for the Weibull distribution: Weibull probability plotting (WPP) methods and hazard plotting methods. For a basic understanding of the two methods, see, e.g., Lai & Xie (2006, p. 145), Breyfogle (1992, p. 163) and Nelson (2004, chap. 3 & 4).

Probability plotting for the Weibull distribution was introduced by Kao (1959). Some discussions on the Weibull probability paper can be found in, e.g., Nelson & Thompson (1971). White (1969) suggested using some analytic techniques such as least squares to fit the straight line on the WPP instead of eye-fitting. Cran (1976) gave several numerical examples of using probability plotting to estimate the Weibull parameters. The WPP technique has been also used on the modified or extended Weibull distributions, see, e.g., Murthy et al. (2004).

The related work on WPP has been centered on the determination of the *Y*-axis plotting positions. Conventionally, the *Y*-axis plotting positions on the Weibull probability paper, which denote failure probabilities or unreliability, are estimated by some non-parametric estimators of the form $(i-c_1)/(n+c_2)$. Professor Weibull originally used $i/(n+1)$ to obtain the plotting positions (Weibull, 1939). This is then named Weibull plotting position or Weibull estimator. Theoretically, it is the exact mean rank plotting position of each data point. The Weibull estimator had been used for many years until the Bernard estimator became more popular. The Bernard estimator, i.e., $(i-0.3)/(n+0.4)$, was proposed by Bernard & Bosi-Levenbach (1953) as an approximation to the median rank plotting position. It is a good approximation to the exact median rank value of each data point shown by Mischke (1979) via analytical methods and Fothergill (1990) via Monte Carlo simulations. Compared to the mean rank plotting position, one of the good properties of the median rank plotting position is that it is distribution free (Mischke, 1979; Yu & Hung, 2001). With Monte Carlo simulations, many researchers, see, e.g., Fothergill (1990) and Cacciari & Montanari (1991), have compared several plotting positions including Weibull (Weibull, 1939), Bernard (Bernard & Bosi-Levenbach, 1953), Hazen (Hazen, 1930), Blom (Blom, 1958), Filliben (Filliben, 1975), etc., on estimating Weibull parameters for complete samples of different sample sizes. Most agreement has been achieved on the Bernard estimator and hence it is most widely used today. Many textbooks on reliability data analysis have adopted the Bernard estimator as the standard method for estimating failure probabilities, see, e.g., Tobias & Trindade (1995).

Besides the Weibull estimator and the Bernard estimator, a few other estimators for failure probability or *Y*-axis plotting positions were discussed in the last decade. Ross (1994b) suggested a *Y*-axis plotting position that he called the expected plotting position. Two formulas were provided. One is used to calculate the exact expected plotting position for each data point, which has a complex form, and the other is a simple approximation to the exact values and the formula is $(i-0.44)/(n+0.25)$. However, these formulas, especially the simplified one, have not received as much attention as they should have. Drapella & Kosznik (1999) suggested a similar approach as Ross' for calculating *Y*-axis plotting positions and their formula is basically same as that of Ross' for the exact expected plotting position. The formula has then been cited many times in recent years and is considered to be a bias correction method for the conventional LSE method, see, e.g., Xie et al. (2000), Yang & Xie (2003), Hung (2004) and Lu et al. (2004). The recent work of Wu & Lu (2004) and Wu et al. (2006) examined the idea of using different failure probability estimators for different sample sizes. The authors tabulated the optimal estimators for certain sample sizes. Tiryakioglu & Hudak (2007), in a similar way, tabulated another set of optimal estimators for different sample sizes between 9 and 50. However, since there is no certain pattern in these tabulations, this kind of method is apparently inconvenient in view of practical application.

The above non-parametric plotting positions are mainly designed for complete data, though it is not uncommon to see that they are wrongly used for censored data in the literature. For censored data, to best use the information from all the observations, new methods are needed to obtain plotting positions. The Kaplan-Meier estimator (Kaplan & Meier, 1958) is the oldest non-parametric estimator of failure probabilities applied to censored data. A big disadvantage of the estimator is that the unreliability for the last failure data point is always 1, and hence it tends to underestimate the failures in the tail of the distribution. Herd (1960) proposed a method to calculate the reliability at each failure data point recursively in the case of multiply censored data, and Johnson (1964) decomposed the Herd's method into two steps: first is to calculate the modified failure order number (MFON) of each failure data point and then use the MFON in the Weibull estimator to estimate the reliability or failure probability. The combination of their work is commonly known as the Herd-Johnson method. Nelson once commented the Johnson's method (Johnson, 1964) as a small and laborious refinement compared to the original estimator of Herd (Herd, 1960), see, e.g., Nelson (2004, pp. 147-148). However, the two-step estimation of the failure probability with the identification of the MFON as the first step gained its popularity in the last decade as the age sensitive methods were proposed, see, e.g., the age sensitive method of Hastings & Bartlett (1997) and the exponential age sensitive method of Campean (2000). More recently, Skinner et al. (2001) and Hossain & Zimmer (2003) modified the Herd-Johnson method and proposed a simple formula which can directly calculate the failure probability. Wang (2001, 2004) proposed a so-called refined rank regression method which is a parametric method and must be solved iteratively. Despite the calculation complexity, Wang's method has a good theoretical background and does not need many assumptions. Although these recently proposed methods have been shown by the authors to outperform the Kaplan-Meier estimator or the Herd-Johnson estimator, none of them have become popular or widely recognized. The practitioners have not been aware of them. Therefore, a systematic comparison of the existing methods in view of parameter estimation for the Weibull parameters will be useful.

Obviously, the research on the estimation of failure probabilities or the *Y*-axis plotting positions in the cases of both complete data and multiply censored data has not reached a final conclusion. In Section 4.3, a detailed summary on the existing plotting positions is presented for complete data and multiply censored data, respectively, and the recommendations are given both from the theoretical point of view and from Monte Carlo simulation results.

Another graphical estimation method is the hazard plotting estimation method proposed by Nelson, see, e.g., Nelson (1972, 2004), and it also received many agreements. Many years ago, the graphical methods were all done manually and the big advantage of using hazard plotting for censored data is to save human labor (Breyfogle, 1992). In view of estimation accuracy, however, hazard plotting will probably not outperform probability plotting because its estimation for the hazard function (i.e., $h(t) = 1/$the reserve rank of each failure data point) is very simple and there are few alternatives. In contrast, the probability plotting technique has the variety because of the various plotting positions that can be applied. Obviously, by changing the plotting positions, the probability plot can achieve a better fit of sample data then the hazard plot.

As mentioned, hazard plotting is a simple but less flexible method compared to WPP. Besides, the programs of WPP are available in many statistical software packages, e.g., MATLAB 7, and hence WPP is readily applicable.

1.3.2 Analytical Estimation Methods

Analytical estimation methods for the Weibull distribution have a large family. Typical methods include: method of moment estimation (MME) or modified method of moment estimation (MMME), maximum likelihood estimation (MLE), least squares estimation (LSE), method of percentiles and Bayesian estimation method.

Earlier studies have been mainly confined to MLE and MME/MMME. The references on MME and MMME can be found in Dubey (1966), Mann (1968), Newby (1980), Arora (2000), etc. It has been found that MLE outperforms MME/MMME in most cases, see, e.g., Mann (1968), and MME/MMME is usually not efficient compared to other methods such as MLE (Murthy et al., 2004, p. 62). In fact, the MME/MMME methods are seldom discussed by Weibull researchers nowadays.

MLE, in contrast, is preferred by a majority of Weibull researchers because of its good statistical perspectives. Cohen (1965) first presented the estimating equations of the MLE method of the two-parameter Weibull distribution for different types of samples including complete samples, Type I or Type II singly censored samples and progressively censored samples (i.e., removing one or more items from life testing at various times prior to the termination of the test). Harter & Moore (1965) presented the MLE method of the three-parameter Weibull distribution when all the three parameters are unknown for complete samples and Type II singly censored samples.

The existence and uniqueness of the maximum likelihood estimators (MLE) have been discussed by many researchers. McCool (1970) proved that the MLEs of the shape and scale parameters always exist and are unique when the location parameter is known (for example, the two-parameter Weibull distribution). Farnum & Booth (1997) presented similar results for the MLE applied to complete data and singly censored data, and introduced a statistic which can be used to get a quick approximation of the shape parameter estimate. However, the existence and uniqueness of the MLE does not necessarily apply to the three-parameter Weibull distribution when all three parameters are unknown, see, e.g., Rockette et al. (1974) and Hirose (1996).

The large sample properties of the MLE have been extensively studied. Cohen (1965) presented the information matrix of the MLE of the two Weibull parameters for complete samples, singly censored samples and progressively censored samples, respectively. Harter & Moore (1967) presented the maximum-likelihood information matrix for doubly censored samples from the three-parameter Weibull distribution. Thoman et al. (1969) proved the existence of the two pivotal functions of the MLE, i.e., $\hat{\beta}/\beta$ and $\hat{\beta}\ln(\hat{\alpha}/\alpha)$, whose distributions are independent of α and β. With Monte Carlo simulations, they tabulated the percentage points of the distributions of the two pivotal functions which can be used to construct confidence intervals and conduct hypothesis testing regarding the parameters. The authors also pointed out that the distributions of the two pivotal functions are asymptotically normal and provided suggestions on the required sample size to apply the large sample theory for MLE. Billmann et al. (1972) extended the analysis of Thoman et al. (1969) to singly censored samples and proposed their modified pivotal functions of the MLEs.

Numerical methods such as the Newton-Raphson method have to be used to solve the estimating equations of MLE, which were inconvenient at about a half-century ago, and hence simple and closed form approximations for the MLE have been proposed, see, e.g., a series of papers by Bain (1972), Engelhardt (1975) and Engelhardt & Bain (1973, 1974, 1977).

As pointed out by Mann (1967), the MLEs of the Weibull distribution enjoy the properties of consistency, asymptotic efficiency, asymptotic unbiasedness and asymptotic normality. In other words, the estimators have outstanding large sample properties. The small sample properties of the MLEs have become a hot topic since 1990s, and surprisingly, it has been found that the estimators can be highly biased in the cases of small samples and highly censored samples (see, e.g., Jacquelin, 1993; Ross, 1994a; Cacciari et al., 1996). Different methods have been proposed to eliminate or reduce the bias of the ML estimators, especially for the shape parameter estimator. Ross (1994a, 1996) and Hirose (1999) both based on the pivotal function $\hat{\beta}/\beta$, proposed simple bias correcting formulas that can be directly applied to the original ML estimators.

In the meantime, much work can be found that provides analytical or experimental results on the comparison among different parameter estimation methods, see, e.g., Cacciari et al. (1996), Montanari et al. (1997a, b, 1998).

In recent years, the related work of MLE is more for the three-parameter Weibull distribution or the modified/extended Weibull distributions. Abbasi et al. (2006) proposed a new procedure to solve the MLE of the three-parameter Weibull distribution.

Like MLE and MME/MMME, linear order statistics estimation methods have existed for a long time. A great deal of work emerged during the late 1960s and the early 1970s, see, e.g., White (1964), McCool (1965), Mann (1967, 1968), D'Agostino (1971) and Thoman (1972). A common feature of these methods involves transferring the Weibull distribution to the extreme value distribution which has a location-scale form. After the transformation, the estimating equations for the location parameter (i.e., $u = \ln\alpha$) and the scale parameter (i.e., $\sigma = 1/\beta$) of the extreme-value distribution can be expressed by the linear combinations of the order statistics of the transformed observations (i.e., $x = \ln t$) and solved. Several estimators with good statistical properties have been proposed including best linear unbiased estimators (BLUE) (see, e.g., White, 1964; McCool, 1965), best linear invariant estimators (BLIE) (see, e.g., Mann, 1967) and nearly best linear unbiased or invariant estimators (NBLUE or NBLIE) (see, e.g., Thoman, 1972). The estimators of α and β can be obtained from the estimators of u and σ, respectively, based on the relationships of $u = \ln\alpha$ and $\sigma = 1/\beta$; however, since both are of nonlinear relationships, the estimators of α and β will probably not be unbiased. Moreover, these methods normally involve one or several reference tables proposed by the respective authors and a look-up of the reference tables is required upon practical application. This greatly limits their applications.

The LSE method is basically the analytical version of the WPP method. Like WPP, it involves the estimation of failure probability at each failure data point. The related work on the estimation of failure probabilities, or similarly, the determination of the *Y*-axis plotting positions, has been described in Section 1.3.1. The LSE method can also be treated as a special case of the linear order statistics estimation methods.

LSE is less discussed compared to MLE, MME and other linear estimation methods and the traditional opinion among researchers considers it as a simple and inaccurate method (similar to the graphical estimation methods) and it is suggested to provide the start values of parameters for other more sophisticated estimation methods such as MLE. However, in the 1990s, some researchers, see, e.g., Montanari et al. (1997a, b, 1998), in their examination of MLE, compared MLE and LSE via Monte Carlo simulations, and their results showed that the bias of the least squares estimator (LSE) can be much smaller than the bias of the MLE for estimating the shape parameter for complete data, singly censored data and multiply censored data. Ross (1999) presented another intensive comparison between MLE and LSE (with several plotting positions) and reached the similar conclusion that for estimating the shape parameter, the performance of the LSE method with either the median rank plotting positions or the mean plotting positions, is not worse than that of the MLE method in dealing with small samples, and both are biased. Based on the results, Ross suggested that ANSI/IEEE Std 930-1987 (IEEE Guide for the Statistical Analysis of Electrical Insulation Voltage Endurance Data, 1987, sec. 4.1) change the statement that LSE is less accurate than MLE.

Weighted least squares estimation (WLSE) methods for the Weibull distribution have been discussed by some researchers. White (1969) briefly described a WLSE method and gave a numerical example. The weights used in the White's method are tabulated for certain sample sizes. This method can be treated as the traditional WLSE method but the calculation of weights is rather complicated. More recently, Bergman (1986), Faucher & Tyson (1988), Hung (2001) and Lu et al. (2004) each proposed a simple formula for calculating weights based on different approaches to approximate the variances of the predictor variable values. They all demonstrated that their WLSE techniques are more efficient than LSE for estimating the Weibull parameters. Lu et al. (2004) also presented an overview of the WLSE methods, except the traditional method of White (White, 1969), and compared them via Monte Carlo simulations. Theoretically, the traditional method has the best statistical foundation while the ‘new’ ones are simpler and more convenient for application. It is necessary to check the performance of these ‘new’ methods on parameter estimation using the traditional method as a reference.

Besides LSE and WLSE, Lawson et al. (1997) examined some robust M-estimators for the Weibull parameters and compared them with the LSEs for complete and censored data sets with and without outliers. The authors concluded that the robust M-estimation methods outperform LSE in view of both model statistics and parameter estimates. With a bunch of existing robust regression techniques, the robust regression estimation (RRE) methods can be further explored.

Nonlinear estimation methods have also been discussed by some researchers. Berger & Lawrence (1974), via Monte Carlo simulations, concluded that the nonlinear regression technique performs similar to, if not worse than, the LSE method. Somboonsavatdee et al. (2007) pointed out that the graphical estimators (WPP and LSE) are especially useful with censored data.

Finally, there are other estimation methods such as methods of percentiles, see, e.g., Seki & Yokoyama (1993), Wang & Keats (1995), Mark (2005), Bayesian estimation methods, see, e.g., Kaminskiy & Krivtsov (2005), Soliman et al. (2006), and modified profile likelihood methods, see, e.g., Yang & Xie (2003), Ferrari et al. (2007).

1.3.3 Summary and Research Gaps

A summary of the existing parameter estimation methods for the Weibull distribution is shown in Table 1-3. The methods are divided into two large categories: graphical estimation methods and analytical estimation methods. Analytical estimation methods are further divided into five small groups: MME/MMME, MLE, linear order statistics estimation methods, linear regression estimation methods and others. LSE, WLSE and RRE are all related to linear regression techniques and hence this category of methods is named *linear regression estimation methods*.

Table 1-3: Summary of existing parameter estimation methods for the Weibull distribution.

Category	*Methods*	*Related Work*
Graphical Estimation methods	WPP	Weibull (1939), Bernard & Bosi-Levenbach (1953), Kaplan & Meier (1958), Kao (1959), Herd (1960), Johnson (1964), Weibull (1967), Nelson & Thompson (1971), Filliben (1975), Cran (1976), Mischke (1979), Fothergill (1990), Ross (1994b), Hastings & Bartlett (1997), Campean (2000), Skinner et al. (2001), Hossain & Zimmer (2003), Wang (2001, 2004), Wu et al. (2006), Tiryakioglu & Hudak (2007), etc.
	Hazard plotting	Nelson (1972, 2004), Breyfogle (1992), etc.
Analytical Estimation Methods	MME/MMME	Dubey (1966), Mann (1968), Newby (1980), Arora (2000), Murthy et al. (2004), etc.
	MLE	Cohen (1965), Harter & Moore (1965, 1967), Mann (1967), Thoman et al. (1969), McCool (1970), Billmann et al. (1972), Bain (1972), Rockette et al. (1974), Engelhardt (1975), Engelhardt & Bain (1973, 1974 and 1977), Jacquelin (1993), Cacciari et al. (1996), Ross (1994a, 1996), Hirose (1996, 1999), , Montanari et al. (1997a,b, 1998), Abbasi et al. (2006), etc.
	Linear order statistics estimation methods (BLUE, BLIE, NBLUE, NBLIE, etc.)	White (1964), McCool (1965), Mann (1967, 1968), D'Agostino (1971), Thoman (1972), etc.
	Linear regression estimation methods (LSE, WLSE, RRE)	White (1969), Berger & Lawrence (1974), Bergman (1986), Faucher & Tyson (1988), Hung (2001), Lawson et al. (1997), Montanari et al. (1997a,b, 1998), Ross (1994b, 1999), Lu et al. (2004), etc.
	Others (nonlinear estimation methods, method of percentile, Bayesian methods, etc.)	Berger & Lawrence (1974), Seki & Yokoyama (1993), Wang & Keats (1995), Yang & Xie (2003), Kaminskiy & Krivtsov (2005), Mark (2005), Soliman et al. (2006), Ferrari et al. (2007), etc.

As it can be observed from Table 1-3, a majority of the work on Weibull parameter estimation methods was conducted between 1960 and 1980. However, many of them are seldom used nowadays such as the traditional MME/MMME methods and the linear order statistics estimation methods. Recently, Tiryakioglu & Hudak (2007) pointed out that the moments method should be used only when the sample size is more than 14 and the shape parameter is larger than 20. The linear order statistics estimation methods, as mentioned previously, can generate estimators of u and σ ($u = \ln\alpha$ and $\sigma = 1/\beta$) with good statistical properties, but the estimators of α and β are biased. Besides, the methods in this group are normally inconvenient in view of practical applications.

The recent work on the Weibull parameter estimation methods has focused on one or several of the following aspects:

- Bias correction methods
- Estimation based on small samples
- Estimation based on censored data or field data
- Robust estimation methods
- Bayesian estimation methods or others

In fact, WPP, MLE and LSE have become the most popular and widely used parameter estimation methods for the Weibull distribution. WPP is a graphical method which can serve as a simple tool for model validation and outlier detection. MLE is considered to have good statistical perspectives and is preferred by researchers, while WPP and LSE are frequently used by practitioners because of the simplicity and graphical presentation. For example, LSE is the standard parameter estimation method for the Weibull distribution in soil studies (Munkholm & Perfect, 2005). MLE has been intensively examined in the literature, where both large sample properties and small sample properties of its estimators have been investigated. In the 1990s, some researchers found that the estimators of MLE and LSE are both highly biased in the cases of small samples and censored samples, see, e.g., Montanari et al. (1997a, b, 1998), which could raise a warning message. Several bias correction methods have been proposed for the MLE method, see, e.g., Ross (1994a, 1996) and Hirose (1999). However, there are no bias correction methods for LSE. Indeed, the LSE method is less discussed by researchers. Previously we have mentioned that reliability data analysis requires different approaches for different types of data, and the group of linear regression methods can satisfy this purpose because, as is well-known, different regression techniques, such as WLS and robust regression, are good at handling certain data types. LSE, as the simplest method in the group of linear regression estimation methods, can be refined or replaced by other methods in the group to achieve better estimation results. In summary, LSE and other linear regression estimation methods have good potentials compared to MLE, but little work has been done to explore them.

1.4 Scope of the Thesis

This thesis focuses on the linear regression estimation methods including LSE for the Weibull distribution. WPP is presented together with the linear regression estimation methods because they can be easily combined. The proposed estimation methods are frequently compared with the MLE method because of its wide application. Other estimation methods in Table 1-3 such as MME/MMME are beyond the scope of this thesis.

Harsh data conditions including small samples, highly censored samples, and/or samples with outliers are central to this study mainly because they are very common in the field and they are the recent interests of Weibull researchers.

1.5 Research Objectives and Significance

The purpose of this thesis is to refine the conventional LSE (or ordinary LSE, or OLSE) method and develop new linear regression estimation methods for the Weibull distribution to deal with harsh data conditions such as small samples, highly censored samples, and/or samples with outliers. Several simple methods are proposed that can be easily applied and understood. The specific aims are listed as follows:

1) Thoroughly investigate the properties of the OLS estimators of the two Weibull parameters via both theoretical analysis and intensive Monte Carlo simulation experiments (Chapter 3).

2) Provide suggestions on the application procedures of the LSE method including the selection of failure probability estimator and the regression direction, applied to complete data and censored data, respectively (Chapter 4).

3) Propose simple bias correcting formulas for the OLS shape parameter estimator, applied to small and complete data, and censored data with low censoring levels (Chapter 5).

4) Discuss the existing WLSE methods for the Weibull distribution and propose new methods for calculating weights for complete data and censored data, respectively (Chapter 6).

5) Examine various robust regression techniques and develop robust M-estimation methods for the Weibull distribution to replace OLSE in order to deal with outliers (Chapter 7).

6) Provide application instructions on the linear regression estimation methods discussed in this study with numerical examples (Chapter 8).

The LSE method is basically the application of simple linear regression. Therefore, it is clear that the existing theories and various linear regression techniques can be applied to improve or replace the LSE method to deal with various data types. We will examine WLS regression techniques and robust regression techniques. The step-by-step procedures will be provided for the application of these methods. Moreover, the names and versions of common statistical software packages that can be used to obtain quick results will be mentioned. To reduce the bias of the OLSE of the shape parameter, bias correction methods will be proposed. The proposed simple bias correcting formulas can be added to the end of the conventional OLSE procedure to provide more accurate estimates without adding computation complexity.

The results of this study should give researchers a better understanding of the theories of LSE and other linear regression estimation methods. The proposed methods will be of great practical value for practitioners conducting reliability data analysis. Moreover, it may lead to a better understanding of the roles of LSE and WPP among all existing Weibull parameter estimation methods.

Chapter 2

Basic Weibull Parameter Estimation Methods

This chapter describes three nowadays most widely used parameter estimation methods for the Weibull distribution, i.e., WPP, LSE and MLE. The theoretical backgrounds of these methods are presented. Common criteria for comparing estimation methods and estimators are described.

2.1 Introduction and Notations

Now suppose there is a random sample from a life testing experiment. Assume the underlying distribution is the Weibull distribution. This sample can be denoted as $t_1, t_2, \cdots, t_i, \cdots, t_n \quad (i = 1,2,\cdots,n)$. Based on the experiment schemes, it can be a complete sample where all the observations are failures, or it can be a censored sample where some of the observations are failures and the others are censors. In this thesis, multiply censored samples are used as the general case for censored life data. For a multiply censored sample, let $t_{f,1}, t_{f,2}, \cdots, t_{f,j}, \cdots, t_{f,r} \ (j = 1, 2, \cdots, r)$ denote the failure times and $t_{c,1}, t_{c,2}, \cdots, t_{c,k}, \cdots, t_{c,(n-r)} \ (k = 1, 2, \cdots, n-r)$ denote the censoring times.

The order statistics of the observations are used in the LSE method since the failures occur in sequence. Let $t_{(i)}$ denotes the i^{th} smallest failure time in a complete sample, i.e., $t_{(1)} \le t_{(2)} \le \cdots \le t_{(i)} \le \cdots \le t_{(n)}$. For a multiply censored sample with r failures and $n-r$ censors, let $t_{f,(j)}$ denotes the j^{th} smallest failure time and $t_{c,(k)}$

denotes the k^{th} smallest censoring time, so $t_{f,(1)} \le t_{f,(2)} \le \cdots \le t_{f,(j)} \le \cdots \le t_{f,(r)}$ and $t_{c,(1)} \le t_{c,(2)} \le \cdots \le t_{c,(k)} \le \cdots \le t_{c,(n-r)}$. Table 2-1 provides an illustration of the notations with a numerical example.

Table 2-1: An illustration of the notations with a numerical example.

Unit No	Failure (F) / Censor (C) Indicator	Age (hour)	Notations of the Original Sample	Notations of the Order Statistics of the Sample Without Failure/Censor Indicator	Notations of the Order Statistics of the Sample With Failure/Censor Indicator
1	F	290	t_1	$t_{(2)}$	$t_{f,(2)}$
2	C	1000	t_2	$t_{(7)}$	$t_{c,(3)}$
3	F	133	t_3	$t_{(1)}$	$t_{f,(1)}$
4	F	470	t_4	$t_{(3)}$	$t_{f,(3)}$
5	C	500	t_5	$t_{(4)}$	$t_{c,(1)}$
6	F	700	t_6	$t_{(5)}$	$t_{f,(4)}$
7	C	800	t_7	$t_{(6)}$	$t_{c,(2)}$

The objective of parameter estimation is to estimate α and β using sample data. In the following of this chapter, the theoretical backgrounds and the estimation equations (except WPP) of three common estimation methods of the Weibull distribution, i.e., WPP, LSE and MLE, are separately presented in Sections 2.2, 2.3 and 2.4. Finally, Section 2.5 describes the common criteria for comparing different estimation methods and their estimators.

2.2 Weibull Probability Plot and *Y*-axis Plotting Positions

WPP is a traditional graphical method for estimating the Weibull parameters. Proposed by Kao (1959), it is still widely used nowadays for Weibull analysis. WPP, in addition to providing simple parameter estimates, it serves the purpose of simple model validation and outlier identification which are very important in any engineering data analysis.

WPP is based on the linearization of the Weibull CDF in Equation (1-1). The linearized Weibull CDF is given by

$$\ln\left[-\ln(1-F(t))\right]=\beta\ln t-\beta\ln\alpha \tag{2-1}$$

Weibull probability paper is specially scaled based on Equation (2-1) so that it shows a straight line if the Weibull distribution fits the sample data. Its *X*-axis represents the observations t (i.e., failure time) from a life testing experiment or the field. The *Y*-axis represents the cumulative probability of failure $F(t)$. From the Weibull CDF, the value of $F(t)$ at each failure data point are unknown without the values of α and β and hence can only be estimated. Similar to other probability plotting methods, for example, the normal probability plotting, non-parametric estimators of $F(t)$ with a general form of $(i-c_1)/(n+c_2)$ are frequently used to obtain the *Y*-axis plotting positions. As is well known, $(i-3/8)/(n+1/4)$ (Blom, 1958) is used for the normal probability plotting. As for WPP, the selection of the method to obtain the *Y*-axis plotting positions depends on whether the sample is complete or censored. In the following, the theoretical backgrounds of the commonly used *Y*-axis plotting positions covered in the reliability textbooks for complete samples and censored samples are briefly presented.

<u>Theoretical Backgrounds of Commonly Used Y-axis Plotting Positions on WPP</u>

The complete samples are considered first. A common practice when using probability theory to analyze the order statistics of random samples from a continuous distribution (the parent distribution) considers the probability F as uniformly distributed between 0 and 1, and hence its order statistic $F_{(i)}$ has a beta distribution with parameters i and $n-i+1$. The mean and median of this beta distribution are commonly used for the *Y*-axis plotting positions. The mean has a simple form, i.e.,

$$E(F_{(i)}) = \frac{i}{n+1} \tag{2-2}$$

Professor Weibull originally used Equation (2-2) (Weibull, 1939) and this is then named Weibull plotting position or Weibull estimator. Theoretically, it is the exact mean rank plotting position of each data point.

The median of $F_{(i)}$ is related to the incomplete beta function. It is the solution of

$$i\binom{n}{i}\int_0^{Median(F_{(i)})} p^{i-1}(1-p)^{n-i}\,dp = \frac{1}{2} \tag{2-3}$$

The exact median values at different combinations of i and n can be obtained using numerical methods. One can also lookup the standard tables of the percentage points of the incomplete beta distribution (see, e.g., Gibbons et al., 1999) to get quick results.

The median rank plotting position in Equation (2-3) is more favored than the mean rank plotting position in Equation (2-2) by Weibull researchers. Simple approximations have been proposed for the median rank plotting position, among which the Bernard estimator (Bernard & Bosi-Levenbach, 1953) has been widely used nowadays. The Bernard estimator is given by

Bernard estimator

$$\hat{F}_{(i)} = \frac{i-0.3}{n+0.4} \tag{2-4}$$

Another popular source for $\hat{F}_{(i)}$ is the Hazen estimator in Equation (2-5). It is also known as the midpoint probability estimator since it is the middle value of the interval from $(i-1)/n$ to i/n (Kimball, 1960). The Hazen estimator is used as the default method for *Y*-axis plotting positions in the WPP program of MATLAB 7.

Hazen estimator $$\hat{F}_{(i)} = \frac{i-0.5}{n} \tag{2-5}$$

From the theoretical backgrounds of the above estimators (Weibull, Bernard and Hazen), it is clear that all of them have no relationship with the Weibull CDF. In other words, these are distribution-free plotting positions. In Section 4.3, plotting positions related to the Weibull CDF will be presented.

Estimation of the failure probabilities for censored Weibull samples is a challenge and the above mentioned estimators should not be directly used. It is important to note that WPP and the group of linear regression estimation methods discussed throughout this study only plots, or in the analytical cases uses, the failure times. The influence of censoring can be reflected in the estimation of failure probability at each failure data point.

Similar to the common estimators of $F_{(i)}$ for complete samples, failure probability estimators that are independent of failure time are frequently used for censored samples. In the following, let $\hat{F}_{f,(j)}$ denotes the failure probability estimator for the j^{th} failure in a censored sample, i.e., $\hat{F}_{f,(j)} = \hat{F}(t_{f,(j)})$. The Herd-Johnson method (Herd, 1960; Johnson, 1964) is most widely used for estimating failure probabilities for censored data. It is given by

Herd-Johnson estimator $$\begin{cases} \hat{R}_{f,(j)} = \left(\dfrac{n+1-I_j}{n+2-I_j} \right) \cdot \hat{R}_{f,(j-1)} \\ \hat{F}_{f,(j)} = 1 - \hat{R}_{f,(j)} \end{cases} \tag{2-6}$$

where I_j denotes the event number of the j^{th} failure in the sample. The occurrence of a failure and a censor are both considered as an event. $\hat{R}_{f,(j)}$ is the supplement of $\hat{F}_{f,(j)}$ and $\hat{R}_{f,(0)} = 1$.

The theoretical background for the derivation of Equation (2-6) is briefly described as follows. Assume there is a multiply censored sample of size n in which r failures ($0 < r < n$) and $n - r$ censors are intermixed along the time axis. Let $t_{(1)}, t_{(2)}, \cdots, t_{(i)}, \cdots, t_{(n)}$ denote the ordered observations. We call $t_{(i)}$ an event and it can be a failure event or a censor event. Let $t_{f,(1)}, t_{f,(2)}, \cdots, t_{f,(j)}, \cdots, t_{f,(r)}$ ($1 \le r \le n$) be the ordered failure events. From the definition of I_j one can obtain $t_{f,(j)} = t_{(I_j)}$. Censoring times can lie in one of the intervals constructed by failure times, i.e., $\left(0, t_{f,(1)}\right]$, $\left[t_{f,(j-1)}, t_{f,(j)}\right]$ ($1 < j \le r$), and $\left[t_{f,(r)}, +\infty\right)$. The Herd-Johnson method first assumes that a censor happens concurrently with a failure event, say for example, if the censoring time lies in the interval $\left[t_{f,(j-1)}, t_{f,(j)}\right]$, it is treated as happening at $t_{f,(j-1)}$. Now consider a censor which occurs at $t_{f,(j-1)}$, if allowed to continue the test, it may fail in any of the intervals between two consecutive events $\left(t_{(i-1)}, t_{(i)}\right]$, where $t_{(i-1)} \ge t_{f,(j-1)}$, or the interval following the final event, denoted by $\left[t_n, +\infty\right)$, and there is a total of $n - I_j + 2$ possible intervals. By assuming the probabilities of failing in any of the intervals are equal, the probability of failing in $\left[t_{f,(j-1)}, t_{f,(j)}\right]$ is then $1/(n - I_j + 2)$, or the probability of surviving in $\left[t_{f,(j-1)}, t_{f,(j)}\right]$ is $(n - I_j + 1)/(n - I_j + 2)$. Applying the multiplication rule of conditional probability, we obtain

$$\Pr\{T \ge t_{f,(j)}\} = \Pr\{T \ge t_{f,(j)} | T \ge t_{f,(j-1)}\} \cdot \Pr\{T \ge t_{f,(j-1)}\} = \frac{n - I_j + 1}{n - I_j + 2} \cdot \Pr\{T \ge t_{f,(j-1)}\}$$

$$\Rightarrow R(t_{f,(j)}) = \frac{n - I_j + 1}{n - I_j + 2} \cdot R(t_{f,(j-1)})$$

A numerical example is given to illustrate the Herd-Johnson method, as shown in Figure 2-1 and Table 2-2. Figure 2-1 plots the ordered observations in the sample along a time axis and Table 2-2 shows the calculation of $\hat{F}_{f,(j)}$ at each failure data point.

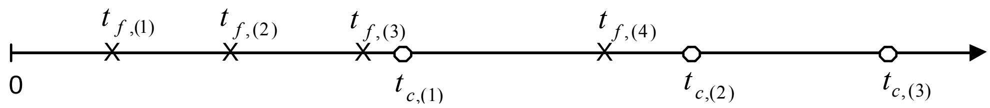


Figure 2-1: A numerical example of the Herd-Johnson method: ordered events along a time axis ("x" denotes failure and "o" denotes censor).

Table 2-2: A numerical example of the Herd-Johnson method: calculation of $\hat{F}_{f,(j)}$.

$t_{f,(j)}$	j	I_j	$\hat{R}_{f,(j)}$	$\hat{F}_{f,(j)}$
133	1	1	$\frac{7-1+1}{7-1+2} = \frac{7}{8}$	$1 - \frac{7}{8} = \frac{1}{8}$
290	2	2	$\frac{7-2+1}{7-2+2} \times \frac{7}{8} = \frac{6}{7} \times \frac{7}{8} = \frac{3}{4}$	$1 - \frac{3}{4} = \frac{1}{4}$
470	3	3	$\frac{7-3+1}{7-3+2} \times \frac{3}{4} = \frac{5}{6} \times \frac{3}{4} = \frac{5}{8}$	$1 - \frac{5}{8} = \frac{3}{8}$
700	4	5	$\frac{7-5+1}{7-5+2} \times \frac{5}{8} = \frac{3}{4} \times \frac{5}{8} = \frac{15}{32}$	$1 - \frac{15}{32} = \frac{17}{32}$

The estimation of failure probability is an important issue that affects both goodness-of-fit and parameter estimation results. Many researchers have investigated the issue and different methods have been compared and favored, see, e.g., Fothergill (1990), Cacciari & Montanari (1991). For censored data, besides the Herd-Johnson estimator which is a non-parametric estimator, some parametric estimators, see, e.g., Wang (2001, 2004), have been proposed. Although some estimators, e.g., the Bernard estimator used for complete samples and the Herd-Johnson estimator used for censored samples, are more frequently used than the others, by now the agreement has not been reached and the discussion is ongoing. Section 4.3 will further explore the issue.

Application Procedure of WPP

A widely used procedure of WPP is to plot t along the horizontal axis and the estimated values of $F(t)$, commonly called *Y*-axis plotting positions, along the vertical axis, on the Weibull probability paper. As a traditional way, a straight line is fitted to the points by eye; however, more objective estimates can be obtained by fitting the straight line via the least squares regression technique. The shape parameter is then estimated by the slope of the regression line and the scale parameter is estimated by either the exponential of the ratio of the regression line's intercept to slope, or the value of t when $F = 0.632$ (see Equation (1-7)).

WPP can be easily generated by common statistical software packages such as MATLAB, SAS, S-PLUS and MINITAB. Table 2-3 summarizes the syntax (for MATLAB and SAS) or dialogs (for S-PLUS and MINITAB) used in these software packages to generate a WPP and their default straight line fitting techniques, including the default *Y*-axis plotting positions, if applicable. As can be seen from the table, MATLAB 7 uses the LS fit with the Hazen estimator (i.e., Equation (2-5)) for *Y*-axis plotting positions by default, S-PLUS 6 provides both LS fit and MLE fit, while SAS 9 and MINITAB 14 use MLE fit by default. The MLE fit is not traditional for the WPP; however, has gained some popularity since researchers favor the MLE method for parameter estimation. If the MLE fit is used, the *Y*-axis plotting positions are directly calculated by the Weibull CDF with the ML estimates of the two Weibull parameters. The practitioners should be cautious about the MLE fit because it tends to overestimate the shape parameter for small samples. The use of the Kaplan-Meier estimator for the *Y*-axis plotting positions in SAS 9 and S-PLUS 6 is inappropriate. As mentioned in Section 1.3.1, a big disadvantage of the Kaplan-Meier estimator is that the unreliability for the last failure data point is always 1, and hence it tends to underestimate the failures in the tail of the distribution.

Table 2-3: Summary of the syntax or dialogs for generating WPP with common statistical software packages and their default straight line fitting techniques*.

Software & Version	WPP Syntax/Dialog for Complete Data	WPP Syntax/Dialog for Censored Data	Default Straight Line Fitting Techniques	References
MATLAB 7	wblplot(x) probplot('weibull', x)	probplot('weibull', x, cens, freq)	Least squares. Default *Y*-axis plotting position is the Hazen estimator (i - 0.5)/n, where n is sample size for complete data and number of failures for censored data	http://www.mathworks.com/access/helpdesk/help/helpdesk.html
SAS 9	PROBPLOT variable</options>	PROBPLOT variable<*censor-variable(values)></options>	By default is MLE fit instead of LS fit. If use LS fit, the default plotting position is the modified Kaplan-Meier rank, but we can also use mean rank or median rank (specified in options)	http://support.sas.com/documentation/cdl/en/qcug/59658/HTML/default/rel_intro_sect34.htm#qcug_rel_intro_probopt
S-PLUS 6	SPLIDA ► Single distribution analysis ► Probability plot with nonparametric confidence intervals SPLIDA ► Single distribution life data analyses ► Probability plot with parametric ML fit		Least squares. Default *Y*-axis plotting position is the Kaplan-Meier rank. Straight line generated by ML estimates.	http://www.public.iastate.edu/~wqmeeker/splida/SplidaGui.pdf
MINITAB 14	Graph ► Probability Plot (specify Weibull distribution)		Straight line generated by ML estimates.	http://www.minitab.com/en-CA/support/answers/answer.aspx?ID=1331

MATLAB 7 is used in this study. The default *Y*-axis plotting positions are calculated by the Hazen estimator but can be easily changed to other options. Figure 2-2 gives an example of a computer-generated WPP in MATLAB 7.

* Online references: http://www.mathworks.com/access/helpdesk/help/helpdesk.html (MATLAB); http://support.sas.com/documentation/cdl/en/qcug/59658/HTML/default/rel_intro_sect34.htm#qcug_rel_intro_probopt (SAS); http://www.public.iastate.edu/~wqmeeker/splida/SplidaGui.pdf (S-PLUS); http://www.public.iastate.edu/~wqmeeker/splida/SplidaGui.pdf (MINITAB).

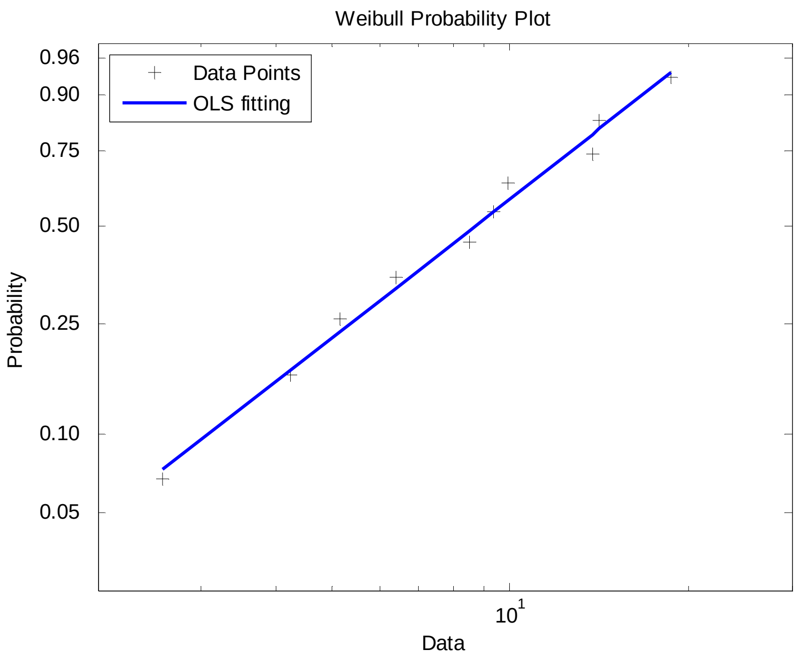


Figure 2-2: An example of a computer-generated WPP in MATLAB 7.

2.3 Least Squares Estimation

The LSE method uses the least squares regression to estimate the two parameters based on the linearized Weibull CDF in Equation (2-1).

As the conventional way, setting $X = \ln T$, $Y = \ln[-\ln(1-F)]$, $A = -\beta \ln \alpha$ and $B = \beta$, Equation (2-1) becomes a simple equation, i.e.,

$$Y = A + BX \tag{2-7}$$

Thus the estimation of α and β can be transferred to the estimation of the regression coefficients for a simple linear regression model of the form $Y = A + BX + e$, where e is the error term.

For a complete data set $t_1, t_2, \cdots, t_i, \cdots, t_n$, the values of X and Y can be calculated by

$$x_i = \ln(t_{(i)}) \text{ and } y_i = \ln\left[-\ln(1-\hat{F}_{(i)})\right] \tag{2-8}$$

For a censored data set where $t_{f,1}, t_{f,2}, \cdots, t_{f,j}, \cdots, t_{f,r}$ denote the failure times, the values of X and Y can be calculated by

$$x_i = \ln(t_{f,(j)}) \text{ and } y_i = \ln\left[-\ln(1-\hat{F}_{f,(j)})\right] \tag{2-9}$$

The common methods used to obtain the values of $\hat{F}_{(i)}$ and $\hat{F}_{f,(j)}$ have been described in Section 2.2.

The objective function of the LSE method is

$$\min\ S = \sum_{i=1}^{r}\left[y_i - (A + Bx_i)\right]^2 \tag{2-10}$$

where for complete data, $r = n$.

By taking partial derivatives of S with regard to A and B, respectively, and setting the results to 0, we obtain

$$\begin{cases} \hat{B} = \dfrac{\sum_{i=1}^{r}\left[(x_i - \bar{x})(y_i - \bar{y})\right]}{\sum_{i=1}^{r}(x_i - \bar{x})^2} = \dfrac{r\sum_{i=1}^{r}x_i y_i - \sum_{i=1}^{r}x_i \cdot \sum_{i=1}^{r}y_i}{r\sum_{i=1}^{r}x_i^2 - \left(\sum_{i=1}^{r}x_i\right)^2} \\ \hat{A} = \bar{y} - \hat{B}\bar{x} = \dfrac{\sum_{i=1}^{r}y_i - \hat{B}\sum_{i=1}^{r}x_i}{r} \end{cases} \tag{2-11}$$

where $\bar{x} = \sum_{i=1}^{r} x_i \Big/ r$ and $\bar{y} = \sum_{i=1}^{r} y_i \Big/ r$.

Based on $A = -\beta \ln \alpha$ and $B = \beta$, the estimating equation related to α and β is then given by

$$\begin{cases} \hat{\beta} = \dfrac{\sum_{i=1}^{r}\left[(x_i - \bar{x})(y_i - \bar{y})\right]}{\sum_{i=1}^{r}(x_i - \bar{x})^2} = \dfrac{r\sum_{i=1}^{r} x_i y_i - \sum_{i=1}^{r} x_i \cdot \sum_{i=1}^{r} y_i}{r\sum_{i=1}^{r} x_i^2 - \left(\sum_{i=1}^{r} x_i\right)^2} \\ \hat{\alpha} = \exp\left(-\dfrac{\bar{y} - \hat{\beta}\bar{x}}{\hat{\beta}}\right) = \exp\left(-\dfrac{\sum_{i=1}^{r} y_i - \hat{\beta}\sum_{i=1}^{r} x_i}{r\hat{\beta}}\right) \end{cases} \quad (2\text{-}12)$$

Equation (2-12) can be applied to both complete data and censored data. For complete data, $r = n$.

2.3.1 The Ordinary/Conventional LSE Method

There are some uncertainties in the LSE method which makes it inappropriate to describe LSE by a single equation like Equation (2-12). Firstly, Equation (2-12) is derived based on the setting of $X = \ln T$ and $Y = \ln[-\ln(1-F)]$. Another option appeared in the literature is to set $Y = \ln T$ and $X = \ln[-\ln(1-F)]$, i.e., to reverse the independent variable and the dependent variable in the regression. This will give another estimating equation for LSE. Discussions for the regression direction are presented in Section 4.4. Secondly, even if the conventional setting of *X* and *Y* is used, different methods for calculating F_i or y_i will result in different estimates for the parameters. A detailed comparison of the various estimators of the *Y*-axis plotting positions on parameter estimation is presented in Section 4.3. Based on the above two points, in fact, LSE has a family of methods.

According to the common practice, the OLSE method refers to the LSE method that 1) sets $X = \ln T$ and $Y = \ln[-\ln(1-F)]$, so that Equation (2-12) is the estimating equation; and 2) for estimating F, the Bernard estimator in Equation (2-4) is used for complete data, and the Herd-Johnson estimator in Equation (2-6) is used for multiply censored data.

Application Procedure of OLSE

Step 1: Rank failure times from smallest to largest and calculate the estimates for failure probability at each failure data point. For complete data, use the Bernard estimator, i.e., Equation (2-4), to calculate $\hat{F}_{(i)}$. For censored data, use the Herd-Johnson estimator, i.e., Equation (2-6), to calculate $\hat{F}_{f,(j)}$.

Step 2: Calculate x_i and y_i. For complete data, use Equation (2-8). For censored data, use Equation (2-9).

Step 3: Estimate α and β using Equation (2-12).

2.4 Maximum Likelihood Estimation

MLE is one of the most widely used tools for statistical inference. Cohen (1965) introduced the maximum likelihood equations for estimating the two Weibull parameters from complete samples, Type I or Type II singly censored samples and multiply censored samples, respectively. The likelihood function for complete Weibull samples is given by

$$L = \prod_{i=1}^{n} f(t_i) = \prod_{i=1}^{n} \left(\beta t_i^{\beta-1} / \alpha^{\beta}\right) \exp\left[-\left(t_i / \alpha\right)^{\beta}\right] \tag{2-13}$$

Taking logarithm of L, differentiating with respect to α and β and equating to 0, the estimating equation can be obtained as

$$\begin{cases} \dfrac{\sum_{i=1}^{n} t_i^{\hat{\beta}} \ln t_i}{\sum_{i=1}^{n} t_i^{\hat{\beta}}} - \dfrac{1}{\hat{\beta}} = \dfrac{1}{n}\sum_{i=1}^{n} \ln t_i \\ \hat{\alpha} = \left(\sum_{i=1}^{n} t_i^{\hat{\beta}} / n \right)^{1/\hat{\beta}} \end{cases} \tag{2-14}$$

The likelihood function for singly censored samples, either Type I or Type II censored, is given by

$$\begin{aligned} L &= C\left[\prod_{j=1}^{r} f(t_i)\right] \cdot \left[1 - F(t_T)\right]^{n-r} \\ &= C\left\{\prod_{j=1}^{r} \left[\left(\beta t_i^{\beta-1} \middle/ \alpha^{\beta}\right) \exp\left[-(t_i/\alpha)^{\beta}\right]\right]\right\} \cdot \exp\left[-(n-r)(t_T/\alpha)^{\beta}\right] \end{aligned} \tag{2-15}$$

where C is a constant. For Type I censoring, t_T is the predetermined time of termination, and for Type II censoring, t_T is the time at the r^{th} failure, i.e., $t_T = t_r$.

The estimating equation of MLE for singly censored data is

$$\begin{cases} \dfrac{\sum_{i=1}^{r} t_i^{\hat{\beta}} \ln t_i}{\sum_{i=1}^{r} t_i^{\hat{\beta}}} - \dfrac{1}{\hat{\beta}} = \dfrac{1}{r}\sum_{i=1}^{r} \ln t_i \\ \hat{\alpha} = \left[\left(\sum_{i=1}^{r} t_i^{\hat{\beta}} \right) \middle/ r \right]^{1/\hat{\beta}} \end{cases} \tag{2-16}$$

For multiply censored samples, the likelihood function is given by

$$\begin{aligned} L &= C\prod_{j=1}^{r} f(t_{f,j}) \prod_{k=1}^{n-r} \left[1 - F(t_{c,k})\right] \\ &= C\prod_{j=1}^{r} \left\{\left(\beta t_{f,j}^{\beta-1} / \alpha^{\beta}\right) \exp\left[-\left(t_{f,j}/\alpha\right)^{\beta}\right]\right\} \cdot \prod_{k=1}^{n-r} \exp\left[-\left(t_{c,k}/\alpha\right)^{\beta}\right] \end{aligned} \tag{2-17}$$

and the estimating equation is

$$\begin{cases} \dfrac{\sum_{j=1}^{r} t_{f,j}^{\hat{\beta}} \ln t_{f,j} + \sum_{k=1}^{n-r} t_{c,k}^{\hat{\beta}} \ln t_{c,k}}{\sum_{j=1}^{r} t_{f,j}^{\hat{\beta}} + \sum_{k=1}^{m} t_{c,k}^{\hat{\beta}}} - \dfrac{1}{\hat{\beta}} = \dfrac{1}{r}\sum_{j=1}^{r} \ln t_{f,j} \\ \hat{\alpha} = \left[\left(\sum_{j=1}^{r} t_{f,j}^{\hat{\beta}} + \sum_{k=1}^{n-r} t_{c,k}^{\hat{\beta}} \right) \Big/ r \right]^{1/\hat{\beta}} \end{cases} \tag{2-18}$$

The Newton-Raphson method is frequently used to solve the estimating equations, i.e., Equations (2-14), (2-16) and (2-18). Although the calculation is complicated, nowadays many statistical software packages such as MATLAB, SAS, S-PLUS and MINITAB have embedded programs for calculating the ML estimates. Electronic spreadsheets such as Excel can also solve the estimating equations of MLE, see, e.g., Tang (2003) for a numerical example.

2.5 Comparison of Estimation Methods and Estimators

As previously stated, parameter estimation usually serves as a preliminary step of Weibull analysis and the parameter estimates may greatly affect the business decisions making on the subsequent steps. Different parameter estimation methods can generate widely differing estimates; therefore, it is important to have objective criteria to instruct the selection of one estimation method over the other alternatives. Tobias & Trindade (1995) gave four most desirable attributes for estimation methods. Their descriptions are quoted below.

- ***Lack of bias***: The expected value of the estimate equals the true parameter.

- ***Minimum variance***: The estimator of the selected method has less variability on the average than any estimators. If this estimator is also unbiased, it is likely to be closer to the true value than other estimators.
- ***Sufficiency***: The estimate makes use of all the statistical information available in the data.
- ***Consistency***: The estimate tends to get closer to true value with larger sample size (infinite samples yield perfect estimates).

In view of the application perspectives for engineers, we add another desirable attribute,

- ***Simplicity***: The method does not involve complicated calculation and sophisticated statistical knowledge. In short, it can be easily understood and easily applied.

Also there are commonly used criteria for comparing parameter estimators including bias, variance (or standard deviation), mean square error (MSE), efficiency, consistency and robustness. The following descriptions talk about how these terms are measured.

- ***Bias***: The difference between the expected value of a statistic and the parameter value which it estimates. An estimator is said to be unbiased if in the long run it takes on the value of the population parameter.
- ***Variance***: The expected value of the squares of the difference between the values of the estimates and the mean of them.
- ***MSE***: The expected value of the squares of the differences between the values of the estimates and the parameter value. MSE can also be calculated by the sum of the variance and the squared bias of the estimator.

- ***Efficiency***: The ratio of the variances of two estimators. Sometimes, we will select an estimator with a small amount of bias but a high efficiency.
- ***Consistency***: Estimator that converges in probability to the quantity being estimated as the sample size grows. The performance of a consistent estimator improves with the increase of sample size.
- ***Robustness***: The properties of the estimator when the assumptions used in the parameter estimation method are not valid. A common situation is the properties of the estimator in the presence of outliers.

For the three basic estimation methods described in this chapter, WPP is the simplest method and it can serve as a simple tool for model validation and outlier identification. MLE is considered to have good statistical perspectives since it is asymptotically unbiased, asymptotically efficient and consistent. Compared to MLE, the LSE method has some advantages: 1) it has a closed form solution which can be easily calculated; 2) it can be easily incorporated into WPP and the different ways of obtaining the *Y*-axis plotting positions adds its flexibility; and 3) the properties of the LS estimators including bias and MSE are not inferior to those of the ML estimators, especially under harsh data conditions such as small samples and highly censored samples. In the next chapter, the properties of the OLS estimators are discussed in details. The simulation results will be given on the comparison of OLSE and MLE for both parameter estimators.

Chapter 3

Properties of the OLS Estimators

This chapter explores the properties of the OLS estimators for the Weibull distribution through two approaches: analytical examination and Monte Carlo experimental examination. The results suggest the possibility and directions to improve the OLSE method.

3.1 Introduction

The OLSE method is widely used by practitioners conducting Weibull analysis. The analytical background and application procedure of the method, and the relationship between OLSE and LSE in the general sense, have been described in Section 2.3.

As previously mentioned, the traditional viewpoint toward LSE considers it as a simple but inaccurate method for Weibull parameter estimation, compared with other analytical estimation methods such as MLE. As a result, this method has been overlooked by many researchers and it was not until the last decade that some researchers, based on Monte Carlo simulations, pointed out that the properties such as bias and MSE of the OLSE of the Weibull shape parameter outperform those of the MLE for small samples and highly censored samples, see, e.g., Montanari et al., (1997a, b, 1998).

This chapter presents a detailed examination of the OLS estimators of both Weibull parameters. Firstly, using the knowledge of least squares regression or the Gauss-Markov theorem, we clarify why the OLS estimators of the Weibull parameters are not BLUE and discuss how the selection of the *Y*-axis plotting positions will affect the bias of the estimators via analytical methods. Moreover, the existence of two pivotal functions, $\hat{\beta}/\beta$ and $\hat{\beta}\ln(\hat{\alpha}/\alpha)$, of the LS estimators, regardless of the determination of the *Y*-axis plotting positions, and for both complete and censored data, is proved. Secondly, the method of using Monte Carlo simulation experiments to determine the bias, variance and MSE of the OLS estimators is described. The experiment procedures, setting of experiment factors, and experiment results are presented. Finally, the results from both analytical examinations and experiment examinations are summarized.

3.2 Analytical Examinations of the OLS Estimators

3.2.1 OLS Estimators Are Not BLUE

As pointed out in Section 2.3, LSE transfers the estimation of α and β to the estimation of the two regression coefficients for a simple linear regression model of the form $Y = A + BX + e$, where $A = -\beta\ln\alpha$, $B = \beta$ and e is the error term. The LS estimators of α and β can be obtained via the LS estimators of A and B.

According to the Gauss-Markov theorem, for a simple linear regression model $Y = A + BX + e$, if certain assumptions are satisfied, the LS estimators of A and B will be BLUE, i.e., unbiased and have minimum MSE among all linear estimators of A and B (Allen, 1997, pp. 182-185). These assumptions are

i. The expected value or mean of the population errors is zero. This assumption can be mathematically stated as $E(e_i) = 0$.

ii. The variance of the errors is constant for all values of the independent variable. This assumption is also known as the homoscedasticity condition. Mathematically, the assumption can be expressed by $Var(e_i) = \sigma^2$ for all i.

iii. The errors are independent of each other. Mathematically, it can be expressed by $Cov(e_i, e_j) = 0$ for all i, j.

iv. The errors and the independent variable are independent, i.e., $Cov(e_i, x_i) = 0$.

There is no specification on the distribution of the error; however, if the error is normality distributed, the LS estimators of A and B will be the best unbiased estimator (BUE) among all linear and nonlinear estimators.

If the above assumptions are satisfied, the BLUE of β and $\ln\alpha$ can be obtained via the BLUE of A and B based on $B = \beta$ and $A = -\beta\ln\alpha$. The estimator of α is not BLUE because *ln* is not a linear operation, but this is not especially problematic since in most times only β is of importance.

In the most common simple linear regression scenario, the values of *X* are treated as known constants set by a design and the values of *Y* are measured conditionally on the values of *X* in an experiment. This does not meet the background of the LSE method because here both *X* and *Y* are random variables and the values of *Y* cannot be measured but estimated. To check the assumptions ***i*** – ***iv*** for the linear regression model of the OLSE method, first assume the uncertainty of T or $X = \ln T$ is much smaller than the uncertainty of Y ($Y = \ln[-\ln(1-F)]$), thus the uncertainty of e can be confined to Y. This assumption justifies the regression direction of *Y* on *X* used in OLSE. With this assumption and also note that the values of x_i and y_i used in the estimating equation, i.e., Equation (2-12), come from the order statistics of the variables, the problem now is to examine $E(Y_{(i)}), Var(Y_{(i)})$ and $Cov(Y_{(i)}, Y_{(j)})$. From the knowledge of order statistics, it is clear that $Var(Y_{(i)})$, which is a function of the order number *i*, is not a constant, and any two order statistics, e.g., $Y_{(i)}$ and $Y_{(j)}$, are correlated. Therefore, assumptions ***ii*** and ***iii*** are usually inappropriate. The analytical expressions for $E(Y_{(i)})$, $Var(Y_{(i)})$ and $Cov(Y_{(i)}, Y_{(j)})$ are presented in Section 3.2.2. Assumption ***i*** is also not true for the OLSE; however, as shown in Section 3.2.3, the sensible selection of the *Y*-axis plotting positions, which is not by the Bernard estimator used in OLSE, can satisfy this assumption in the case of complete data. Finally, under the assumption that the error can be confined to *Y*, assumption ***iv*** is satisfied.

The analytical examination clearly shows that the OLSE of β and $\ln\alpha$ are not BLUE. It is very likely that the OLSE of α is not BLUE as well.

3.2.2 Derivations of the Mean, Variance and Covariance of the Order Statistics of *Y*

As is well known, if the random variable *T* follows the Weibull distribution with scale parameter α and shape parameter β, then the variable $X = \ln T$ follows the extreme value distribution whose CDF has a location-scale form given by

$$F(x) = 1 - \exp[-\exp(x - \mu)/\sigma] \tag{3-1}$$

where $\mu = \ln\alpha$ and $\sigma = 1/\beta$.

For location-scale distributions such as the normal distribution and the extreme value distribution, the variable $Z = (X - \mu)/\sigma$ follows a parameter-free distribution and hence is frequently used to simply the analytical analysis. This variable Z is frequently called reduced variable.

From Equation (3-1), the reduced variable Z related to the Weibull distribution is given by

$$Z = (X-\mu)/\sigma = (\ln T - \ln\alpha)/(1/\beta) = \beta\ln T - \beta\ln\alpha = \ln\left[(T/\alpha)^{\beta}\right] \quad (3\text{-}2)$$

Based on the Weibull CDF, the CDF of Z can be determined as

$$\begin{aligned} F(z) &= P(Z \le z) = P(\ln\left[(T/\alpha)^{\beta}\right] \le z) = P(T \le \alpha\cdot e^{z/\beta}) \\ &= 1-\exp\left[-\left(\frac{\alpha\cdot e^{z/\beta}}{\alpha}\right)^{\beta}\right] = 1-\exp(-e^{z}) \end{aligned} \quad (3\text{-}3)$$

Obviously, Z follows the standard smallest extreme value Type I distribution or the standard Gumbel distribution.

The linearized Weibull CDF is

$$\ln\left[-\ln(1-F(t))\right] = \beta\ln t - \beta\ln\alpha \quad (3\text{-}4)$$

Comparing Equation (3-4) with Equation (3-2), we obtain

$$Z = \ln\left[-\ln(1-F(t))\right] \quad (3\text{-}5)$$

Recall that $Y = \ln\left[-\ln(1-F(t))\right]$ which is exactly the same as the expression for Z in Equation (3-5), therefore, the values of $y_i = \ln\left[-\ln(1-\hat{F}_{(i)})\right]$ can be looked on as the values taken on by the order statistic of Z, i.e., $Z_{(i)}$. Thus,

$$E(Y_{(i)}) = E(Z_{(i)}), Var(Y_{(i)}) = Var(Z_{(i)}), Cov(Y_{(i)}, Y_{(j)}) = Cov(Z_{(i)}, Z_{(j)}) \quad (3\text{-}6)$$

From Equation (3-3), the CDF of $Z_{(i)}$ can be determined as

$$\begin{aligned} F(z_i) &= i\binom{n}{i}\int_{-\infty}^{z_i}\left[F(z)\right]^{i-1}\left[1-F(z)\right]^{n-i} f(z)dz \\ &= i\binom{n}{i}\int_{-\infty}^{z_i}(1-\exp(-e^z))^{i-1}(\exp(-e^z))^{n-i}\, d(1-\exp(-e^z)) \end{aligned} \qquad (3\text{-}7)$$

$E(Z_{(i)}), Var(Z_{(i)})$ and $Cov(Z_{(i)}, Z_{(j)})$ can then be derived from the CDF. The results are

$$E(Z_{(i)}) = i\binom{n}{i}\cdot\sum_{k=0}^{i-1}\left\{\binom{i-1}{k}(-1)^k\,\frac{-\gamma-\ln(n-i+k+1)}{n-i+k+1}\right\} \qquad (3\text{-}8)$$

$$\begin{aligned} E(Z_{(i)}^2) = 1.978112 + i\binom{n}{i}\sum_{k=0}^{i-1}\left\{\binom{i-1}{k}(-1)^k \cdot \right. \\ \left. \frac{2\gamma\ln(n-i+k+1)+\left[\ln(n-i+k+1)\right]^2}{n-i+k+1}\right\} \end{aligned} \qquad (3\text{-}9)$$

where $\gamma = 0.577216$ is the Euler's constant.

$$E(Z_{(i)}Z_{(j)}) = \binom{n}{j}\binom{j}{i-1}\int_0^{+\infty}\int_0^{v}\ln u\ln v[1-e^{-u}]^{i-1}[e^{-u}-e^{-v}]^{j-i-1}e^{-u}e^{-(n-j+1)v}dudv \qquad (3\text{-}10)$$

where $u = \ln z_i$ and $v = \ln z_j$.

$$Var(Z_{(i)}) = E(Z_{(i)}^2) - [E(Z_{(i)})]^2 \qquad (3\text{-}11)$$

$$Cov(Z_{(i)}, Z_{(j)}) = E(Z_{(i)}Z_{(j)}) - E(Z_{(i)})\cdot E(Z_{(j)}) \qquad (3\text{-}12)$$

Appendix A gives the detailed derivation of Equations (3-8) – (3-10).

From the above results, obviously, the variance of $Z_{(i)}$ or equally $Y_{(i)}$ is not constant. In Chapter 6, the values of $Var(Y_{(i)})$ at selected sample sizes will be tabulated which are used to calculate the exact weights for the WLSE method.

3.2.3 Sensible Selection for y_i

Several numerical expressions for calculating the values of $F_{(i)}$ or y_i have been presented in Section 2.2 with their theoretical backgrounds. It is noteworthy that $F_{(i)}$ is treated as a random variable rather than a probability in the process of determining the analytical expressions of its estimators. Let the i^{th} smallest observation $T_{(i)}$ or $X_{(i)} = \ln T_{(i)}$, which is also a random variable having a different value in different samples, has the plotting position $F_{(i)}$ or y_i. It is sensible to select y_i so that the point $((E(X_{(i)}), y_i)$ lies on the linear regression line. Numerically this means

$$y_i = A + B \cdot E(X_{(i)}) \tag{3-13}$$

where $A = -\beta \ln \alpha$, $B = \beta$.

Let $\mu = \ln \alpha$ and $\sigma = 1/\beta$, then $A = -\mu/\sigma$ and $B = 1/\sigma$, and Equation (3-13) becomes

$$y_i = \lfloor E(X_{(i)}) - \mu \rfloor / \sigma = E\lfloor (X_{(i)} - \mu)/\sigma \rfloor \tag{3-14}$$

Thus the plotting positions y_i are uniquely defined as the expected values of the order statistics of the reduced variable $Z = (X - \mu)/\sigma$, i.e.,

$$y_i = E(Z_{(i)}) \tag{3-15}$$

The reduced variable Z and its order statistics are defined in the previous section and the values of $E(Z_{(i)})$ can be obtained by Equation (3-8).

Based on the relationship $y_i = \ln\left[-\ln\left(1-\hat{F}_{(i)}\right)\right]$, the plotting positions of $F_{(i)}$ can be obtained by

$$\hat{F}_{(i)} = 1 - \exp(-\exp(E(Z_{(i)})) \qquad (3\text{-}16)$$

This way of determining the Y-axis plotting positions makes the points (x_i, y_i), where the values of x_i come from sample observations and are different in different samples, and the values of y_i are determined by Equation (3-15) and are fixed for a certain sample size, on the average will achieve a linear plot if the Weibull distribution fits.

3.2.4 Relationship between Plotting Positions and Bias of LS Estimators

Now suppose the plotting positions $\hat{F}_{(i)}$ or $y_i = \ln\left[-\ln\left(1-\hat{F}_{(i)}\right)\right]$ are predetermined by some convention, e.g., the Bernard estimator $\hat{F}_{(i)} = (i-0.3)/(n+0.4)$, the Hazen estimator $\hat{F}_{(i)} = (i-0.5)/n$, the Weibull estimator $\hat{F}_{(i)} = i/(n+1)$, or the expected values of the order statistics of the reduced variable, i.e., Equation (3-15) or Equation (3-16). Thus the Y-axis values are fixed at a specific sample size and have no uncertainty (this assumption is different from the one presented in Section 3.2.1 for the OLSE method), suggesting that one should minimize the sum of squares of the deviations in the X-axis direction (i.e., failure time) when applying the LSE method for parameter estimation.

With $\mu = \ln\alpha$ and $\sigma = 1/\beta$, the linear regression model $Y = A + BX + e$ can be transferred to

$$X = \mu + \sigma Y + e' \tag{3-17}$$

where e' is the error term.

The LS estimating equations for μ and σ, by minimizing the sum of $\left[x_i - (\hat{\mu} + \hat{\sigma} y_i)\right]^2$, are given by

$$\hat{\sigma} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})x_i}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3-18}$$

$$\hat{\mu} = \bar{x} - \hat{\sigma} \cdot \bar{y} \tag{3-19}$$

where x_i denote the sample values taken on from the random variable $X_{(i)}$ and the values of y_i are predetermined and not random. $\bar{x} = \sum_{i=1}^{n} x_i \Big/ n$ and $\bar{y} = \sum_{i=1}^{n} y_i \Big/ n$.

It can be easily proved that $\hat{\sigma}$ and $\hat{\mu}$ computed by Equation (3-18) and Equation (3-19), respectively, are unbiased if the values of y_i are determined by the method presented in Section 3.2.3, i.e., Equation (3-15). The proof is given below.

Proof for the Unbiasedness of the LS Estimators of μ and σ When Equation (3-15) is Used for y_i

The unbiasedness of $\hat{\sigma}$ and $\hat{\mu}$ can be proved by $E(\hat{\sigma}) = \sigma$ and $E(\hat{\mu}) = \mu$.

Since y_i are treated as fixed values, from Equation (3-18), the expected value of $\hat{\sigma}$ is given by

$$E(\hat{\sigma}) = E\left(\frac{\sum_{i=1}^{n}(y_i - \bar{y})x_i}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right) = \frac{\sum_{i=1}^{n}(y_i - \bar{y})E(X_{(i)})}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3-20}$$

In Section 3.2.3, Equation (3-15) is derived based on the relationship given by Equation (3-13). Rewrite this equation as

$$E(X_{(i)}) = \mu + \sigma y_i \tag{3-21}$$

and substituting it in Equation (3-20) yields

$$\begin{aligned}
E(\hat{\sigma}) &= \frac{\sum_{i=1}^{n}(y_i - \bar{y})E(X_{(i)})}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\mu + \sigma y_i)}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \\
&= \mu \frac{\sum_{i=1}^{n}(y_i - \bar{y})}{\sum_{i=1}^{n}(y_i - \bar{y})^2} + \sigma \frac{\sum_{i=1}^{n}(y_i - \bar{y}) \cdot y_i}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \\
&= \mu \frac{\sum_{i=1}^{n} y_i - n\bar{y}}{\sum_{i=1}^{n}(y_i - \bar{y})^2} + \sigma \frac{\sum_{i=1}^{n} y_i^2 - \bar{y} \cdot \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} y_i^2 - 2\bar{y} \cdot \sum_{i=1}^{n} y_i + n\bar{y}^2} \\
&= \mu \frac{n\bar{y} - n\bar{y}}{\sum_{i=1}^{n}(y_i - \bar{y})^2} + \sigma \frac{\sum_{i=1}^{n} y_i^2 - n\bar{y}^2}{\sum_{i=1}^{n} y_i^2 - n\bar{y}^2} \\
&= \sigma
\end{aligned} \tag{3-22}$$

Then from Equation (3-19), the expected value of $\hat{\mu}$ is

$$E(\hat{\mu}) = E(\overline{X}) - E(\hat{\sigma}) \cdot \bar{y} = \mu + \sigma\bar{y} - \sigma\bar{y} = \mu \tag{3-23}$$

Therefore, the LS estimators $\hat{\sigma}$ and $\hat{\mu}$ are both unbiased when Equation (3-15) is used for determining y_i. This kind of plotting position has a good statistical background and has been recommended by a few researchers, see, e.g., (Ross, 1994b). However, since the relationships between μ and α, and σ and β, are both nonlinear, there is no guarantee that $\hat{\alpha}$ and $\hat{\beta}$ are also unbiased.

In the following, assuming the values of y_i are predetermined by any plotting convention, the analytical expressions of the relative bias of $\hat{\sigma}$ and $\hat{\mu}$, respectively, are presented.

Let
$$s_i = \frac{(y_i - \bar{y})}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3-24}$$

From Equation (3-18) and Equation (3-19), the relative bias of the LS estimators of μ and σ can be obtained by

$$\begin{aligned}\frac{E(\hat{\sigma})}{\sigma} &= E\left(\frac{\sum_{i=1}^{n}(y_i - \bar{y})\cdot X_{(i)}}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right)\Bigg/\sigma \\ &= E\left(\sum_{i=1}^{n} s_i X_{(i)}\right)\Bigg/\sigma = \sum_{i=1}^{n} s_i E(X_{(i)})\Big/\sigma\end{aligned} \tag{3-25}$$

and

$$\frac{E(\hat{\mu})}{\mu} = \frac{E(\bar{X} - \hat{\sigma}\cdot\bar{y})}{\mu} = \frac{E(\bar{X}) - E(\hat{\sigma})\cdot\bar{y}}{\mu} \tag{3-26}$$

From the definitions of the reduced variable Z and its order statistic $Z_{(i)}$,

$$E(Z_{(i)}) = \left[E(X_{(i)}) - \mu\right]\big/\sigma \tag{3-27}$$

and

$$E(\overline{Z}) = \left[E(\overline{X}) - \mu\right]/\sigma \tag{3-28}$$

Let $\omega_i = E(Z_{(i)})$ and rewrite Equation (3-27) as

$$E(X_{(i)}) = \mu + \sigma \cdot E(Z_{(i)}) = \mu + \sigma \cdot \omega_i \tag{3-29}$$

where the values of ω_i can be calculated by Equation (3-8).

Substituting Equation (3-29) in Equation (3-25) for $E(X_{(i)})$ yields

$$\begin{aligned} E(\hat{\sigma})/\sigma &= \sum_{i=1}^{n} s_i E(X_{(i)}) \Big/ \sigma = \sum_{i=1}^{n} s_i \left(\mu + \sigma\omega_i\right) \Big/ \sigma \\ &= \mu \sum_{i=1}^{n} s_i \Big/ \sigma + \sum_{i=1}^{n} s_i \omega_i \\ &= \sum_{i=1}^{n} s_i \omega_i \end{aligned} \tag{3-30}$$

where $\sum_{i=1}^{n} s_i = 0$ can be easily obtained from Equation (3-24).

Since Z follows the standard smallest extreme value Type I distribution, we have $E(\overline{Z}) = E(Z) = -\gamma$, where $\gamma = 0.577216$ is the Euler's constant. Rewrite Equation (3-28) as

$$E(\overline{X}) = \mu + \sigma \cdot E(\overline{Z}) = \mu - \sigma \cdot \gamma \tag{3-31}$$

Then substituting Equation (3-30) and Equation (3-31) in Equation (3-26) for $E(\hat{\sigma})$ and $E(\overline{X})$ yields

$$\begin{aligned} E(\hat{\mu})/\mu &= \left[E(\overline{X}) - E(\hat{\sigma}) \cdot \overline{y}\right]/\mu \\ &= \left[\mu - \sigma\gamma - \sigma\overline{y} \sum_{i=1}^{n} s_i \omega_i \right] \Big/ \mu \end{aligned} \tag{3-32}$$

Thus the relative bias of $\hat{\sigma}$ can be numerically calculated by Equation (3-30) given a sample size, a predetermined method for calculating the plotting positions, and the values of ω_i which can be calculated by Equation (3-8). The relative bias of $\hat{\mu}$ involves the true values of μ and σ which are normally unknown and hence can only be estimated.

3.2.5 Pivotal Functions of LS Estimators

The definition of pivotal function is a function, e.g., $g(\theta)$ of θ whose distribution is known and is independent of θ (Garthwaite et al., 2002, p. 98). For the Weibull distribution, as is well known, there are two pivotal functions for the ML estimators of the Weibull parameters, i.e., $\hat{\beta}/\beta$ and $\hat{\beta}\ln(\hat{\alpha}/\alpha)$. Their distributions can be determined via the Monte Carlo method and are independent of α and β. Bain & Antle (1967) presented three theorems of $\hat{\beta}/\beta$ and $\hat{\beta}\ln(\hat{\alpha}/\alpha)$ that clearly address the properties of the two pivotal functions for their proposed estimators of the Weibull parameters (neither LSE nor MLE). Let $\hat{\alpha}_{1,1}$, $\hat{\beta}_{1,1}$ denote the estimators of α and β when the sample is actually from a normalized Weibull distribution, i.e., $\alpha = \beta = 1$, the three theorems are as follows.

Theorem 1. $\hat{\beta}/\beta$ has the same distribution as $\hat{\beta}_{1,1}$ and is distributed independently of α and β.

Theorem 2. $\hat{\alpha}/\alpha$ has the same distribution as $\hat{\alpha}_{1,\beta}$ and depends only on β.

Theorem 3. $\hat{\beta}\ln(\hat{\alpha}/\alpha)$ has the same distribution as $\hat{\beta}_{1,1}\ln(\hat{\alpha}_{1,1})$, or $(\hat{\alpha}/\alpha)^{\beta}$ has the same distribution $\hat{\alpha}_{1,1}$.

Thoman et al. (1969) examined the above theorems for the MLE of the Weibull parameters and pointed out that $\hat{\beta}/\beta$ and $\hat{\beta}\ln(\hat{\alpha}/\alpha)$ are two pivotal functions.

The pivotal functions for the LS estimators are seldom mentioned by Weibull researchers. It can be proved that the properties of $\hat{\beta}/\beta$ and $\hat{\beta}\ln(\hat{\alpha}/\alpha)$ also apply to the LS estimated α and β. The proof is given below.

<u>Proof for the Two Pivotal Functions of the LSE</u>

Let $T_1, T_2, \cdots, T_i, \cdots, T_n$ $(i = 1, 2, \cdots, n)$ denotes a random sample from a normalized Weibull distribution (i.e., $\alpha = \beta = 1$). Substituting $t_i = (T_i/\alpha)^{\beta}$ can generate a new random sample, denoted by $t_1, t_2, \cdots, t_i, \cdots, t_n$ $(i = 1, 2, \cdots, n)$, from the Weibull distribution with arbitrary α and β. Applying the LSE method for this new sample, the LS shape parameter estimator can be obtained by

$$\hat{\beta} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{n\sum_{i=1}^{n} \ln t_i \cdot y_i - \sum_{i=1}^{n} \ln t_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} (\ln t_i)^2 - \left(\sum_{i=1}^{n} \ln t_i\right)^2} \tag{3-33}$$

Substituting $\alpha T_i^{1/\beta}$ for t_i in Equation (3-33) yields

$$\begin{aligned}\hat{\beta} &= \frac{n\sum_{i=1}^{n} \ln\left(\alpha T_i^{1/\beta}\right) \cdot y_i - \sum_{i=1}^{n} \ln\left(\alpha T_i^{1/\beta}\right) \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} \left[\ln\left(\alpha T_i^{1/\beta}\right)\right]^2 - \left[\sum_{i=1}^{n} \ln\left(\alpha T_i^{1/\beta}\right)\right]^2} = \frac{\left(n\sum_{i=1}^{n} \ln T_i \cdot y_i - \sum_{i=1}^{n} \ln T_i \sum_{i=1}^{n} y_i\right) \Big/ \beta}{\left(n\beta\sum_{i=1}^{n} (\ln T_i)^2 - \beta\left(\sum_{i=1}^{n} \ln T_i\right)^2\right) \Big/ \beta^2} \\ &= \beta \cdot \frac{n\sum_{i=1}^{n} \ln T_i \cdot y_i - \sum_{i=1}^{n} \ln T_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} (\ln T_i)^2 - \left(\sum_{i=1}^{n} \ln T_i\right)^2}\end{aligned} \tag{3-34}$$

For the normalized Weibull sample $T_1, T_2, \cdots, T_i, \cdots, T_n$ $(i = 1, 2, \cdots, n)$, the LS shape parameter estimator is given by

$$\hat{\beta}_{1,1} = \frac{n\sum_{i=1}^{n} \ln T_i \cdot y_i - \sum_{i=1}^{n} \ln T_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} (\ln T_i)^2 - \left(\sum_{i=1}^{n} \ln T_i\right)^2} \tag{3-35}$$

Comparing Equation (3-34) and Equation (3-35), we obtain

$$\hat{\beta} = \beta \cdot \hat{\beta}_{1,1} \tag{3-36}$$

It follows that $\hat{\beta}/\beta$ has the same distribution as $\hat{\beta}_{1,1}$.

Similarly, the second pivotal function $\hat{\beta}\ln(\hat{\alpha}/\alpha)$ can be proved. For the Weibull sample $t_1, t_2, \cdots, t_i, \cdots, t_n$ $(i = 1, 2, \cdots, n)$, the LS estimators satisfy

$$\hat{\beta}\ln\hat{\alpha} = -\frac{\sum_{i=1}^{n} y_i - \hat{\beta}\sum_{i=1}^{n} x_i}{n} = -\frac{\sum_{i=1}^{n} y_i - \hat{\beta}\sum_{i=1}^{n} \ln t_i}{n} \tag{3-37}$$

Substituting $\alpha T_i^{1/\beta}$ for t_i in Equation (3-37) yields

$$\begin{aligned} \hat{\beta}\ln\hat{\alpha} &= -\frac{\sum_{i=1}^{n} y_i - \hat{\beta}\sum_{i=1}^{n} \ln\left(\alpha T_i^{1/\beta}\right)}{n} = -\frac{\sum_{i=1}^{n} y_i - (\hat{\beta}/\beta)\sum_{i=1}^{n} \ln T_i - n\hat{\beta}\ln\alpha}{n} \\ &= -\frac{\sum_{i=1}^{n} y_i - (\hat{\beta}/\beta)\sum_{i=1}^{n} \ln T_i}{n} + \hat{\beta}\ln\alpha \end{aligned} \tag{3-38}$$

Thus

$$\hat{\beta}\ln\left(\hat{\alpha}/\alpha\right) = \hat{\beta}\ln\hat{\alpha} - \hat{\beta}\ln\alpha = -\frac{\sum_{i=1}^{n} y_i - (\hat{\beta}/\beta)\sum_{i=1}^{n} \ln T_i}{n} \tag{3-39}$$

For the normalized Weibull sample $T_1, T_2, \cdots, T_i, \cdots, T_n$ $(i = 1, 2, \cdots, n)$, the LS estimators satisfy

$$\hat{\beta}_{1,1} \ln \hat{\alpha}_{1,1} = -\frac{\sum_{i=1}^{n} y_i - \hat{\beta}_{1,1} \sum_{i=1}^{n} x_i}{n} = -\frac{\sum_{i=1}^{n} y_i - \hat{\beta}_{1,1} \sum_{i=1}^{n} \ln T_i}{n} \qquad (3\text{-}40)$$

From Equation (3-36), $\hat{\beta}/\beta = \hat{\beta}_{1,1}$. Therefore, comparing Equation (3-39) and Equation (3-40) yields

$$\hat{\beta} \ln(\hat{\alpha}/\alpha) = \hat{\beta}_{1,1} \ln \hat{\alpha}_{1,1} \qquad (3\text{-}41)$$

It follows that $\hat{\beta} \ln(\hat{\alpha} / \alpha)$ has the same distribution as $\hat{\beta}_{1,1} \ln \hat{\alpha}_{1,1}$.

The above proof applies to complete data; however, in the case of censored data, one can simply change n in the equations to r, and the results still holds. In addition, the values of y_i are treated as fixed values in the proof, and it does not matter which method is used for calculating y_i.

The two pivotal functions, $\hat{\beta} / \beta$ and $\hat{\beta} \ln(\hat{\alpha} / \alpha)$, are very useful in parameter estimation. An important application is to correct the bias of the Weibull estimators. Investigations on the bias correction methods for the LS estimators based on the first pivotal function are shown in Chapter 5. Moreover, the pivotal functions also play a significant role in the Monte Carlo experiment examination for the LS estimators. Especially for the examination of the shape parameter, in most times the true parameter values of α and β can be fixed to 1 in the experiment, since $\hat{\beta} / \beta$ has the same distribution as $\hat{\beta}_{1,1}$. Theoretically it is unnecessary to try different parameter values and hence a lot of simulation work can be saved. This, unfortunately, has not been noticed by many researchers. The effects of the pivotal functions on the setting of the true parameter values of α and β in a Monte Carlo simulation experiment are presented in Section 3.3.2.

3.3 Monte Carlo Experiment Examination of the OLS Estimators

As previously stated in Section 2.5, bias, variance (or standard deviation) and MSE of the estimators are the common criteria for assessing the performance of a parameter estimation method. The analytical examinations described in last section show that the OLS estimators of α and β are not BLUE: they are biased and may have large variance. However, it is difficult to give analytical expressions for the bias or variance of the OLS estimated α and β. For this reason, the Monte Carlo method is frequently used. With Monte Carlo simulations, the sampling distributions of the estimators can be approximated and hence the bias, variance and MSE of the estimators can be determined.

Ambrozic & Vidovic (2007) summarized the three typical aims of Monte Carlo simulations in reliability data analysis as: comparing different parameter estimation methods, discovering the optimal probability estimators (i.e., $\hat{F}_{(i)}$ or $\hat{F}_{f,(j)}$) in the linear regression method, and analyzing the type of distribution functions for Weibull estimators. All of these purposes are covered in this study.

In the following, Section 3.3.1 describes the common procedures of a Monte Carlo simulation experiment to obtain the bias, variance and MSE of the OLS estimators in the case of complete data and multiply censored data, respectively.

Section 3.3.2 presents the settings of the experiment factors and Section 3.3.3 presents the important simulation results for the OLS estimators.

3.3.1 Monte Carlo Experiment Procedures

Monte Carlo simulations can be executed by many statistical software packages such as MATLAB, SAS, S-PLUS, Mathematic, etc. Most of these software packages have reliable algorithms for generating the uniformly distributed random numbers. Based on these uniformly distributed numbers, random Weibull samples, either complete or censored, can be generated. The software MATLAB 7 is used in this study.

Monte Carlo Experiment Procedure for Complete Data

The objective of the experiment is to calculate the bias, variance and MSE of the LSE of α and β under different combinations of the predetermined factors including the true parameter values of α and β (denoted by α_T and β_T), and sample size n. The step-by-step experiment procedure is described as follows.

Step 1: Generate n random numbers $p_1, p_2, \cdots, p_n$ from a uniform distribution, $p_i \in U(0,1)$.

Step 2: For any specified values of α_T, β_T and p_i, a random Weibull sample $t_1, t_2, \cdots, t_n$ can be obtained by calculating $t_i = \alpha_T\left[-\ln(1-p_i)\right]^{1/\beta_T}$ $(i = 1, 2, \cdots, n)$.

Step 3: For the current Weibull sample, estimate α and β using the LSE method (refer to Section 2.3).

Step 4: Repeat Step 1 to Step 3 for M times (M is called iteration number or repetition number).

Step 5: Calculate the bias, variance and MSE of the estimators with the following formula,

$$\begin{aligned} B(\hat{\theta}) &= \overline{\theta} - \theta \\ Var(\hat{\theta}) &= \frac{1}{M}\sum_{i=1}^{M}(\hat{\theta} - \overline{\theta})^2 \\ MSE(\hat{\theta}) &= \frac{1}{M}\sum_{i=1}^{M}(\hat{\theta} - \theta)^2 \end{aligned} \tag{3-42}$$

where θ can be replaced by α and β, and $\overline{\theta} = \frac{1}{M}\sum_{i=1}^{M}\hat{\theta}_i$.

Monte Carlo Experiment Procedure for Multiply Censored Data

The experiment procedure for multiply censored data is more complicated because it involves generating multiply censored samples (failure times and censoring times are intermixed in such a sample). The step-by-step experiment procedure for multiply censored data used in this study is described as follows.

Step 1: Generate a random complete sample $t_1, t_2, \cdots, t_n$ from the Weibull distribution with specified values of α_T, β_T and n (refer to the first two steps of the procedure for complete data).

Step 2: From the complete sample $t_1, t_2, \cdots, t_n$, randomly select $n-r$ observations, denoted by $t_{c,k}$ $(k = 1, 2, \cdots, n-r)$, as the candidates to be modified to generate censoring times. The remaining observations, denoted by $t_{f,j}$ $(j = 1, 2, \cdots, r)$, are unchanged as failure times.

Step 3: Generate $n-r$ random numbers $p_1, p_2, \cdots, p_{n-r}$ from a uniform distribution, $p_k \in U(0,1)$.

Step 4: Change $t_{c,k}$ to $p_k \cdot t_{c,k}$ to create the censoring times.

Step 5: Merge failure times and censoring times to produce a multiply censored sample.

Step 6: For the current sample, estimate α and β using the LSE method (refer to Section 2.3).

Step 7: Repeat Step 1 to Step 6 for M times.

Step 8: Calculate the bias, variance and MSE of the estimators by Equation (3-42).

The steps for generating the censoring times (Step 3 and Step 4) are based on an underlying assumption that the censoring times are independent of the failure times. This means the mechanism that causes the censoring is independent of the mechanism that causes the failure. Thus the simplest way can be used to generate the censoring times.

3.3.2 Setting of Experiment Factors

Simulation results are often presented under different combinations of the experiment factors. For complete data, there are four factors of concern: the true parameter values α_T and β_T, sample size n and iteration number M. For censored data, there is one more factor, i.e., the censoring level c. The setting of the experiment factors is by no means arbitrary. In the following, some general guidelines on the selection of the values for each experiment factor are summarized.

Selection of True Values of α and β

For the Weibull distribution, α is the scale parameter that can take on any positive value and β is the shape parameter that usually takes values between 0.1 and 10.

Without considering the existence of the two pivotal functions, i.e., $\hat{\beta}/\beta$ and $\hat{\beta}\ln(\hat{\alpha}/\alpha)$, of the LS estimators, one must examine every combination of common values of α_T and β_T to have a full picture of the performance of the LS estimators. Luckily, as mentioned in Section 3.2.5, the pivotal functions can theoretically save the simulation work. *Theorem 1* says that $\hat{\beta}/\beta$ has the same distribution as $\hat{\beta}_{1,1}$, thus the properties such as bias and MSE of the LS estimated β under any combination of α_T and β_T can be obtained from the properties of $\hat{\beta}_{1,1}$ ($\alpha_T = \beta_T = 1$). Therefore, to examine the LS shape parameter estimator, the values of α_T and β_T can be fixed to 1 in the whole simulation experiment. On the other side, for examining the LS scale parameter estimator, based on *Theorem 2*, i.e., $\hat{\alpha}/\alpha$ has the same distribution as $\hat{\alpha}_{1,\beta}$ and depends only on β, the values of α_T can still be fixed to 1; however, different values of β_T should be used.

In summary, the value of α_T can always be fixed to 1. For the purpose of examining $\hat{\beta}$, β_T can be fixed to 1 as well. However, for examining $\hat{\alpha}$, different β_T should be used. For example, $\beta_T = 0.5, 0.8, 1, 2, 4, 5, 6, 8, 10$.

Selection of Sample Size

Small sample properties and large sample properties of the estimators are frequently examined separately. The selection of sample size depends on the focus of the study. In this thesis, the focus is the small sample properties of the LSE, which is also the recent focus of Weibull researchers. For a Weibull sample of size n, commonly it is known as a small sample if $n \le 20$, a medium sample if $20 < n \le 100$, and a large sample if $n > 100$ (Abernethy, 2000). With the focus on small to medium sized samples, n is frequently set in the range of 3 to 30.

For censored samples, however, the selection of the sample size is more arbitrary. The common range of the sample size used in this study for censored samples is from 10 to 200.

Selection of Iteration Number

The accuracy of the simulation results is closely related to the iteration number or repetition number. Usually increasing the iteration number can achieve a higher accuracy; however, the simulation time is also increased. A trade-off between accuracy and simulation time should be made. The accuracy of the simulation results at an iteration number can be simply estimated by repeating the whole simulation process for several times. Therefore, by setting a tolerance of accuracy, the required iteration number can be determined by trial and error. In the literature, 10000 is the commonly used iteration number, and we found that in most cases, this number can achieve an accuracy of at least two decimal places. To have a higher accuracy, 50000 repetitions can be used.

Selection of Censoring Level

Censoring level is often presented by percentage. Commonly $c \geq 50\%$ refers to a highly censored sample. Both low censoring levels and high censoring levels are examined in this study and the range is frequently from 10% to 80%. For simplicity, the censoring levels selected should satisfy that $c \cdot n$ is an integer and $c \cdot n \geq 2$ (required by the LSE method as at least two data points are needed for conducting regression).

3.3.3 Simulation Results for the OLS Estimators

A Monte Carlo experiment was conducted to examine the bias, standard deviation and MSE of the OLS estimators, especially for the shape parameter estimator, in the cases of complete data and multiply censored data, respectively. The experiment follows the procedures described in Section 3.3.1.

Table 3-1 shows the setting of simulation factors in this experiment. For each combination of the simulation factors (α_T, β_T, n and c), 50000 random samples were generated and the parameter estimates of both parameters were obtained from OLSE and MLE simultaneously. The mean, standard deviation and MSE of both parameter estimates were calculated and analyzed. The experiment was executed in MATLAB 7. The iteration number 50000 in most cases can guarantee an accuracy of 0.5%.

Table 3-1: Setting of experiment factors. The experiment is to examine the OLS estimators.

Factors	Values
α_T	1
β_T	0.5, 1, 2, 3, 5, 8
n	3 – 20, 22, 24,…, 28, 30, 35, …, 45, 50, 60, …, 90, 100, 200 (complete data) 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200 (censored data)
c	10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%
M	50000
Methods	OLSE, MLE

3.3.3.1 Simulation Results for Complete Data

The simulation results of the shape parameter estimator, in the case of complete data, are shown in Table 3-2. The relative values $E(\hat{\beta})/\beta_T$, $S(\hat{\beta})/\beta_T$ and $MSE(\hat{\beta})/\beta_T^2$ are tabulated in the table. The results for the scale parameter are shown in Table 3-3. Since α_T is fixed to 1 all the time, the values of $E(\hat{\alpha}), S(\hat{\alpha})$ and $MSE(\hat{\alpha})$ equal to the relative values. Note that not all the simulation results are tabulated in the two tables; however, the omitted results will not affect the following conclusions which can be observed from the tabulated values.

<u>Simulation Results for Estimators of β (Table 3-2)</u>

1) ***The reliability of the simulation results is judged by the pivotal quantity*** $\hat{\beta}/\beta_T$ *:* In theory, the distribution of $\hat{\beta}/\beta_T$, obtained by both MLE and OLSE, should be independent of β_T . This can be used to check the reliability of the simulations. From Table 3-2, it can be seen that the values of $E(\hat{\beta})/\beta_T$, $S(\hat{\beta})/\beta_T$ and $MSE(\hat{\beta})/\beta_T^2$ for both methods almost do not vary with β_T at all sample sizes examined, especially from $n = 8$ onwards.

2) ***Bias of the OLSE of the shape parameter****:* $\hat{\beta}/\beta_T$ is inconsistent with n . The bias is most significant at $n = 3, 4$; however, it reaches smallest between $n = 6$ and $n = 7$. From $n = 5$ onwards, the relative bias is typically within 5%. During $10 < n < 30$, the relative bias is like a constant and remains at 4% or so. Typically, β is overestimated when $n \le 6$ and underestimated for the remaining conditions.

3) ***Standard deviation and MSE of the OLSE of the shape parameter****:* The magnitude of $S(\hat{\beta})/\beta_T$ and $MSE(\hat{\beta})/\beta_T^2$ decreases as the sample size n increases. The magnitude of the absolute standard deviation ($S(\hat{\beta})$) is much larger than that of the absolute bias ($E(\hat{\beta}) - \beta_T$) under all combinations of the experiment factors, especially when n is very small. In other words, the

MSE of $\hat{\beta}$ is mainly contributed by the standard deviation instead of the bias.

4) ***Comparison between OLSE and MLE**:* The relative bias $\hat{\beta}/\beta_T$ of the OLSE is significantly smaller than that of the MLE for small samples, i.e., $n \le 20$, is slightly smaller than that of the MLE for $20 < n < 50$, and is slightly larger for $n \ge 50$. The magnitude of $S(\hat{\beta})/\beta_T$ and $MSE(\hat{\beta})/\beta_T^2$ of the OLSE is significantly smaller than that of the MLE for $n \le 10$. The differences are small for $10 < n < 20$, and from $n = 20$ onwards, MLE has slightly smaller values than OLSE.

Simulation Results for Estimators of α (Table 3-3)

1) ***General observations**:* The magnitude of the bias, standard deviation and MSE of $\hat{\alpha}$ decreases dramatically as β_T increases and decreases slowly as n increases.

2) ***Bias of the OLSE of the scale parameter**:* At $\beta_T = 0.5$, for the relative bias of $\hat{\alpha}$ to be within 10% requires $n > 35$, and to be within 5% requires $n > 90$. At $\beta_T = 1$, for the relative bias of $\hat{\alpha}$ to be within 10% only requires $n > 7$, and to be within 5% requires $n > 19$. At $\beta_T = 2$, for the relative bias to be within 3% requires $n > 4$, to be within 2% requires $n > 14$, and to be within 1% requires $n > 50$. At $\beta_T = 3$ onwards, the relative bias of $\hat{\alpha}$ is always smaller than 2% and typically within 1%.

3) ***Standard deviation and MSE of the OLSE of the scale parameter**:* The magnitude of $S(\hat{\alpha})$ and $MSE(\hat{\alpha})$ decreases as either β_T or n increases. The

largest values of them happen at $n=3$ and $\beta_T=0.5$. At $n \geq 100$ or $\beta_T \geq 3$, their values are very small and close to 0. Same as the results for the MSE of $\hat{\beta}$, $MSE(\hat{\alpha})$ is mainly contributed by the standard deviation instead of bias.

4) ***Comparison between OLSE and MLE:*** MLE is significantly better than OLSE for estimating α when $\beta_T \leq 3$ in view of bias, standard deviation and MSE, and especially at small β_T and small n. For $\beta_T > 3$, the two estimators of α are very close.

5) ***Comparison between the results for $\hat{\alpha}$ and $\hat{\beta}$ of the OLSE:*** The bias of $\hat{\alpha}$ depends on β_T but the bias of $\hat{\beta}$ is independent of β_T. The bias of $\hat{\alpha}$ seems to be not an issue when $\beta_T \geq 3$, as the bias is typically within 1% at all sample sizes investigated. However, the bias of $\hat{\beta}$ is 4% to 5% for small to medium sized samples. The standard deviation and MSE of $\hat{\alpha}$ also depend on β_T but those of $\hat{\beta}$ is independent of β_T. The magnitude of $S(\hat{\alpha})$ and $MSE(\hat{\alpha})$ becomes very small when $\beta_T \geq 3$, and are smaller than that of $S(\hat{\beta})/\beta_T$ and $MSE(\hat{\beta})/\beta_T^2$; however, when $\beta_T < 3$, the magnitude of $S(\hat{\alpha})$ and $MSE(\hat{\alpha})$ is typically larger than that of $S(\hat{\beta})/\beta_T$ and $MSE(\hat{\beta})/\beta_T^2$.

Table 3-2: Simulation results of $\hat{\beta}$ for complete data, generated by OLSE and MLE, at different n and β_T: the values of $E(\hat{\beta})/\beta_T \pm S(\hat{\beta})/\beta$ and $MSE(\hat{\beta})/\beta_T^2$ (in parentheses).

$\beta_T=0.5$

Method	3	4	5	6	7	8	9	10	11	12	13
OLSE	1.425 ± 5.042 (25.603)	1.136 ± 0.943 (0.908)	1.045 ± 0.582 (0.341)	1.012 ± 0.480 (0.231)	0.991 ± 0.414 (0.172)	0.982 ± 0.376 (0.142)	0.972 ± 0.344 (0.119)	0.971 ± 0.325 (0.107)	0.964 ± 0.303 (0.093)	0.963 ± 0.286 (0.083)	0.960 ± 0.274 (0.077)
MLE	2.284 ± 8.265 (69.954)	1.658 ± 1.342 (2.234)	1.436 ± 0.782 (0.802)	1.334 ± 0.608 (0.481)	1.267 ± 0.499 (0.321)	1.224 ± 0.434 (0.239)	1.192 ± 0.387 (0.186)	1.173 ± 0.357 (0.157)	1.149 ± 0.323 (0.127)	1.135 ± 0.298 (0.107)	1.124 ± 0.282 (0.095)
Method	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**
OLSE	0.961 ± 0.264 (0.071)	0.961 ± 0.253 (0.066)	0.959 ± 0.244 (0.061)	0.960 ± 0.238 (0.058)	0.960 ± 0.232 (0.056)	0.960 ± 0.225 (0.052)	0.961 ± 0.220 (0.050)	0.959 ± 0.210 (0.046)	0.962 ± 0.201 (0.042)	0.959 ± 0.193 (0.039)	0.962 ± 0.186 (0.036)
MLE	1.114 ± 0.265 (0.083)	1.105 ± 0.252 (0.075)	1.097 ± 0.240 (0.067)	1.092 ± 0.231 (0.062)	1.086 ± 0.222 (0.057)	1.080 ± 0.213 (0.052)	1.078 ± 0.208 (0.049)	1.067 ± 0.193 (0.042)	1.063 ± 0.184 (0.038)	1.056 ± 0.173 (0.033)	1.053 ± 0.166 (0.030)
Method	**30**	**35**	**40**	**45**	**50**	**60**	**70**	**80**	**90**	**100**	**200**
OLSE	0.963 ± 0.181 (0.034)	0.964 ± 0.168 (0.030)	0.965 ± 0.157 (0.026)	0.965 ± 0.149 (0.023)	0.968 ± 0.142 (0.021)	0.971 ± 0.129 (0.018)	0.971 ± 0.121 (0.015)	0.974 ± 0.114 (0.014)	0.976 ± 0.108 (0.012)	0.977 ± 0.102 (0.011)	0.984 ± 0.073 (0.006)
MLE	1.051 ± 0.159 (0.028)	1.042 ± 0.145 (0.023)	1.035 ± 0.132 (0.019)	1.030 ± 0.124 (0.016)	1.028 ± 0.117 (0.014)	1.024 ± 0.106 (0.012)	1.020 ± 0.097 (0.010)	1.017 ± 0.090 (0.008)	1.016 ± 0.085 (0.007)	1.015 ± 0.080 (0.007)	1.007 ± 0.056 (0.003)

$\beta_T=1$

Method	3	4	5	6	7	8	9	10	11	12	13
OLSE	1.428 ± 3.225 (10.586)	1.125 ± 0.925 (0.871)	1.053 ± 0.597 (0.359)	1.009 ± 0.482 (0.233)	0.996 ± 0.423 (0.179)	0.983 ± 0.376 (0.142)	0.974 ± 0.343 (0.119)	0.970 ± 0.322 (0.104)	0.963 ± 0.303 (0.093)	0.961 ± 0.286 (0.083)	0.961 ± 0.274 (0.077)
MLE	2.288 ± 5.830 (35.648)	1.641 ± 1.353 (2.243)	1.445 ± 0.803 (0.844)	1.331 ± 0.619 (0.492)	1.274 ± 0.510 (0.335)	1.226 ± 0.437 (0.242)	1.193 ± 0.386 (0.186)	1.171 ± 0.351 (0.153)	1.150 ± 0.322 (0.126)	1.134 ± 0.299 (0.108)	1.123 ± 0.281 (0.094)
Method	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**
OLSE	0.960 ± 0.262 (0.070)	0.960 ± 0.255 (0.067)	0.962 ± 0.244 (0.061)	0.959 ± 0.238 (0.058)	0.958 ± 0.231 (0.055)	0.960 ± 0.225 (0.052)	0.960 ± 0.219 (0.049)	0.960 ± 0.209 (0.045)	0.959 ± 0.201 (0.042)	0.959 ± 0.191 (0.038)	0.962 ± 0.185 (0.036)
MLE	1.113 ± 0.265 (0.083)	1.105 ± 0.254 (0.075)	1.099 ± 0.240 (0.067)	1.091 ± 0.230 (0.061)	1.084 ± 0.221 (0.056)	1.081 ± 0.213 (0.052)	1.076 ± 0.206 (0.048)	1.069 ± 0.194 (0.042)	1.062 ± 0.182 (0.037)	1.057 ± 0.171 (0.032)	1.053 ± 0.165 (0.030)
Method	**30**	**35**	**40**	**45**	**50**	**60**	**70**	**80**	**90**	**100**	**200**
OLSE	0.961 ± 0.180 (0.034)	0.964 ± 0.168 (0.030)	0.966 ± 0.158 (0.026)	0.967 ± 0.149 (0.023)	0.966 ± 0.141 (0.021)	0.971 ± 0.130 (0.018)	0.973 ± 0.121 (0.015)	0.974 ± 0.114 (0.014)	0.975 ± 0.108 (0.012)	0.977 ± 0.102 (0.011)	0.984 ± 0.073 (0.006)
MLE	1.049 ± 0.158 (0.027)	1.043 ± 0.145 (0.023)	1.037 ± 0.133 (0.019)	1.032 ± 0.124 (0.017)	1.027 ± 0.117 (0.014)	1.024 ± 0.106 (0.012)	1.020 ± 0.097 (0.010)	1.018 ± 0.090 (0.009)	1.015 ± 0.085 (0.007)	1.014 ± 0.080 (0.007)	1.007 ± 0.056 (0.003)

$\beta_T=2$

Method	3	4	5	6	7	8	9	10	11	12	13
OLSE	1.414 ± 2.459 (6.218)	1.132 ± 0.871 (0.776)	1.050 ± 0.605 (0.368)	1.009 ± 0.480 (0.230)	0.993 ± 0.418 (0.174)	0.980 ± 0.375 (0.141)	0.974 ± 0.347 (0.121)	0.967 ± 0.319 (0.103)	0.966 ± 0.303 (0.093)	0.963 ± 0.288 (0.084)	0.961 ± 0.274 (0.077)
MLE	2.258 ± 3.803 (16.047)	1.652 ± 1.268 (2.033)	1.444 ± 0.807 (0.848)	1.329 ± 0.605 (0.475)	1.272 ± 0.504 (0.328)	1.226 ± 0.436 (0.241)	1.194 ± 0.389 (0.189)	1.168 ± 0.348 (0.149)	1.152 ± 0.323 (0.127)	1.135 ± 0.300 (0.108)	1.122 ± 0.281 (0.094)
Method	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**
OLSE	0.961 ± 0.262 (0.070)	0.960 ± 0.254 (0.066)	0.959 ± 0.244 (0.061)	0.960 ± 0.238 (0.058)	0.960 ± 0.232 (0.056)	0.957 ± 0.225 (0.052)	0.960 ± 0.220 (0.050)	0.959 ± 0.208 (0.045)	0.959 ± 0.200 (0.042)	0.960 ± 0.193 (0.039)	0.962 ± 0.186 (0.036)
MLE	1.113 ± 0.264 (0.083)	1.103 ± 0.252 (0.074)	1.096 ± 0.241 (0.068)	1.091 ± 0.230 (0.061)	1.086 ± 0.223 (0.057)	1.079 ± 0.213 (0.051)	1.077 ± 0.207 (0.049)	1.068 ± 0.192 (0.042)	1.061 ± 0.181 (0.036)	1.057 ± 0.172 (0.033)	1.053 ± 0.166 (0.031)
Method	**30**	**35**	**40**	**45**	**50**	**60**	**70**	**80**	**90**	**100**	**200**
OLSE	0.962 ± 0.180 (0.034)	0.963 ± 0.167 (0.029)	0.965 ± 0.157 (0.026)	0.967 ± 0.149 (0.023)	0.969 ± 0.141 (0.021)	0.971 ± 0.130 (0.018)	0.972 ± 0.121 (0.015)	0.974 ± 0.113 (0.013)	0.976 ± 0.107 (0.012)	0.976 ± 0.103 (0.011)	0.984 ± 0.073 (0.006)
MLE	1.049 ± 0.158 (0.028)	1.041 ± 0.143 (0.022)	1.036 ± 0.133 (0.019)	1.032 ± 0.124 (0.016)	1.028 ± 0.117 (0.014)	1.023 ± 0.106 (0.012)	1.020 ± 0.097 (0.010)	1.017 ± 0.090 (0.008)	1.015 ± 0.084 (0.007)	1.014 ± 0.081 (0.007)	1.007 ± 0.056 (0.003)

$\beta_T=3$

Method	3	4	5	6	7	8	9	10	11	12	13
OLSE	1.419 ± 2.395 (5.914)	1.137 ± 0.884 (0.799)	1.047 ± 0.592 (0.352)	1.014 ± 0.486 (0.237)	0.991 ± 0.423 (0.179)	0.984 ± 0.378 (0.143)	0.973 ± 0.344 (0.119)	0.968 ± 0.321 (0.104)	0.966 ± 0.302 (0.092)	0.962 ± 0.287 (0.084)	0.962 ± 0.275 (0.077)
MLE	2.268 ± 3.837 (16.330)	1.656 ± 1.260 (2.017)	1.439 ± 0.792 (0.821)	1.337 ± 0.612 (0.489)	1.268 ± 0.511 (0.332)	1.230 ± 0.443 (0.249)	1.193 ± 0.387 (0.187)	1.168 ± 0.351 (0.151)	1.152 ± 0.322 (0.127)	1.135 ± 0.301 (0.109)	1.123 ± 0.282 (0.095)
Method	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**
OLSE	0.962 ± 0.263 (0.070)	0.959 ± 0.254 (0.066)	0.958 ± 0.246 (0.062)	0.960 ± 0.239 (0.059)	0.959 ± 0.230 (0.054)	0.959 ± 0.224 (0.052)	0.961 ± 0.218 (0.049)	0.958 ± 0.208 (0.045)	0.960 ± 0.200 (0.042)	0.961 ± 0.194 (0.039)	0.962 ± 0.187 (0.036)
MLE	1.114 ± 0.265 (0.083)	1.105 ± 0.252 (0.074)	1.097 ± 0.241 (0.067)	1.092 ± 0.231 (0.062)	1.085 ± 0.220 (0.055)	1.080 ± 0.212 (0.051)	1.077 ± 0.205 (0.048)	1.068 ± 0.192 (0.041)	1.062 ± 0.182 (0.037)	1.057 ± 0.174 (0.034)	1.054 ± 0.167 (0.031)
Method	**30**	**35**	**40**	**45**	**50**	**60**	**70**	**80**	**90**	**100**	**200**
OLSE	0.962 ± 0.180 (0.034)	0.962 ± 0.167 (0.029)	0.965 ± 0.157 (0.026)	0.966 ± 0.148 (0.023)	0.968 ± 0.141 (0.021)	0.970 ± 0.130 (0.018)	0.973 ± 0.121 (0.015)	0.974 ± 0.114 (0.014)	0.976 ± 0.108 (0.012)	0.977 ± 0.102 (0.011)	0.985 ± 0.073 (0.006)
MLE	1.048 ± 0.159 (0.028)	1.040 ± 0.144 (0.022)	1.036 ± 0.133 (0.019)	1.031 ± 0.124 (0.016)	1.029 ± 0.116 (0.014)	1.023 ± 0.105 (0.012)	1.020 ± 0.097 (0.010)	1.017 ± 0.090 (0.008)	1.016 ± 0.085 (0.008)	1.014 ± 0.080 (0.007)	1.007 ± 0.056 (0.003)

$\beta_T=5$

Method	3	4	5	6	7	8	9	10	11	12	13
OLSE	1.406 ± 2.205 (5.028)	1.136 ± 1.063 (1.149)	1.051 ± 0.622 (0.389)	1.018 ± 0.499 (0.249)	0.993 ± 0.418 (0.175)	0.983 ± 0.377 (0.143)	0.971 ± 0.340 (0.117)	0.969 ± 0.318 (0.102)	0.966 ± 0.302 (0.092)	0.961 ± 0.285 (0.083)	0.962 ± 0.274 (0.076)
MLE	2.249 ± 3.524 (13.980)	1.658 ± 1.601 (2.995)	1.445 ± 0.829 (0.886)	1.340 ± 0.627 (0.509)	1.270 ± 0.509 (0.332)	1.226 ± 0.435 (0.241)	1.192 ± 0.383 (0.183)	1.170 ± 0.349 (0.151)	1.152 ± 0.322 (0.127)	1.133 ± 0.297 (0.106)	1.123 ± 0.280 (0.094)
Method	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**
OLSE	0.962 ± 0.265 (0.071)	0.959 ± 0.253 (0.066)	0.961 ± 0.246 (0.062)	0.959 ± 0.237 (0.058)	0.959 ± 0.230 (0.055)	0.959 ± 0.225 (0.052)	0.958 ± 0.218 (0.049)	0.960 ± 0.210 (0.046)	0.959 ± 0.199 (0.041)	0.960 ± 0.192 (0.039)	0.960 ± 0.185 (0.036)
MLE	1.116 ± 0.267 (0.085)	1.105 ± 0.253 (0.075)	1.099 ± 0.240 (0.067)	1.090 ± 0.228 (0.060)	1.086 ± 0.220 (0.056)	1.081 ± 0.213 (0.052)	1.075 ± 0.205 (0.048)	1.069 ± 0.193 (0.042)	1.061 ± 0.181 (0.037)	1.056 ± 0.172 (0.033)	1.051 ± 0.165 (0.030)
Method	**30**	**35**	**40**	**45**	**50**	**60**	**70**	**80**	**90**	**100**	**200**
OLSE	0.962 ± 0.181 (0.034)	0.963 ± 0.167 (0.029)	0.964 ± 0.157 (0.026)	0.967 ± 0.148 (0.023)	0.968 ± 0.141 (0.021)	0.970 ± 0.130 (0.018)	0.973 ± 0.121 (0.015)	0.974 ± 0.113 (0.014)	0.975 ± 0.108 (0.012)	0.977 ± 0.102 (0.011)	0.984 ± 0.073 (0.006)
MLE	1.049 ± 0.158 (0.028)	1.041 ± 0.144 (0.022)	1.035 ± 0.133 (0.019)	1.032 ± 0.124 (0.016)	1.028 ± 0.116 (0.014)	1.023 ± 0.106 (0.012)	1.021 ± 0.097 (0.010)	1.018 ± 0.090 (0.008)	1.016 ± 0.085 (0.007)	1.014 ± 0.080 (0.007)	1.007 ± 0.056 (0.003)

$\beta_T=8$

Method	3	4	5	6	7	8	9	10	11	12	13
OLSE	1.440 ± 3.298 (11.073)	1.138 ± 0.932 (0.888)	1.047 ± 0.597 (0.359)	1.009 ± 0.478 (0.229)	0.993 ± 0.420 (0.176)	0.978 ± 0.373 (0.139)	0.973 ± 0.343 (0.118)	0.969 ± 0.321 (0.104)	0.964 ± 0.302 (0.092)	0.962 ± 0.284 (0.082)	0.963 ± 0.274 (0.077)
MLE	2.303 ± 5.028 (26.982)	1.659 ± 1.337 (2.222)	1.439 ± 0.796 (0.826)	1.331 ± 0.606 (0.477)	1.270 ± 0.507 (0.329)	1.223 ± 0.435 (0.239)	1.192 ± 0.387 (0.187)	1.171 ± 0.353 (0.154)	1.151 ± 0.322 (0.126)	1.136 ± 0.298 (0.107)	1.123 ± 0.281 (0.094)
Method	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**
OLSE	0.959 ± 0.262 (0.070)	0.961 ± 0.252 (0.065)	0.961 ± 0.245 (0.062)	0.960 ± 0.237 (0.058)	0.958 ± 0.230 (0.054)	0.961 ± 0.225 (0.052)	0.959 ± 0.219 (0.050)	0.961 ± 0.207 (0.045)	0.960 ± 0.201 (0.042)	0.961 ± 0.193 (0.039)	0.961 ± 0.187 (0.036)
MLE	1.113 ± 0.265 (0.083)	1.107 ± 0.252 (0.075)	1.099 ± 0.241 (0.068)	1.091 ± 0.230 (0.061)	1.084 ± 0.220 (0.055)	1.081 ± 0.213 (0.052)	1.076 ± 0.206 (0.048)	1.069 ± 0.192 (0.042)	1.062 ± 0.181 (0.037)	1.057 ± 0.174 (0.034)	1.051 ± 0.166 (0.030)
Method	**30**	**35**	**40**	**45**	**50**	**60**	**70**	**80**	**90**	**100**	**200**
OLSE	0.962 ± 0.180 (0.034)	0.964 ± 0.166 (0.029)	0.965 ± 0.156 (0.026)	0.967 ± 0.150 (0.023)	0.968 ± 0.142 (0.021)	0.970 ± 0.131 (0.018)	0.973 ± 0.121 (0.016)	0.974 ± 0.114 (0.014)	0.975 ± 0.107 (0.012)	0.977 ± 0.103 (0.011)	0.984 ± 0.073 (0.006)
MLE	1.047 ± 0.158 (0.027)	1.041 ± 0.144 (0.022)	1.035 ± 0.132 (0.019)	1.033 ± 0.125 (0.017)	1.028 ± 0.117 (0.014)	1.023 ± 0.106 (0.012)	1.020 ± 0.098 (0.010)	1.018 ± 0.091 (0.009)	1.016 ± 0.085 (0.007)	1.014 ± 0.080 (0.007)	1.007 ± 0.056 (0.003)

Table 3-3: Simulation results of $\hat{\alpha}$ for complete data, generated by OLSE and MLE, at different n and β_T: the values of $E(\hat{\alpha}) \pm S(\hat{\alpha})$ and $MSE(\hat{\alpha})$ (in parentheses).

β_T	Method	n										
β_T=0.5		3	4	5	6	7	8	9	10	11	12	13
	OLSE	1.823 ± 2.291 (5.928)	1.650 ± 1.792 (3.635)	1.534 ± 1.490 (2.504)	1.448 ± 1.299 (1.888)	1.401 ± 1.159 (1.504)	1.353 ± 1.054 (1.236)	1.323 ± 0.966 (1.038)	1.294 ± 0.895 (0.888)	1.272 ± 0.841 (0.782)	1.256 ± 0.799 (0.704)	1.235 ± 0.761 (0.635)
	MLE	1.477 ± 1.886 (3.783)	1.364 ± 1.483 (2.333)	1.292 ± 1.254 (1.658)	1.236 ± 1.095 (1.254)	1.209 ± 0.988 (1.021)	1.179 ± 0.898 (0.839)	1.160 ± 0.824 (0.705)	1.144 ± 0.776 (0.622)	1.131 ± 0.726 (0.545)	1.122 ± 0.691 (0.493)	1.108 ± 0.658 (0.444)
		14	15	16	17	18	19	20	22	24	26	28
	OLSE	1.222 ± 0.712 (0.556)	1.211 ± 0.686 (0.515)	1.202 ± 0.664 (0.482)	1.196 ± 0.646 (0.455)	1.185 ± 0.614 (0.411)	1.177 ± 0.596 (0.386)	1.167 ± 0.577 (0.360)	1.158 ± 0.550 (0.327)	1.142 ± 0.512 (0.282)	1.133 ± 0.487 (0.255)	1.124 ± 0.469 (0.235)
	MLE	1.103 ± 0.626 (0.403)	1.097 ± 0.603 (0.373)	1.092 ± 0.581 (0.346)	1.090 ± 0.563 (0.325)	1.084 ± 0.544 (0.303)	1.079 ± 0.527 (0.284)	1.072 ± 0.508 (0.263)	1.069 ± 0.487 (0.242)	1.060 ± 0.458 (0.213)	1.054 ± 0.436 (0.193)	1.050 ± 0.424 (0.182)
		30	35	40	45	50	60	70	80	90	100	200
	OLSE	1.125 ± 0.453 (0.221)	1.105 ± 0.410 (0.179)	1.094 ± 0.381 (0.154)	1.086 ± 0.358 (0.135)	1.081 ± 0.338 (0.121)	1.068 ± 0.305 (0.098)	1.061 ± 0.278 (0.081)	1.055 ± 0.259 (0.070)	1.051 ± 0.244 (0.062)	1.047 ± 0.230 (0.055)	1.026 ± 0.159 (0.026)
	MLE	1.052 ± 0.408 (0.170)	1.041 ± 0.372 (0.140)	1.036 ± 0.349 (0.123)	1.032 ± 0.327 (0.108)	1.030 ± 0.310 (0.097)	1.024 ± 0.282 (0.080)	1.021 ± 0.258 (0.067)	1.019 ± 0.242 (0.059)	1.018 ± 0.229 (0.053)	1.015 ± 0.215 (0.046)	1.007 ± 0.151 (0.023)
β_T=1		3	4	5	6	7	8	9	10	11	12	13
	OLSE	1.163 ± 0.684 (0.494)	1.136 ± 0.584 (0.359)	1.122 ± 0.516 (0.281)	1.111 ± 0.470 (0.233)	1.102 ± 0.439 (0.203)	1.090 ± 0.405 (0.172)	1.085 ± 0.383 (0.154)	1.081 ± 0.361 (0.137)	1.073 ± 0.343 (0.123)	1.071 ± 0.330 (0.114)	1.068 ± 0.316 (0.105)
	MLE	1.045 ± 0.621 (0.388)	1.031 ± 0.533 (0.285)	1.030 ± 0.474 (0.226)	1.027 ± 0.433 (0.188)	1.026 ± 0.407 (0.166)	1.018 ± 0.375 (0.141)	1.018 ± 0.355 (0.126)	1.017 ± 0.336 (0.113)	1.013 ± 0.318 (0.102)	1.014 ± 0.306 (0.094)	1.013 ± 0.294 (0.087)
		14	15	16	17	18	19	20	22	24	26	28
	OLSE	1.066 ± 0.305 (0.098)	1.060 ± 0.293 (0.089)	1.056 ± 0.282 (0.083)	1.057 ± 0.276 (0.079)	1.054 ± 0.269 (0.075)	1.051 ± 0.259 (0.070)	1.050 ± 0.254 (0.067)	1.048 ± 0.241 (0.060)	1.045 ± 0.232 (0.056)	1.045 ± 0.223 (0.052)	1.041 ± 0.212 (0.047)
	MLE	1.014 ± 0.284 (0.081)	1.010 ± 0.273 (0.075)	1.009 ± 0.264 (0.070)	1.010 ± 0.257 (0.066)	1.009 ± 0.252 (0.063)	1.008 ± 0.243 (0.059)	1.008 ± 0.238 (0.057)	1.008 ± 0.226 (0.051)	1.007 ± 0.217 (0.047)	1.009 ± 0.209 (0.044)	1.007 ± 0.200 (0.040)
		30	35	40	45	50	60	70	80	90	100	200
	OLSE	1.038 ± 0.206 (0.044)	1.035 ± 0.188 (0.037)	1.033 ± 0.177 (0.032)	1.029 ± 0.166 (0.028)	1.027 ± 0.156 (0.025)	1.023 ± 0.143 (0.021)	1.021 ± 0.133 (0.018)	1.020 ± 0.124 (0.016)	1.017 ± 0.117 (0.014)	1.017 ± 0.110 (0.012)	1.010 ± 0.077 (0.006)
	MLE	1.005 ± 0.193 (0.037)	1.006 ± 0.178 (0.032)	1.005 ± 0.167 (0.028)	1.004 ± 0.157 (0.025)	1.003 ± 0.148 (0.022)	1.002 ± 0.136 (0.018)	1.003 ± 0.126 (0.016)	1.003 ± 0.118 (0.014)	1.001 ± 0.111 (0.012)	1.002 ± 0.105 (0.011)	1.001 ± 0.074 (0.006)
β_T=2		3	4	5	6	7	8	9	10	11	12	13
	OLSE	1.033 ± 0.309 (0.097)	1.034 ± 0.267 (0.072)	1.030 ± 0.241 (0.059)	1.030 ± 0.221 (0.050)	1.028 ± 0.205 (0.043)	1.028 ± 0.193 (0.038)	1.026 ± 0.181 (0.034)	1.026 ± 0.173 (0.031)	1.024 ± 0.165 (0.028)	1.023 ± 0.157 (0.025)	1.022 ± 0.153 (0.024)
	MLE	0.978 ± 0.296 (0.088)	0.985 ± 0.256 (0.066)	0.986 ± 0.231 (0.054)	0.990 ± 0.212 (0.045)	0.992 ± 0.197 (0.039)	0.994 ± 0.185 (0.034)	0.994 ± 0.174 (0.030)	0.995 ± 0.166 (0.027)	0.996 ± 0.158 (0.025)	0.996 ± 0.151 (0.023)	0.996 ± 0.146 (0.021)
		14	15	16	17	18	19	20	22	24	26	28
	OLSE	1.021 ± 0.146 (0.022)	1.020 ± 0.141 (0.020)	1.019 ± 0.136 (0.019)	1.020 ± 0.132 (0.018)	1.018 ± 0.129 (0.017)	1.018 ± 0.125 (0.016)	1.018 ± 0.122 (0.015)	1.017 ± 0.116 (0.014)	1.016 ± 0.113 (0.013)	1.016 ± 0.108 (0.012)	1.015 ± 0.103 (0.011)
	MLE	0.997 ± 0.140 (0.020)	0.997 ± 0.135 (0.018)	0.996 ± 0.131 (0.017)	0.998 ± 0.127 (0.016)	0.997 ± 0.124 (0.015)	0.997 ± 0.120 (0.014)	0.998 ± 0.117 (0.014)	0.998 ± 0.112 (0.012)	0.998 ± 0.108 (0.012)	0.998 ± 0.103 (0.011)	0.998 ± 0.099 (0.010)
		30	35	40	45	50	60	70	80	90	100	200
	OLSE	1.015 ± 0.100 (0.010)	1.013 ± 0.092 (0.009)	1.012 ± 0.087 (0.008)	1.011 ± 0.081 (0.007)	1.011 ± 0.077 (0.006)	1.009 ± 0.071 (0.005)	1.009 ± 0.065 (0.004)	1.008 ± 0.061 (0.004)	1.007 ± 0.057 (0.003)	1.007 ± 0.055 (0.003)	1.004 ± 0.039 (0.002)
	MLE	0.999 ± 0.096 (0.009)	0.999 ± 0.088 (0.008)	0.999 ± 0.083 (0.007)	0.999 ± 0.078 (0.006)	0.999 ± 0.075 (0.006)	0.999 ± 0.068 (0.005)	0.999 ± 0.063 (0.004)	0.999 ± 0.059 (0.003)	0.999 ± 0.055 (0.003)	1.000 ± 0.053 (0.003)	1.000 ± 0.037 (0.001)
β_T=3		3	4	5	6	7	8	9	10	11	12	13
	OLSE	1.010 ± 0.205 (0.042)	1.015 ± 0.178 (0.032)	1.014 ± 0.159 (0.026)	1.014 ± 0.145 (0.021)	1.015 ± 0.135 (0.019)	1.014 ± 0.127 (0.016)	1.014 ± 0.119 (0.014)	1.013 ± 0.114 (0.013)	1.012 ± 0.108 (0.012)	1.012 ± 0.104 (0.011)	1.012 ± 0.101 (0.010)
	MLE	0.974 ± 0.200 (0.041)	0.983 ± 0.174 (0.031)	0.985 ± 0.155 (0.024)	0.988 ± 0.141 (0.020)	0.991 ± 0.132 (0.017)	0.992 ± 0.124 (0.015)	0.993 ± 0.116 (0.014)	0.993 ± 0.110 (0.012)	0.994 ± 0.105 (0.011)	0.994 ± 0.101 (0.010)	0.995 ± 0.098 (0.010)
		14	15	16	17	18	19	20	22	24	26	28
	OLSE	1.012 ± 0.096 (0.009)	1.011 ± 0.094 (0.009)	1.011 ± 0.090 (0.008)	1.011 ± 0.088 (0.008)	1.011 ± 0.085 (0.007)	1.010 ± 0.083 (0.007)	1.010 ± 0.081 (0.007)	1.010 ± 0.077 (0.006)	1.009 ± 0.074 (0.006)	1.009 ± 0.071 (0.005)	1.009 ± 0.069 (0.005)
	MLE	0.995 ± 0.093 (0.009)	0.995 ± 0.090 (0.008)	0.996 ± 0.088 (0.008)	0.997 ± 0.085 (0.007)	0.997 ± 0.082 (0.007)	0.997 ± 0.081 (0.007)	0.997 ± 0.078 (0.006)	0.997 ± 0.075 (0.006)	0.997 ± 0.071 (0.005)	0.998 ± 0.069 (0.005)	0.998 ± 0.066 (0.004)
		30	35	40	45	50	60	70	80	90	100	200
	OLSE	1.009 ± 0.066 (0.004)	1.007 ± 0.061 (0.004)	1.007 ± 0.057 (0.003)	1.007 ± 0.054 (0.003)	1.006 ± 0.051 (0.003)	1.006 ± 0.047 (0.002)	1.005 ± 0.043 (0.002)	1.005 ± 0.041 (0.002)	1.005 ± 0.038 (0.001)	1.004 ± 0.036 (0.001)	1.003 ± 0.026 (0.001)
	MLE	0.998 ± 0.064 (0.004)	0.998 ± 0.059 (0.003)	0.998 ± 0.055 (0.003)	0.999 ± 0.052 (0.003)	0.999 ± 0.050 (0.002)	0.999 ± 0.045 (0.002)	0.999 ± 0.042 (0.002)	0.999 ± 0.039 (0.002)	0.999 ± 0.037 (0.001)	0.999 ± 0.035 (0.001)	1.000 ± 0.025 (0.001)
β_T=5		3	4	5	6	7	8	9	10	11	12	13
	OLSE	1.001 ± 0.124 (0.015)	1.005 ± 0.107 (0.012)	1.006 ± 0.096 (0.009)	1.006 ± 0.088 (0.008)	1.007 ± 0.081 (0.007)	1.006 ± 0.076 (0.006)	1.006 ± 0.072 (0.005)	1.007 ± 0.068 (0.005)	1.007 ± 0.065 (0.004)	1.006 ± 0.062 (0.004)	1.006 ± 0.060 (0.004)
	MLE	0.979 ± 0.123 (0.016)	0.986 ± 0.106 (0.011)	0.988 ± 0.095 (0.009)	0.991 ± 0.086 (0.008)	0.992 ± 0.080 (0.006)	0.993 ± 0.075 (0.006)	0.994 ± 0.070 (0.005)	0.995 ± 0.067 (0.004)	0.996 ± 0.063 (0.004)	0.995 ± 0.061 (0.004)	0.996 ± 0.059 (0.003)
		14	15	16	17	18	19	20	22	24	26	28
	OLSE	1.006 ± 0.058 (0.003)	1.006 ± 0.056 (0.003)	1.006 ± 0.054 (0.003)	1.006 ± 0.052 (0.003)	1.006 ± 0.051 (0.003)	1.005 ± 0.050 (0.003)	1.006 ± 0.048 (0.002)	1.005 ± 0.046 (0.002)	1.005 ± 0.044 (0.002)	1.005 ± 0.042 (0.002)	1.004 ± 0.041 (0.002)
	MLE	0.996 ± 0.056 (0.003)	0.996 ± 0.055 (0.003)	0.997 ± 0.053 (0.003)	0.997 ± 0.051 (0.003)	0.997 ± 0.050 (0.002)	0.997 ± 0.048 (0.002)	0.998 ± 0.047 (0.002)	0.997 ± 0.045 (0.002)	0.998 ± 0.043 (0.002)	0.998 ± 0.041 (0.002)	0.998 ± 0.040 (0.002)
		30	35	40	45	50	60	70	80	90	100	200
	OLSE	1.004 ± 0.040 (0.002)	1.004 ± 0.037 (0.001)	1.004 ± 0.034 (0.001)	1.004 ± 0.032 (0.001)	1.003 ± 0.031 (0.001)	1.003 ± 0.028 (0.001)	1.003 ± 0.026 (0.001)	1.003 ± 0.024 (0.001)	1.002 ± 0.023 (0.001)	1.002 ± 0.022 (0.000)	1.001 ± 0.015 (0.000)
	MLE	0.998 ± 0.039 (0.001)	0.999 ± 0.036 (0.001)	0.999 ± 0.033 (0.001)	0.999 ± 0.031 (0.001)	0.999 ± 0.030 (0.001)	0.999 ± 0.027 (0.001)	0.999 ± 0.025 (0.001)	0.999 ± 0.024 (0.001)	0.999 ± 0.022 (0.000)	0.999 ± 0.021 (0.000)	1.000 ± 0.015 (0.000)
β_T=8		3	4	5	6	7	8	9	10	11	12	13
	OLSE	0.999 ± 0.078 (0.006)	1.001 ± 0.068 (0.005)	1.003 ± 0.060 (0.004)	1.003 ± 0.055 (0.003)	1.003 ± 0.051 (0.003)	1.003 ± 0.048 (0.002)	1.003 ± 0.045 (0.002)	1.004 ± 0.042 (0.002)	1.003 ± 0.041 (0.002)	1.003 ± 0.039 (0.002)	1.003 ± 0.037 (0.001)
	MLE	0.985 ± 0.078 (0.006)	0.989 ± 0.067 (0.005)	0.992 ± 0.060 (0.004)	0.993 ± 0.054 (0.003)	0.994 ± 0.050 (0.003)	0.995 ± 0.047 (0.002)	0.995 ± 0.044 (0.002)	0.996 ± 0.042 (0.002)	0.996 ± 0.040 (0.002)	0.997 ± 0.038 (0.001)	0.997 ± 0.037 (0.001)
		14	15	16	17	18	19	20	22	24	26	28
	OLSE	1.003 ± 0.036 (0.001)	1.003 ± 0.035 (0.001)	1.003 ± 0.034 (0.001)	1.003 ± 0.033 (0.001)	1.003 ± 0.032 (0.001)	1.003 ± 0.031 (0.001)	1.003 ± 0.030 (0.001)	1.003 ± 0.029 (0.001)	1.003 ± 0.028 (0.001)	1.003 ± 0.027 (0.001)	1.003 ± 0.026 (0.001)
	MLE	0.997 ± 0.035 (0.001)	0.997 ± 0.034 (0.001)	0.998 ± 0.033 (0.001)	0.998 ± 0.032 (0.001)	0.998 ± 0.031 (0.001)	0.998 ± 0.030 (0.001)	0.998 ± 0.030 (0.001)	0.998 ± 0.028 (0.001)	0.998 ± 0.027 (0.001)	0.999 ± 0.026 (0.001)	0.999 ± 0.025 (0.001)
		30	35	40	45	50	60	70	80	90	100	200
	OLSE	1.002 ± 0.025 (0.001)	1.002 ± 0.023 (0.001)	1.002 ± 0.021 (0.000)	1.002 ± 0.020 (0.000)	1.002 ± 0.019 (0.000)	1.002 ± 0.018 (0.000)	1.002 ± 0.016 (0.000)	1.002 ± 0.015 (0.000)	1.002 ± 0.014 (0.000)	1.001 ± 0.014 (0.000)	1.001 ± 0.010 (0.000)
	MLE	0.999 ± 0.024 (0.001)	0.999 ± 0.022 (0.000)	0.999 ± 0.021 (0.000)	0.999 ± 0.020 (0.000)	0.999 ± 0.019 (0.000)	0.999 ± 0.017 (0.000)	0.999 ± 0.016 (0.000)	0.999 ± 0.015 (0.000)	1.000 ± 0.014 (0.000)	1.000 ± 0.013 (0.000)	1.000 ± 0.009 (0.000)

3.3.3.2 Simulation Results for Multiply Censored Data

The simulation results for multiply censored data, as can be seen from Table 3-4 – Table 3-7, are presented in four parts: the results for $\hat{\beta}$ at low censoring levels, i.e., $c = 10\% - 40\%$ (Table 3-4), the results for $\hat{\beta}$ at high censoring levels, i.e., $c = 50\% - 80\%$ (Table 3-5), the results for $\hat{\alpha}$ at low censoring levels (Table 3-6) and the results for $\hat{\alpha}$ at high censoring levels (Table 3-7). Please note not all the simulation results are tabulated in the four tables; however, the omitted results will not affect the following conclusions.

<u>Simulation Results for Estimators of β (Table 3-4 and Table 3-5)</u>

1) ***The reliability of the simulation results is judged by the pivotal quantity*** $\hat{\beta}/\beta_T$ *:* For censored data, the properties of the pivotal function $\hat{\beta}/\beta_T$ still apply for MLE and LSE. Therefore, the pivotal quantity can be used to check the reliability of the simulations as it does for the case of complete data. In theory, the values of $E(\hat{\beta})/\beta_T$, $S(\hat{\beta})/\beta_T$ and $MSE(\hat{\beta})/\beta_T^2$, generated by OLSE and MLE, should be constant at different values of β_T. At low censoring levels (refer to Table 3-4), the values of $E(\hat{\beta})/\beta_T$, $S(\hat{\beta})/\beta_T$ and $MSE(\hat{\beta})/\beta_T^2$ at six values of β_T almost do not change for a specific n. Some discrepancies can be observed from the results at high censoring levels (refer to Table 3-5), and the difference of $E(\hat{\beta})/\beta_T$ at different β_T, for example, can be larger than 10% when $c = 80\%$ and $n \le 40$. This is probably due to the complexity in the generation of a highly censored sample. In general, the difference is still acceptable.

2) ***Bias of the OLSE of the shape parameter:*** The relative bias of $\hat{\beta}$ of the OLSE can be smaller than 1, equal to 1, and larger than 1, depending on the combination of n and c. The general trend of the relative bias as a function of n at any specific censoring level, or the general trend of the relative bias as a function of c at any specific sample size, is similar, that is, the bias first decreases with the variable (n or c), then at certain point the bias reaches 0, and after that the bias increases with the variable (n or c). The bias is obviously inconsistent with either n or c. As shown in Table 3-4, at low censoring levels, the bias reaches smallest at the combination of $c = 30\%$ and $n = 150 - 200$, or the combination of $c = 40\%$ and $n = 100 - 150$. At high censoring levels (refer to Table 3-5), the bias reaches smallest at the combination of $c = 50\%$ and $n = 80 - 100$, or the combination of $c = 60\%$ and $n = 50 - 60$, or the combination of $c = 70\%$ and $n = 20 - 30$. The bias is largest at the combination of $c = 10\%$ and $n = 20$ and the combination of $c = 80\%$ and $n = 200$. Although the bias presents a strange pattern as a function of n and c, the relative bias is typically within 5%. The pattern of the bias is further examined in Section 5.4.2.

3) ***Standard deviation and MSE of the OLSE of the shape parameter:*** The values of the relative standard deviation and relative MSE both decrease with the increase of n at a specific c, and consistently increase with the increase of c (from 10% to 80%) at a specific n. The values of $S(\hat{\beta})/\beta_T$ and $MSE(\hat{\beta})/\beta_T^2$ are significant for small samples with very high censoring levels, e.g., at the combination of $c = 80\%$ and $n = 20$.

4) ***Comparison between OLSE and MLE for estimating*** β*:* In view of bias, standard deviation or MSE, OLSE outperforms MLE for estimating β in most cases, except when $c = 10\% - 20\%$. The relative bias, relative standard deviation or relative MSE of $\hat{\beta}$ of the OLSE is significantly smaller than that of the MLE at high censoring levels (50% – 80%). Especially at $c = 80\%$, the relative bias of $\hat{\beta}$ of the OLSE is 20% – 40% smaller than that of the MLE. Although MLE performs inferior to OLSE for estimating β in most times under the simulation conditions examined, $\hat{\beta}$ of the MLE has good consistency and is asymptotically unbiased as sample size increases.

Simulation Results for Estimators of α (Table 3-6 and Table 3-7)

1) ***General results****:* The bias, standard deviation and MSE of $\hat{\alpha}$ of both methods decrease as β_T increases. The decrease is dramatic from $\beta_T = 0.5$ to $\beta_T = 1$. From the results at high censoring levels (refer to Table 3-6), both methods, especially OLSE, are unstable for estimating α at $\beta_T = 0.5$, and both methods result in extremely large estimates especially when the sample size is small.

2) ***Bias of the OLSE of the scale parameter****:* The bias is extremely large at $\beta_T = 0.5$ at all censoring levels and the results are unstable when $c \geq 50\%$. At low censoring levels (refer to Table 3-6), the bias increases as c increases at all combinations of β_T and n, and the bias decreases as n increases at all combinations of β_T and c. The bias is significant ($\geq 10\%$) when $\beta_T \leq 2$ and $c \geq 30\%$, but is typically within 2% at $\beta_T = 5$ and within 1% at $\beta_T = 8$. On

the other hand, the bias at high censoring levels (refer to Table 3-7) is inconsistent with c and consistent with n. At $\beta_T = 3, 5, 8$, the bias reaches smallest at $c = 70\%$. Generally the bias at high censoring levels at any combination of β_T and n is larger than that at low censoring levels. At high censoring levels (50% – 80%), the bias is larger than 10% when $\beta_T \leq 3$ and all combinations of n and c. At $\beta_T = 8$, the bias is typically within 5%.

3) ***Standard deviation and MSE of the OLSE of the scale parameter:*** The results regarding the standard deviation and MSE of $\hat{\alpha}$ of the OLSE are similar to those of the bias. Unstable results and extremely large values can be observed at $\beta_T = 0.5$. Good consistency of standard deviation or MSE as a function of n and c can be observed at low censoring levels (10% – 40%); however, the standard deviation or MSE is inconsistent with c at high censoring levels. The standard deviation and MSE reach smallest when $\beta_T = 8$ and $c = 70\%$.

4) ***Comparison between OLSE and MLE:*** MLE outperforms OLSE for estimating α in view of bias, standard deviation and MSE at all conditions examined, and is significantly better than OLSE when $\beta_T = 0.5$ and $\beta_T = 1$. The difference between the two methods decreases as β_T increases, and at $\beta_T = 8$, both estimators of α are nearly unbiased and have very small standard deviation and MSE.

5) ***Comparison between the results for $\hat{\alpha}$ and $\hat{\beta}$ of the OLSE:*** The bias, standard deviation and MSE of $\hat{\alpha}$ highly depend on β_T while those of $\hat{\beta}$ is independent of β_T. MLE is generally better for estimating α while OLSE is better for estimating β when $c \geq 30\%$.

Table 3-4: Simulation results of $\hat{\beta}$ for multiply censored data, generated by OLSE and MLE, at different n, β_T and c (part I – low censoring levels): the values of $E(\hat{\beta})/\beta_T \pm S(\hat{\beta})/\beta$ and $MSE(\hat{\beta})/\beta_T^2$ (in parentheses).

			n									
		Method	20	30	40	50	60	70	80	100	150	200
βT=0.5	10%	OLSE	0.899 ± 0.214 (0.056)	0.918 ± 0.181 (0.040)	0.926 ± 0.159 (0.031)	0.934 ± 0.144 (0.025)	0.942 ± 0.136 (0.022)	0.950 ± 0.127 (0.019)	0.953 ± 0.118 (0.016)	0.958 ± 0.107 (0.013)	0.969 ± 0.088 (0.009)	0.976 ± 0.077 (0.006)
		MLE	1.096 ± 0.220 (0.057)	1.068 ± 0.170 (0.033)	1.053 ± 0.142 (0.023)	1.045 ± 0.124 (0.017)	1.041 ± 0.113 (0.015)	1.038 ± 0.105 (0.012)	1.034 ± 0.097 (0.010)	1.030 ± 0.085 (0.008)	1.025 ± 0.069 (0.005)	1.022 ± 0.059 (0.004)
	20%	OLSE	0.905 ± 0.232 (0.063)	0.924 ± 0.193 (0.043)	0.932 ± 0.170 (0.033)	0.946 ± 0.153 (0.026)	0.949 ± 0.142 (0.023)	0.954 ± 0.132 (0.019)	0.961 ± 0.125 (0.017)	0.964 ± 0.113 (0.014)	0.976 ± 0.094 (0.009)	0.984 ± 0.083 (0.007)
		MLE	1.117 ± 0.241 (0.072)	1.089 ± 0.183 (0.042)	1.070 ± 0.152 (0.028)	1.066 ± 0.132 (0.022)	1.057 ± 0.121 (0.018)	1.055 ± 0.110 (0.015)	1.052 ± 0.103 (0.013)	1.045 ± 0.090 (0.010)	1.040 ± 0.074 (0.007)	1.038 ± 0.063 (0.005)
	30%	OLSE	0.918 ± 0.248 (0.068)	0.930 ± 0.204 (0.046)	0.942 ± 0.184 (0.037)	0.953 ± 0.168 (0.030)	0.956 ± 0.152 (0.025)	0.962 ± 0.142 (0.022)	0.967 ± 0.135 (0.019)	0.973 ± 0.121 (0.015)	0.985 ± 0.101 (0.010)	0.993 ± 0.089 (0.008)
		MLE	1.148 ± 0.266 (0.093)	1.109 ± 0.199 (0.051)	1.093 ± 0.168 (0.037)	1.086 ± 0.146 (0.029)	1.077 ± 0.132 (0.023)	1.072 ± 0.119 (0.019)	1.069 ± 0.112 (0.017)	1.064 ± 0.099 (0.014)	1.058 ± 0.080 (0.010)	1.055 ± 0.069 (0.008)
	40%	OLSE	0.926 ± 0.265 (0.076)	0.934 ± 0.221 (0.053)	0.950 ± 0.197 (0.041)	0.962 ± 0.179 (0.033)	0.964 ± 0.167 (0.029)	0.971 ± 0.154 (0.025)	0.972 ± 0.146 (0.022)	0.978 ± 0.131 (0.018)	0.994 ± 0.110 (0.012)	1.001 ± 0.096 (0.009)
		MLE	1.179 ± 0.295 (0.119)	1.130 ± 0.219 (0.065)	1.116 ± 0.184 (0.047)	1.109 ± 0.161 (0.038)	1.097 ± 0.144 (0.030)	1.093 ± 0.133 (0.026)	1.088 ± 0.122 (0.023)	1.083 ± 0.107 (0.018)	1.077 ± 0.087 (0.014)	1.074 ± 0.075 (0.011)
βT=1	10%	OLSE	0.902 ± 0.215 (0.056)	0.917 ± 0.179 (0.039)	0.928 ± 0.160 (0.031)	0.938 ± 0.144 (0.025)	0.943 ± 0.132 (0.021)	0.949 ± 0.126 (0.018)	0.955 ± 0.119 (0.016)	0.962 ± 0.106 (0.013)	0.972 ± 0.088 (0.009)	0.979 ± 0.077 (0.006)
		MLE	1.100 ± 0.220 (0.058)	1.070 ± 0.169 (0.033)	1.056 ± 0.141 (0.023)	1.049 ± 0.125 (0.018)	1.042 ± 0.112 (0.014)	1.037 ± 0.104 (0.012)	1.037 ± 0.096 (0.011)	1.032 ± 0.085 (0.008)	1.026 ± 0.069 (0.005)	1.024 ± 0.059 (0.004)
	20%	OLSE	0.908 ± 0.227 (0.060)	0.927 ± 0.192 (0.042)	0.938 ± 0.169 (0.032)	0.948 ± 0.154 (0.026)	0.954 ± 0.142 (0.022)	0.961 ± 0.133 (0.019)	0.963 ± 0.125 (0.017)	0.973 ± 0.113 (0.014)	0.983 ± 0.094 (0.009)	0.990 ± 0.082 (0.007)
		MLE	1.123 ± 0.234 (0.070)	1.094 ± 0.183 (0.042)	1.077 ± 0.152 (0.029)	1.069 ± 0.134 (0.023)	1.063 ± 0.120 (0.018)	1.061 ± 0.111 (0.016)	1.055 ± 0.102 (0.013)	1.052 ± 0.091 (0.011)	1.045 ± 0.074 (0.007)	1.043 ± 0.063 (0.006)
	30%	OLSE	0.917 ± 0.241 (0.065)	0.935 ± 0.203 (0.046)	0.946 ± 0.180 (0.035)	0.957 ± 0.164 (0.029)	0.965 ± 0.151 (0.024)	0.970 ± 0.141 (0.021)	0.972 ± 0.134 (0.019)	0.982 ± 0.122 (0.015)	0.995 ± 0.101 (0.010)	1.003 ± 0.089 (0.008)
		MLE	1.157 ± 0.259 (0.092)	1.119 ± 0.199 (0.054)	1.101 ± 0.166 (0.038)	1.093 ± 0.145 (0.030)	1.085 ± 0.131 (0.024)	1.080 ± 0.119 (0.020)	1.076 ± 0.111 (0.018)	1.073 ± 0.099 (0.015)	1.067 ± 0.079 (0.011)	1.064 ± 0.068 (0.009)
	40%	OLSE	0.930 ± 0.269 (0.078)	0.945 ± 0.216 (0.050)	0.961 ± 0.198 (0.041)	0.970 ± 0.176 (0.032)	0.975 ± 0.164 (0.028)	0.983 ± 0.155 (0.024)	0.988 ± 0.143 (0.021)	0.994 ± 0.132 (0.017)	1.008 ± 0.110 (0.012)	1.015 ± 0.097 (0.010)
		MLE	1.195 ± 0.300 (0.128)	1.148 ± 0.217 (0.069)	1.132 ± 0.184 (0.051)	1.121 ± 0.159 (0.040)	1.110 ± 0.143 (0.033)	1.108 ± 0.133 (0.029)	1.103 ± 0.120 (0.025)	1.096 ± 0.108 (0.021)	1.090 ± 0.087 (0.016)	1.087 ± 0.075 (0.013)
βT=2	10%	OLSE	0.901 ± 0.215 (0.056)	0.918 ± 0.180 (0.039)	0.929 ± 0.159 (0.030)	0.938 ± 0.145 (0.025)	0.944 ± 0.133 (0.021)	0.950 ± 0.124 (0.018)	0.955 ± 0.118 (0.016)	0.962 ± 0.106 (0.013)	0.974 ± 0.088 (0.009)	0.980 ± 0.077 (0.006)
		MLE	1.101 ± 0.220 (0.059)	1.070 ± 0.169 (0.033)	1.055 ± 0.141 (0.023)	1.047 ± 0.125 (0.018)	1.041 ± 0.112 (0.014)	1.037 ± 0.103 (0.012)	1.035 ± 0.096 (0.010)	1.031 ± 0.085 (0.008)	1.025 ± 0.069 (0.005)	1.024 ± 0.059 (0.004)
	20%	OLSE	0.907 ± 0.227 (0.060)	0.926 ± 0.188 (0.041)	0.938 ± 0.168 (0.032)	0.948 ± 0.152 (0.026)	0.953 ± 0.141 (0.022)	0.960 ± 0.131 (0.019)	0.965 ± 0.124 (0.017)	0.972 ± 0.112 (0.013)	0.983 ± 0.094 (0.009)	0.992 ± 0.082 (0.007)
		MLE	1.125 ± 0.239 (0.073)	1.092 ± 0.181 (0.041)	1.075 ± 0.152 (0.029)	1.067 ± 0.133 (0.022)	1.060 ± 0.120 (0.018)	1.056 ± 0.110 (0.015)	1.053 ± 0.102 (0.013)	1.048 ± 0.090 (0.010)	1.043 ± 0.074 (0.007)	1.041 ± 0.063 (0.006)
	30%	OLSE	0.919 ± 0.244 (0.066)	0.935 ± 0.201 (0.045)	0.946 ± 0.178 (0.035)	0.958 ± 0.162 (0.028)	0.964 ± 0.149 (0.024)	0.970 ± 0.139 (0.020)	0.976 ± 0.131 (0.018)	0.984 ± 0.119 (0.014)	0.996 ± 0.099 (0.010)	1.004 ± 0.087 (0.008)
		MLE	1.159 ± 0.266 (0.096)	1.119 ± 0.198 (0.053)	1.098 ± 0.164 (0.037)	1.090 ± 0.145 (0.029)	1.081 ± 0.130 (0.023)	1.077 ± 0.118 (0.020)	1.074 ± 0.111 (0.018)	1.070 ± 0.097 (0.014)	1.063 ± 0.078 (0.010)	1.060 ± 0.067 (0.008)
	40%	OLSE	0.930 ± 0.262 (0.073)	0.947 ± 0.216 (0.050)	0.958 ± 0.191 (0.038)	0.967 ± 0.172 (0.031)	0.978 ± 0.161 (0.026)	0.982 ± 0.150 (0.023)	0.989 ± 0.142 (0.020)	0.996 ± 0.129 (0.017)	1.011 ± 0.108 (0.012)	1.019 ± 0.095 (0.009)
		MLE	1.198 ± 0.297 (0.127)	1.151 ± 0.221 (0.071)	1.129 ± 0.182 (0.049)	1.116 ± 0.157 (0.038)	1.111 ± 0.143 (0.033)	1.103 ± 0.129 (0.027)	1.100 ± 0.121 (0.025)	1.094 ± 0.107 (0.020)	1.086 ± 0.086 (0.015)	1.084 ± 0.074 (0.013)
βT=3	10%	OLSE	0.898 ± 0.214 (0.056)	0.912 ± 0.179 (0.040)	0.928 ± 0.159 (0.031)	0.936 ± 0.143 (0.024)	0.944 ± 0.132 (0.021)	0.949 ± 0.122 (0.017)	0.954 ± 0.117 (0.016)	0.960 ± 0.107 (0.013)	0.971 ± 0.087 (0.008)	0.978 ± 0.076 (0.006)
		MLE	1.098 ± 0.219 (0.057)	1.067 ± 0.169 (0.033)	1.054 ± 0.142 (0.023)	1.043 ± 0.122 (0.017)	1.039 ± 0.111 (0.014)	1.035 ± 0.101 (0.011)	1.032 ± 0.095 (0.010)	1.029 ± 0.085 (0.008)	1.024 ± 0.068 (0.005)	1.021 ± 0.059 (0.004)
	20%	OLSE	0.905 ± 0.225 (0.060)	0.924 ± 0.188 (0.041)	0.935 ± 0.164 (0.031)	0.943 ± 0.150 (0.026)	0.950 ± 0.139 (0.022)	0.958 ± 0.130 (0.019)	0.962 ± 0.123 (0.016)	0.968 ± 0.111 (0.013)	0.981 ± 0.093 (0.009)	0.989 ± 0.080 (0.007)
		MLE	1.122 ± 0.235 (0.070)	1.089 ± 0.182 (0.041)	1.072 ± 0.151 (0.028)	1.062 ± 0.131 (0.021)	1.056 ± 0.119 (0.017)	1.051 ± 0.109 (0.014)	1.049 ± 0.101 (0.013)	1.044 ± 0.091 (0.010)	1.039 ± 0.073 (0.007)	1.036 ± 0.062 (0.005)
	30%	OLSE	0.916 ± 0.245 (0.067)	0.933 ± 0.203 (0.046)	0.943 ± 0.176 (0.034)	0.949 ± 0.160 (0.028)	0.962 ± 0.147 (0.023)	0.966 ± 0.138 (0.020)	0.973 ± 0.131 (0.018)	0.981 ± 0.117 (0.014)	0.993 ± 0.098 (0.010)	1.001 ± 0.086 (0.007)
		MLE	1.156 ± 0.268 (0.096)	1.114 ± 0.200 (0.053)	1.095 ± 0.163 (0.036)	1.081 ± 0.143 (0.027)	1.077 ± 0.129 (0.023)	1.070 ± 0.117 (0.019)	1.068 ± 0.110 (0.017)	1.063 ± 0.096 (0.013)	1.056 ± 0.078 (0.009)	1.053 ± 0.067 (0.007)
	40%	OLSE	0.923 ± 0.260 (0.074)	0.938 ± 0.214 (0.050)	0.950 ± 0.191 (0.039)	0.961 ± 0.173 (0.031)	0.969 ± 0.158 (0.026)	0.981 ± 0.149 (0.023)	0.982 ± 0.139 (0.020)	0.991 ± 0.127 (0.016)	1.006 ± 0.106 (0.011)	1.014 ± 0.093 (0.009)
		MLE	1.193 ± 0.297 (0.125)	1.141 ± 0.217 (0.067)	1.119 ± 0.182 (0.047)	1.107 ± 0.157 (0.036)	1.100 ± 0.140 (0.029)	1.097 ± 0.130 (0.026)	1.091 ± 0.118 (0.022)	1.085 ± 0.106 (0.018)	1.078 ± 0.085 (0.013)	1.074 ± 0.073 (0.011)
βT=5	10%	OLSE	0.894 ± 0.213 (0.056)	0.910 ± 0.179 (0.040)	0.925 ± 0.160 (0.031)	0.934 ± 0.146 (0.026)	0.939 ± 0.132 (0.021)	0.947 ± 0.124 (0.018)	0.950 ± 0.117 (0.016)	0.958 ± 0.105 (0.013)	0.969 ± 0.087 (0.009)	0.976 ± 0.076 (0.006)
		MLE	1.092 ± 0.219 (0.056)	1.062 ± 0.170 (0.033)	1.050 ± 0.144 (0.023)	1.041 ± 0.125 (0.017)	1.035 ± 0.112 (0.014)	1.032 ± 0.103 (0.012)	1.030 ± 0.095 (0.010)	1.026 ± 0.084 (0.008)	1.020 ± 0.068 (0.005)	1.018 ± 0.059 (0.004)
	20%	OLSE	0.899 ± 0.228 (0.062)	0.915 ± 0.188 (0.043)	0.931 ± 0.165 (0.032)	0.938 ± 0.152 (0.027)	0.945 ± 0.138 (0.022)	0.950 ± 0.130 (0.019)	0.957 ± 0.123 (0.017)	0.963 ± 0.110 (0.013)	0.977 ± 0.092 (0.009)	0.984 ± 0.081 (0.007)
		MLE	1.114 ± 0.240 (0.071)	1.081 ± 0.182 (0.040)	1.066 ± 0.151 (0.027)	1.056 ± 0.133 (0.021)	1.049 ± 0.120 (0.017)	1.045 ± 0.110 (0.014)	1.043 ± 0.102 (0.012)	1.037 ± 0.090 (0.009)	1.032 ± 0.073 (0.006)	1.029 ± 0.062 (0.005)
	30%	OLSE	0.907 ± 0.246 (0.069)	0.927 ± 0.201 (0.046)	0.935 ± 0.174 (0.035)	0.945 ± 0.160 (0.029)	0.952 ± 0.147 (0.024)	0.958 ± 0.139 (0.021)	0.966 ± 0.133 (0.019)	0.969 ± 0.117 (0.015)	0.985 ± 0.098 (0.010)	0.992 ± 0.085 (0.007)
		MLE	1.150 ± 0.267 (0.094)	1.109 ± 0.200 (0.052)	1.085 ± 0.162 (0.034)	1.076 ± 0.145 (0.027)	1.067 ± 0.129 (0.021)	1.063 ± 0.119 (0.018)	1.058 ± 0.111 (0.016)	1.051 ± 0.097 (0.012)	1.046 ± 0.078 (0.008)	1.043 ± 0.067 (0.006)
	40%	OLSE	0.910 ± 0.256 (0.073)	0.931 ± 0.215 (0.051)	0.942 ± 0.190 (0.040)	0.950 ± 0.172 (0.032)	0.959 ± 0.159 (0.027)	0.966 ± 0.147 (0.023)	0.973 ± 0.140 (0.020)	0.978 ± 0.124 (0.016)	0.995 ± 0.104 (0.011)	1.003 ± 0.092 (0.008)
		MLE	1.181 ± 0.293 (0.118)	1.136 ± 0.221 (0.067)	1.111 ± 0.181 (0.045)	1.097 ± 0.158 (0.034)	1.086 ± 0.141 (0.027)	1.083 ± 0.128 (0.023)	1.080 ± 0.121 (0.021)	1.070 ± 0.104 (0.016)	1.065 ± 0.084 (0.011)	1.061 ± 0.073 (0.009)
βT=8	10%	OLSE	0.893 ± 0.217 (0.058)	0.911 ± 0.180 (0.040)	0.923 ± 0.161 (0.032)	0.931 ± 0.144 (0.025)	0.939 ± 0.133 (0.022)	0.943 ± 0.124 (0.019)	0.949 ± 0.116 (0.016)	0.957 ± 0.106 (0.013)	0.966 ± 0.088 (0.009)	0.973 ± 0.077 (0.007)
		MLE	1.095 ± 0.222 (0.058)	1.062 ± 0.169 (0.032)	1.047 ± 0.143 (0.023)	1.039 ± 0.124 (0.017)	1.033 ± 0.111 (0.014)	1.029 ± 0.103 (0.011)	1.027 ± 0.095 (0.010)	1.023 ± 0.086 (0.008)	1.017 ± 0.069 (0.005)	1.015 ± 0.059 (0.004)
	20%	OLSE	0.894 ± 0.228 (0.063)	0.914 ± 0.190 (0.043)	0.925 ± 0.168 (0.034)	0.935 ± 0.151 (0.027)	0.943 ± 0.139 (0.023)	0.946 ± 0.129 (0.020)	0.952 ± 0.122 (0.017)	0.959 ± 0.111 (0.014)	0.971 ± 0.091 (0.009)	0.978 ± 0.080 (0.007)
		MLE	1.111 ± 0.241 (0.071)	1.077 ± 0.181 (0.039)	1.062 ± 0.153 (0.027)	1.052 ± 0.134 (0.021)	1.046 ± 0.119 (0.016)	1.039 ± 0.109 (0.014)	1.038 ± 0.101 (0.012)	1.033 ± 0.090 (0.009)	1.026 ± 0.073 (0.006)	1.024 ± 0.062 (0.004)
	30%	OLSE	0.892 ± 0.234 (0.066)	0.912 ± 0.199 (0.048)	0.926 ± 0.177 (0.037)	0.938 ± 0.157 (0.029)	0.944 ± 0.145 (0.024)	0.951 ± 0.136 (0.021)	0.958 ± 0.130 (0.019)	0.963 ± 0.117 (0.015)	0.976 ± 0.098 (0.010)	0.985 ± 0.085 (0.007)
		MLE	1.130 ± 0.257 (0.083)	1.095 ± 0.199 (0.048)	1.076 ± 0.166 (0.033)	1.066 ± 0.143 (0.025)	1.058 ± 0.128 (0.020)	1.054 ± 0.116 (0.016)	1.051 ± 0.109 (0.014)	1.044 ± 0.096 (0.011)	1.038 ± 0.078 (0.007)	1.035 ± 0.067 (0.006)
	40%	OLSE	0.896 ± 0.259 (0.078)	0.916 ± 0.213 (0.052)	0.928 ± 0.186 (0.040)	0.942 ± 0.169 (0.032)	0.949 ± 0.158 (0.027)	0.957 ± 0.146 (0.023)	0.965 ± 0.138 (0.020)	0.969 ± 0.125 (0.017)	0.984 ± 0.105 (0.011)	0.992 ± 0.091 (0.008)
		MLE	1.165 ± 0.297 (0.115)	1.119 ± 0.218 (0.061)	1.096 ± 0.179 (0.041)	1.085 ± 0.156 (0.032)	1.077 ± 0.142 (0.026)	1.072 ± 0.128 (0.022)	1.068 ± 0.119 (0.019)	1.060 ± 0.106 (0.015)	1.052 ± 0.085 (0.010)	1.048 ± 0.073 (0.008)

Table 3-5: Simulation results of $\hat{\beta}$ for multiply censored data, generated by OLSE and MLE, at different n, β_T and c (part II – high censoring levels): the values of $E(\hat{\beta})/\beta_T \pm S(\hat{\beta})/\beta$ and $MSE(\hat{\beta})/\beta_T^2$ (in parentheses).

		Method	n = 20	30	40	50	60	70	80	100	150	200
β_T=0.5	50%	OLSE	0.938 ± 0.285 (0.085)	0.949 ± 0.241 (0.061)	0.955 ± 0.211 (0.046)	0.968 ± 0.195 (0.039)	0.976 ± 0.181 (0.033)	0.981 ± 0.171 (0.030)	0.984 ± 0.161 (0.026)	0.992 ± 0.146 (0.021)	1.004 ± 0.120 (0.014)	1.010 ± 0.106 (0.011)
		MLE	1.205 ± 0.331 (0.151)	1.167 ± 0.250 (0.090)	1.136 ± 0.202 (0.059)	1.131 ± 0.181 (0.050)	1.124 ± 0.162 (0.042)	1.120 ± 0.150 (0.037)	1.112 ± 0.137 (0.031)	1.107 ± 0.121 (0.026)	1.098 ± 0.097 (0.019)	1.094 ± 0.083 (0.016)
	60%	OLSE	0.958 ± 0.333 (0.113)	0.958 ± 0.272 (0.076)	0.967 ± 0.236 (0.057)	0.976 ± 0.217 (0.048)	0.980 ± 0.201 (0.041)	0.988 ± 0.187 (0.035)	0.989 ± 0.177 (0.032)	1.000 ± 0.164 (0.027)	1.011 ± 0.137 (0.019)	1.019 ± 0.120 (0.015)
		MLE	1.261 ± 0.398 (0.227)	1.193 ± 0.290 (0.122)	1.174 ± 0.235 (0.085)	1.159 ± 0.206 (0.068)	1.146 ± 0.185 (0.055)	1.141 ± 0.169 (0.048)	1.135 ± 0.155 (0.042)	1.132 ± 0.139 (0.037)	1.123 ± 0.112 (0.028)	1.117 ± 0.095 (0.023)
	70%	OLSE	1.003 ± 0.429 (0.184)	0.983 ± 0.315 (0.100)	0.984 ± 0.276 (0.076)	0.994 ± 0.253 (0.064)	0.994 ± 0.230 (0.053)	0.997 ± 0.215 (0.046)	1.004 ± 0.204 (0.042)	1.009 ± 0.186 (0.035)	1.022 ± 0.155 (0.025)	1.029 ± 0.139 (0.020)
		MLE	1.327 ± 0.510 (0.367)	1.245 ± 0.356 (0.187)	1.211 ± 0.290 (0.128)	1.195 ± 0.249 (0.100)	1.181 ± 0.221 (0.082)	1.173 ± 0.204 (0.071)	1.168 ± 0.187 (0.063)	1.159 ± 0.164 (0.052)	1.149 ± 0.132 (0.040)	1.143 ± 0.113 (0.033)
	80%	OLSE	1.098 ± 0.610 (0.382)	1.042 ± 0.417 (0.176)	1.024 ± 0.347 (0.121)	1.019 ± 0.310 (0.096)	1.019 ± 0.283 (0.081)	1.020 ± 0.264 (0.070)	1.025 ± 0.251 (0.064)	1.024 ± 0.225 (0.051)	1.031 ± 0.192 (0.038)	1.042 ± 0.169 (0.030)
		MLE	1.449 ± 0.735 (0.741)	1.328 ± 0.493 (0.350)	1.275 ± 0.382 (0.222)	1.248 ± 0.331 (0.171)	1.228 ± 0.287 (0.134)	1.217 ± 0.263 (0.116)	1.212 ± 0.242 (0.103)	1.197 ± 0.206 (0.081)	1.179 ± 0.169 (0.061)	1.176 ± 0.144 (0.052)
β_T=1	50%	OLSE	0.944 ± 0.287 (0.086)	0.953 ± 0.238 (0.059)	0.967 ± 0.212 (0.046)	0.983 ± 0.193 (0.038)	0.989 ± 0.178 (0.032)	0.994 ± 0.170 (0.029)	0.999 ± 0.160 (0.026)	1.006 ± 0.144 (0.021)	1.022 ± 0.121 (0.015)	1.030 ± 0.106 (0.012)
		MLE	1.236 ± 0.336 (0.169)	1.182 ± 0.249 (0.095)	1.161 ± 0.206 (0.068)	1.152 ± 0.179 (0.055)	1.142 ± 0.160 (0.046)	1.136 ± 0.148 (0.040)	1.132 ± 0.138 (0.037)	1.123 ± 0.119 (0.029)	1.117 ± 0.097 (0.023)	1.113 ± 0.083 (0.020)
	60%	OLSE	0.971 ± 0.339 (0.116)	0.976 ± 0.268 (0.072)	0.992 ± 0.239 (0.057)	0.999 ± 0.216 (0.046)	1.001 ± 0.197 (0.039)	1.010 ± 0.186 (0.035)	1.014 ± 0.177 (0.031)	1.023 ± 0.161 (0.026)	1.036 ± 0.136 (0.020)	1.046 ± 0.120 (0.016)
		MLE	1.298 ± 0.408 (0.255)	1.233 ± 0.291 (0.139)	1.209 ± 0.242 (0.102)	1.191 ± 0.205 (0.079)	1.176 ± 0.185 (0.065)	1.171 ± 0.168 (0.057)	1.166 ± 0.156 (0.052)	1.158 ± 0.138 (0.044)	1.117 ± 0.097 (0.023)	1.113 ± 0.083 (0.020)
	70%	OLSE	1.013 ± 0.417 (0.174)	1.005 ± 0.320 (0.103)	1.008 ± 0.275 (0.076)	1.014 ± 0.246 (0.061)	1.020 ± 0.229 (0.053)	1.029 ± 0.218 (0.048)	1.029 ± 0.203 (0.042)	1.036 ± 0.184 (0.035)	1.053 ± 0.157 (0.028)	1.063 ± 0.139 (0.023)
		MLE	1.386 ± 0.510 (0.409)	1.297 ± 0.370 (0.225)	1.258 ± 0.294 (0.153)	1.236 ± 0.246 (0.116)	1.224 ± 0.223 (0.100)	1.215 ± 0.206 (0.089)	1.205 ± 0.188 (0.078)	1.197 ± 0.160 (0.064)	1.188 ± 0.131 (0.052)	1.182 ± 0.113 (0.046)
	80%	OLSE	1.111 ± 0.621 (0.398)	1.057 ± 0.410 (0.171)	1.043 ± 0.335 (0.114)	1.053 ± 0.312 (0.100)	1.046 ± 0.282 (0.081)	1.053 ± 0.263 (0.072)	1.053 ± 0.247 (0.064)	1.059 ± 0.226 (0.054)	1.074 ± 0.190 (0.042)	1.083 ± 0.170 (0.036)
		MLE	1.524 ± 0.779 (0.882)	1.395 ± 0.497 (0.403)	1.333 ± 0.390 (0.263)	1.313 ± 0.341 (0.215)	1.285 ± 0.294 (0.168)	1.274 ± 0.263 (0.144)	1.262 ± 0.241 (0.127)	1.251 ± 0.211 (0.107)	1.236 ± 0.167 (0.084)	1.228 ± 0.144 (0.073)
β_T=2	50%	OLSE	0.939 ± 0.284 (0.084)	0.956 ± 0.235 (0.057)	0.978 ± 0.208 (0.044)	0.981 ± 0.187 (0.035)	0.992 ± 0.175 (0.031)	0.995 ± 0.164 (0.027)	1.001 ± 0.155 (0.024)	1.010 ± 0.139 (0.020)	1.026 ± 0.119 (0.015)	1.035 ± 0.103 (0.012)
		MLE	1.242 ± 0.337 (0.172)	1.189 ± 0.248 (0.097)	1.168 ± 0.206 (0.071)	1.151 ± 0.178 (0.054)	1.143 ± 0.158 (0.045)	1.133 ± 0.144 (0.038)	1.131 ± 0.136 (0.036)	1.124 ± 0.117 (0.029)	1.115 ± 0.096 (0.022)	1.111 ± 0.081 (0.019)
	60%	OLSE	0.964 ± 0.330 (0.110)	0.981 ± 0.267 (0.072)	0.991 ± 0.229 (0.052)	0.998 ± 0.208 (0.043)	1.009 ± 0.195 (0.038)	1.018 ± 0.184 (0.034)	1.019 ± 0.172 (0.030)	1.027 ± 0.155 (0.025)	1.045 ± 0.132 (0.020)	1.056 ± 0.115 (0.016)
		MLE	1.317 ± 0.418 (0.275)	1.250 ± 0.298 (0.151)	1.215 ± 0.240 (0.104)	1.194 ± 0.205 (0.080)	1.184 ± 0.184 (0.068)	1.179 ± 0.168 (0.060)	1.168 ± 0.154 (0.052)	1.159 ± 0.135 (0.043)	1.150 ± 0.107 (0.034)	1.145 ± 0.091 (0.029)
	70%	OLSE	1.003 ± 0.398 (0.158)	1.005 ± 0.310 (0.096)	1.013 ± 0.271 (0.073)	1.017 ± 0.238 (0.057)	1.028 ± 0.224 (0.051)	1.034 ± 0.210 (0.045)	1.038 ± 0.198 (0.041)	1.050 ± 0.180 (0.035)	1.070 ± 0.152 (0.028)	1.076 ± 0.134 (0.024)
		MLE	1.426 ± 0.544 (0.477)	1.326 ± 0.370 (0.243)	1.279 ± 0.296 (0.165)	1.254 ± 0.250 (0.127)	1.238 ± 0.222 (0.106)	1.224 ± 0.201 (0.091)	1.219 ± 0.185 (0.082)	1.209 ± 0.162 (0.070)	1.198 ± 0.129 (0.056)	1.188 ± 0.110 (0.047)
	80%	OLSE	1.113 ± 0.652 (0.437)	1.055 ± 0.405 (0.167)	1.050 ± 0.329 (0.111)	1.051 ± 0.292 (0.088)	1.054 ± 0.270 (0.076)	1.062 ± 0.252 (0.068)	1.066 ± 0.238 (0.061)	1.075 ± 0.217 (0.053)	1.093 ± 0.187 (0.044)	1.104 ± 0.164 (0.038)
		MLE	1.627 ± 0.861 (1.134)	1.449 ± 0.529 (0.481)	1.376 ± 0.395 (0.298)	1.339 ± 0.331 (0.225)	1.315 ± 0.288 (0.182)	1.303 ± 0.263 (0.161)	1.292 ± 0.239 (0.142)	1.274 ± 0.208 (0.118)	1.259 ± 0.165 (0.094)	1.247 ± 0.138 (0.080)
β_T=3	50%	OLSE	0.938 ± 0.288 (0.087)	0.953 ± 0.236 (0.058)	0.966 ± 0.208 (0.044)	0.973 ± 0.186 (0.035)	0.982 ± 0.171 (0.030)	0.992 ± 0.160 (0.026)	0.996 ± 0.151 (0.023)	1.007 ± 0.139 (0.019)	1.022 ± 0.115 (0.014)	1.030 ± 0.102 (0.011)
		MLE	1.247 ± 0.344 (0.179)	1.185 ± 0.251 (0.097)	1.162 ± 0.208 (0.069)	1.140 ± 0.175 (0.051)	1.133 ± 0.158 (0.043)	1.127 ± 0.144 (0.037)	1.121 ± 0.132 (0.032)	1.115 ± 0.118 (0.027)	1.105 ± 0.094 (0.020)	1.100 ± 0.081 (0.016)
	60%	OLSE	0.957 ± 0.336 (0.115)	0.965 ± 0.262 (0.070)	0.977 ± 0.227 (0.052)	0.993 ± 0.206 (0.043)	0.995 ± 0.190 (0.036)	1.005 ± 0.182 (0.033)	1.013 ± 0.172 (0.030)	1.021 ± 0.153 (0.024)	1.039 ± 0.131 (0.019)	1.050 ± 0.113 (0.015)
		MLE	1.319 ± 0.429 (0.286)	1.237 ± 0.293 (0.142)	1.202 ± 0.237 (0.097)	1.189 ± 0.204 (0.077)	1.171 ± 0.181 (0.062)	1.165 ± 0.167 (0.055)	1.159 ± 0.156 (0.050)	1.148 ± 0.134 (0.040)	1.139 ± 0.107 (0.031)	1.134 ± 0.092 (0.026)
	70%	OLSE	0.980 ± 0.399 (0.160)	0.985 ± 0.298 (0.089)	0.999 ± 0.258 (0.067)	1.012 ± 0.238 (0.057)	1.020 ± 0.220 (0.049)	1.021 ± 0.202 (0.041)	1.032 ± 0.193 (0.038)	1.040 ± 0.174 (0.032)	1.060 ± 0.148 (0.025)	1.070 ± 0.131 (0.022)
		MLE	1.423 ± 0.560 (0.493)	1.315 ± 0.364 (0.232)	1.272 ± 0.291 (0.159)	1.248 ± 0.250 (0.124)	1.230 ± 0.223 (0.103)	1.214 ± 0.196 (0.085)	1.209 ± 0.183 (0.077)	1.195 ± 0.160 (0.064)	1.185 ± 0.126 (0.050)	1.175 ± 0.107 (0.042)
	80%	OLSE	1.078 ± 0.625 (0.397)	1.038 ± 0.412 (0.171)	1.023 ± 0.317 (0.101)	1.041 ± 0.292 (0.087)	1.046 ± 0.269 (0.074)	1.054 ± 0.249 (0.065)	1.054 ± 0.230 (0.056)	1.062 ± 0.211 (0.048)	1.086 ± 0.180 (0.040)	1.099 ± 0.159 (0.035)
		MLE	1.628 ± 0.843 (1.105)	1.457 ± 0.549 (0.511)	1.371 ± 0.396 (0.294)	1.339 ± 0.336 (0.228)	1.314 ± 0.289 (0.182)	1.299 ± 0.261 (0.158)	1.281 ± 0.236 (0.135)	1.266 ± 0.204 (0.112)	1.248 ± 0.162 (0.088)	1.240 ± 0.135 (0.076)
β_T=5	50%	OLSE	0.920 ± 0.283 (0.086)	0.935 ± 0.231 (0.057)	0.946 ± 0.201 (0.043)	0.962 ± 0.184 (0.035)	0.970 ± 0.170 (0.030)	0.976 ± 0.160 (0.026)	0.982 ± 0.150 (0.023)	0.991 ± 0.137 (0.019)	1.005 ± 0.113 (0.013)	1.017 ± 0.100 (0.010)
		MLE	1.229 ± 0.342 (0.170)	1.168 ± 0.250 (0.090)	1.138 ± 0.202 (0.060)	1.126 ± 0.176 (0.047)	1.118 ± 0.157 (0.039)	1.108 ± 0.144 (0.032)	1.104 ± 0.133 (0.028)	1.096 ± 0.118 (0.023)	1.086 ± 0.094 (0.016)	1.082 ± 0.079 (0.013)
	60%	OLSE	0.932 ± 0.328 (0.112)	0.946 ± 0.252 (0.066)	0.960 ± 0.222 (0.051)	0.970 ± 0.201 (0.041)	0.982 ± 0.186 (0.035)	0.987 ± 0.175 (0.031)	0.993 ± 0.165 (0.027)	1.005 ± 0.150 (0.023)	1.022 ± 0.126 (0.016)	1.033 ± 0.111 (0.013)
		MLE	1.298 ± 0.419 (0.264)	1.218 ± 0.282 (0.127)	1.183 ± 0.235 (0.089)	1.164 ± 0.201 (0.067)	1.154 ± 0.178 (0.055)	1.143 ± 0.164 (0.047)	1.137 ± 0.150 (0.041)	1.129 ± 0.132 (0.034)	1.117 ± 0.105 (0.025)	1.112 ± 0.090 (0.021)
	70%	OLSE	0.969 ± 0.413 (0.172)	0.959 ± 0.296 (0.089)	0.973 ± 0.253 (0.065)	0.983 ± 0.230 (0.053)	0.997 ± 0.211 (0.045)	1.008 ± 0.200 (0.040)	1.008 ± 0.189 (0.036)	1.022 ± 0.171 (0.030)	1.040 ± 0.142 (0.022)	1.055 ± 0.127 (0.019)
		MLE	1.425 ± 0.582 (0.520)	1.297 ± 0.368 (0.224)	1.248 ± 0.284 (0.142)	1.220 ± 0.245 (0.108)	1.207 ± 0.215 (0.089)	1.197 ± 0.197 (0.078)	1.185 ± 0.180 (0.066)	1.174 ± 0.156 (0.055)	1.157 ± 0.123 (0.040)	1.153 ± 0.105 (0.035)
	80%	OLSE	1.057 ± 0.639 (0.411)	0.998 ± 0.402 (0.161)	1.007 ± 0.326 (0.106)	1.005 ± 0.274 (0.075)	1.019 ± 0.257 (0.066)	1.027 ± 0.240 (0.059)	1.030 ± 0.224 (0.051)	1.049 ± 0.206 (0.045)	1.069 ± 0.175 (0.035)	1.078 ± 0.152 (0.029)
		MLE	1.662 ± 0.906 (1.260)	1.438 ± 0.545 (0.489)	1.366 ± 0.408 (0.301)	1.315 ± 0.324 (0.204)	1.294 ± 0.289 (0.170)	1.277 ± 0.256 (0.143)	1.260 ± 0.234 (0.122)	1.250 ± 0.203 (0.103)	1.226 ± 0.160 (0.077)	1.211 ± 0.133 (0.062)
β_T=8	50%	OLSE	0.910 ± 0.290 (0.092)	0.922 ± 0.229 (0.059)	0.937 ± 0.203 (0.045)	0.946 ± 0.183 (0.036)	0.953 ± 0.168 (0.030)	0.961 ± 0.158 (0.026)	0.970 ± 0.150 (0.023)	0.976 ± 0.137 (0.019)	0.992 ± 0.113 (0.013)	1.003 ± 0.099 (0.010)
		MLE	1.219 ± 0.347 (0.169)	1.151 ± 0.245 (0.083)	1.128 ± 0.203 (0.058)	1.109 ± 0.174 (0.042)	1.098 ± 0.156 (0.034)	1.091 ± 0.141 (0.028)	1.087 ± 0.131 (0.025)	1.078 ± 0.116 (0.020)	1.071 ± 0.093 (0.014)	1.066 ± 0.080 (0.011)
	60%	OLSE	0.919 ± 0.328 (0.114)	0.927 ± 0.256 (0.071)	0.942 ± 0.224 (0.053)	0.954 ± 0.202 (0.043)	0.965 ± 0.190 (0.037)	0.970 ± 0.173 (0.031)	0.978 ± 0.162 (0.027)	0.986 ± 0.148 (0.022)	1.001 ± 0.124 (0.015)	1.015 ± 0.110 (0.012)
		MLE	1.285 ± 0.425 (0.262)	1.200 ± 0.295 (0.127)	1.168 ± 0.236 (0.084)	1.145 ± 0.203 (0.062)	1.133 ± 0.182 (0.051)	1.122 ± 0.163 (0.041)	1.118 ± 0.148 (0.036)	1.106 ± 0.130 (0.028)	1.094 ± 0.106 (0.020)	1.091 ± 0.090 (0.016)
	70%	OLSE	0.948 ± 0.421 (0.180)	0.946 ± 0.306 (0.096)	0.953 ± 0.258 (0.069)	0.964 ± 0.230 (0.054)	0.972 ± 0.212 (0.046)	0.982 ± 0.198 (0.040)	0.984 ± 0.185 (0.034)	0.998 ± 0.169 (0.029)	1.017 ± 0.140 (0.020)	1.029 ± 0.125 (0.016)
		MLE	1.411 ± 0.598 (0.526)	1.283 ± 0.375 (0.221)	1.229 ± 0.295 (0.139)	1.203 ± 0.247 (0.102)	1.183 ± 0.217 (0.081)	1.171 ± 0.197 (0.068)	1.156 ± 0.178 (0.056)	1.147 ± 0.156 (0.046)	1.131 ± 0.122 (0.032)	1.123 ± 0.104 (0.026)
	80%	OLSE	1.020 ± 0.618 (0.382)	0.973 ± 0.405 (0.165)	0.978 ± 0.325 (0.106)	0.980 ± 0.275 (0.076)	0.989 ± 0.255 (0.065)	0.998 ± 0.238 (0.057)	1.005 ± 0.221 (0.049)	1.017 ± 0.201 (0.041)	1.044 ± 0.171 (0.031)	1.050 ± 0.152 (0.026)
		MLE	1.654 ± 0.924 (1.282)	1.428 ± 0.563 (0.500)	1.349 ± 0.420 (0.298)	1.296 ± 0.327 (0.195)	1.263 ± 0.285 (0.150)	1.250 ± 0.258 (0.129)	1.236 ± 0.233 (0.110)	1.216 ± 0.202 (0.088)	1.196 ± 0.159 (0.064)	1.181 ± 0.133 (0.050)

Table 3-6: Simulation results of $\hat{\alpha}$ for multiply censored data, generated by OLSE and MLE, at different n, β_T and c (part I – low censoring levels): the values of $E(\hat{\alpha}) \pm S(\hat{\alpha})$ and $MSE(\hat{\alpha})$ (in parentheses.

			n					
		Method	**20**	**30**	**50**	**100**	**150**	**200**
β_T=0.5	**10%**	OLSE	1.446 ± 0.752 (0.764)	1.372 ± 0.567 (0.460)	1.303 ± 0.419 (0.267)	1.250 ± 0.279 (0.140)	1.226 ± 0.222 (0.100)	1.217 ± 0.190 (0.083)
		MLE	1.238 ± 0.587 (0.401)	1.215 ± 0.465 (0.262)	1.191 ± 0.356 (0.163)	1.178 ± 0.247 (0.093)	1.171 ± 0.199 (0.069)	1.171 ± 0.174 (0.060)
	20%	OLSE	1.756 ± 0.972 (1.516)	1.659 ± 0.729 (0.966)	1.576 ± 0.530 (0.613)	1.502 ± 0.349 (0.374)	1.473 ± 0.274 (0.299)	1.458 ± 0.236 (0.265)
		MLE	1.457 ± 0.688 (0.683)	1.429 ± 0.549 (0.485)	1.409 ± 0.418 (0.342)	1.392 ± 0.291 (0.239)	1.387 ± 0.235 (0.205)	1.384 ± 0.204 (0.189)
	30%	OLSE	2.194 ± 1.402 (3.390)	2.076 ± 1.017 (2.191)	1.959 ± 0.720 (1.438)	1.856 ± 0.457 (0.942)	1.817 ± 0.360 (0.796)	1.793 ± 0.307 (0.723)
		MLE	1.744 ± 0.822 (1.230)	1.725 ± 0.664 (0.967)	1.701 ± 0.507 (0.748)	1.681 ± 0.349 (0.585)	1.678 ± 0.285 (0.540)	1.674 ± 0.246 (0.515)
	40%	OLSE	2.899 ± 2.548 (10.10)	2.720 ± 2.339 (8.430)	2.534 ± 1.155 (3.686)	2.385 ± 0.699 (2.406)	2.320 ± 0.506 (1.997)	2.288 ± 0.425 (1.841)
		MLE	2.165 ± 1.028 (2.416)	2.138 ± 0.830 (1.983)	2.110 ± 0.628 (1.627)	2.094 ± 0.441 (1.392)	2.086 ± 0.358 (1.307)	2.087 ± 0.308 (1.275)
β_T=1	**10%**	OLSE	1.146 ± 0.283 (0.10)	1.127 ± 0.229 (0.069)	1.109 ± 0.174 (0.042)	1.091 ± 0.122 (0.023)	1.085 ± 0.099 (0.017)	1.081 ± 0.085 (0.014)
		MLE	1.067 ± 0.250 (0.067)	1.065 ± 0.206 (0.047)	1.063 ± 0.159 (0.029)	1.062 ± 0.113 (0.016)	1.061 ± 0.092 (0.012)	1.061 ± 0.080 (0.010)
	20%	OLSE	1.233 ± 0.32 (0)	1.213 ± 0.255 (0.11)	1.195 ± 0.19 (0.1)	1.174 ± 0.136 (0.049)	1.165 ± 0.109 (0.039)	1.161 ± 0.094 (0.035)
		MLE	1.134 ± 0.269 (0.090)	1.134 ± 0.220 (0.066)	1.135 ± 0.170 (0.047)	1.135 ± 0.121 (0.033)	1.134 ± 0.099 (0.028)	1.135 ± 0.085 (0.026)
	30%	OLSE	1.352 ± 0.374 (0.264)	1.328 ± 0.299 (0.197)	1.302 ± 0.225 (0.142)	1.277 ± 0.154 (0.101)	1.267 ± 0.125 (0.087)	1.260 ± 0.107 (0.079)
		MLE	1.223 ± 0.295 (0.137)	1.222 ± 0.240 (0.107)	1.224 ± 0.187 (0.085)	1.223 ± 0.132 (0.067)	1.223 ± 0.108 (0.061)	1.223 ± 0.093 (0.058)
	40%	OLSE	1.496 ± 0.451 (0.450)	1.467 ± 0.359 (0.347)	1.440 ± 0.270 (0.266)	1.410 ± 0.183 (0.202)	1.396 ± 0.147 (0.178)	1.389 ± 0.126 (0.168)
		MLE	1.330 ± 0.330 (0.218)	1.330 ± 0.267 (0.180)	1.333 ± 0.206 (0.153)	1.334 ± 0.146 (0.133)	1.335 ± 0.119 (0.126)	1.335 ± 0.103 (0.123)
β_T=2	**10%**	OLSE	1.052 ± 0.131 (0.020)	1.046 ± 0.107 (0.014)	1.040 ± 0.082 (0.008)	1.033 ± 0.058 (0.004)	1.031 ± 0.047 (0.003)	1.029 ± 0.041 (0.003)
		MLE	1.017 ± 0.122 (0.015)	1.018 ± 0.100 (0.010)	1.020 ± 0.077 (0.006)	1.019 ± 0.055 (0.003)	1.021 ± 0.044 (0.002)	1.021 ± 0.039 (0.002)
	20%	OLSE	1.081 ± 0.141 (0.026)	1.074 ± 0.114 (0.019)	1.068 ± 0.088 (0.012)	1.061 ± 0.062 (0.007)	1.056 ± 0.050 (0.006)	1.056 ± 0.043 (0.005)
		MLE	1.039 ± 0.128 (0.018)	1.042 ± 0.105 (0.013)	1.043 ± 0.081 (0.008)	1.044 ± 0.057 (0.005)	1.044 ± 0.047 (0.004)	1.045 ± 0.040 (0.004)
	30%	OLSE	1.115 ± 0.150 (0.036)	1.109 ± 0.123 (0.027)	1.101 ± 0.095 (0.019)	1.093 ± 0.066 (0.013)	1.089 ± 0.054 (0.011)	1.087 ± 0.047 (0.010)
		MLE	1.067 ± 0.134 (0.023)	1.069 ± 0.111 (0.017)	1.071 ± 0.085 (0.012)	1.073 ± 0.060 (0.009)	1.073 ± 0.050 (0.008)	1.073 ± 0.043 (0.007)
	40%	OLSE	1.157 ± 0.167 (0.052)	1.150 ± 0.136 (0.041)	1.142 ± 0.105 (0.031)	1.133 ± 0.073 (0.023)	1.128 ± 0.059 (0.020)	1.126 ± 0.051 (0.019)
		MLE	1.099 ± 0.143 (0.030)	1.103 ± 0.116 (0.024)	1.105 ± 0.090 (0.019)	1.107 ± 0.064 (0.016)	1.108 ± 0.052 (0.014)	1.108 ± 0.045 (0.014)
β_T=3	**10%**	OLSE	1.029 ± 0.086 (0.008)	1.026 ± 0.070 (0.006)	1.022 ± 0.054 (0.003)	1.018 ± 0.038 (0.002)	1.017 ± 0.031 (0.001)	1.016 ± 0.027 (0.001)
		MLE	1.006 ± 0.081 (0.007)	1.008 ± 0.067 (0.004)	1.009 ± 0.051 (0.003)	1.009 ± 0.036 (0.001)	1.010 ± 0.030 (0.001)	1.010 ± 0.026 (0.001)
	20%	OLSE	1.043 ± 0.091 (0.010)	1.041 ± 0.074 (0.007)	1.036 ± 0.057 (0.005)	1.032 ± 0.040 (0.003)	1.030 ± 0.033 (0.002)	1.029 ± 0.028 (0.002)
		MLE	1.018 ± 0.085 (0.008)	1.020 ± 0.069 (0.005)	1.021 ± 0.054 (0.003)	1.022 ± 0.038 (0.002)	1.022 ± 0.031 (0.001)	1.023 ± 0.027 (0.001)
	30%	OLSE	1.061 ± 0.096 (0.013)	1.058 ± 0.080 (0.010)	1.053 ± 0.061 (0.007)	1.049 ± 0.043 (0.004)	1.046 ± 0.035 (0.003)	1.045 ± 0.030 (0.003)
		MLE	1.032 ± 0.088 (0.009)	1.034 ± 0.073 (0.006)	1.035 ± 0.056 (0.004)	1.037 ± 0.040 (0.003)	1.037 ± 0.032 (0.002)	1.037 ± 0.028 (0.002)
	40%	OLSE	1.082 ± 0.105 (0.018)	1.079 ± 0.086 (0.014)	1.075 ± 0.066 (0.010)	1.069 ± 0.047 (0.007)	1.066 ± 0.038 (0.006)	1.065 ± 0.033 (0.005)
		MLE	1.047 ± 0.094 (0.011)	1.051 ± 0.077 (0.008)	1.053 ± 0.059 (0.006)	1.054 ± 0.042 (0.005)	1.055 ± 0.034 (0.004)	1.055 ± 0.030 (0.004)
β_T=5	**10%**	OLSE	1.015 ± 0.051 (0.003)	1.013 ± 0.042 (0.002)	1.011 ± 0.032 (0.001)	1.009 ± 0.023 (0.001)	1.008 ± 0.019 (0.000)	1.007 ± 0.016 (0.000)
		MLE	1.002 ± 0.049 (0.002)	1.002 ± 0.040 (0.002)	1.003 ± 0.031 (0.001)	1.004 ± 0.022 (0.000)	1.004 ± 0.018 (0.000)	1.004 ± 0.015 (0.000)
	20%	OLSE	1.020 ± 0.054 (0.003)	1.019 ± 0.044 (0.002)	1.017 ± 0.034 (0.001)	1.014 ± 0.024 (0.001)	1.013 ± 0.020 (0.001)	1.013 ± 0.017 (0.000)
		MLE	1.006 ± 0.051 (0.003)	1.007 ± 0.042 (0.002)	1.008 ± 0.032 (0.001)	1.009 ± 0.023 (0.001)	1.009 ± 0.019 (0.000)	1.009 ± 0.016 (0.000)
	30%	OLSE	1.028 ± 0.057 (0.004)	1.026 ± 0.047 (0.003)	1.024 ± 0.036 (0.002)	1.021 ± 0.025 (0.001)	1.020 ± 0.021 (0.001)	1.019 ± 0.018 (0.001)
		MLE	1.011 ± 0.054 (0.003)	1.013 ± 0.044 (0.002)	1.014 ± 0.034 (0.001)	1.015 ± 0.024 (0.001)	1.015 ± 0.019 (0.001)	1.015 ± 0.017 (0.001)
	40%	OLSE	1.037 ± 0.061 (0.005)	1.035 ± 0.050 (0.004)	1.032 ± 0.039 (0.003)	1.030 ± 0.027 (0.002)	1.029 ± 0.022 (0.001)	1.028 ± 0.019 (0.001)
		MLE	1.018 ± 0.057 (0.004)	1.020 ± 0.047 (0.003)	1.021 ± 0.036 (0.002)	1.022 ± 0.025 (0.001)	1.023 ± 0.021 (0.001)	1.023 ± 0.018 (0.001)
β_T=8	**10%**	OLSE	1.008 ± 0.032 (0.001)	1.007 ± 0.026 (0.001)	1.005 ± 0.020 (0.000)	1.004 ± 0.014 (0.000)	1.004 ± 0.012 (0.000)	1.003 ± 0.010 (0.000)
		MLE	1.000 ± 0.031 (0.001)	1.000 ± 0.025 (0.001)	1.001 ± 0.019 (0.000)	1.001 ± 0.014 (0.000)	1.002 ± 0.011 (0.000)	1.002 ± 0.010 (0.000)
	20%	OLSE	1.010 ± 0.033 (0.001)	1.009 ± 0.028 (0.001)	1.008 ± 0.021 (0.001)	1.007 ± 0.015 (0.000)	1.006 ± 0.012 (0.000)	1.006 ± 0.011 (0.000)
		MLE	1.001 ± 0.032 (0.001)	1.002 ± 0.026 (0.001)	1.003 ± 0.020 (0.000)	1.003 ± 0.014 (0.000)	1.004 ± 0.012 (0.000)	1.004 ± 0.010 (0.000)
	30%	OLSE	1.014 ± 0.036 (0.001)	1.012 ± 0.029 (0.001)	1.011 ± 0.022 (0.001)	1.010 ± 0.016 (0.000)	1.009 ± 0.013 (0.000)	1.009 ± 0.011 (0.000)
		MLE	1.004 ± 0.034 (0.001)	1.005 ± 0.028 (0.001)	1.006 ± 0.021 (0.000)	1.006 ± 0.015 (0.000)	1.006 ± 0.012 (0.000)	1.006 ± 0.011 (0.000)
	40%	OLSE	1.018 ± 0.038 (0.002)	1.017 ± 0.031 (0.001)	1.015 ± 0.024 (0.001)	1.014 ± 0.017 (0.000)	1.013 ± 0.014 (0.000)	1.013 ± 0.012 (0.000)
		MLE	1.007 ± 0.036 (0.001)	1.008 ± 0.030 (0.001)	1.009 ± 0.023 (0.001)	1.009 ± 0.016 (0.000)	1.010 ± 0.013 (0.000)	1.010 ± 0.011 (0.000)

Table 3-7: Simulation results of $\hat{\alpha}$ for multiply censored data, generated by OLSE and MLE, at different n, β_T and c (part II – high censoring levels): the values of $E(\hat{\alpha}) \pm S(\hat{\alpha})$ and $MSE(\hat{\alpha})$ (in parentheses).

		Method	n = 20	30	50	100	150	200
β_T=0.5	50%	OLSE	453.59 ± 87553 (7665782938)	13.18 ± 124.7 (15694)	9.739 ± 33.32 (1187)	8.086 ± 9.395 (138.49)	7.574 ± 4.807 (66.33)	7.284 ± 3.150 (49.42)
		MLE	6.325 ± 5.202 (55.42)	6.034 ± 3.532 (37.81)	5.833 ± 2.408 (29.16)	5.700 ± 1.567 (24.55)	5.662 ± 1.261 (23.32)	5.647 ± 1.081 (22.77)
	60%	OLSE	38376 ± 5877178 (34542694582961)	591.5 ± 50231 (2523516876)	56.16 ± 1816.8 (3303913)	22.94 ± 547.2 (299856)	17.12 ± 32.17 (1295)	16.02 ± 21.96 (707.9)
		MLE	20.45 ± 163 (26815)	13.81 ± 30.33 (1084)	11.72 ± 9.273 (200.9)	10.80 ± 4.659 (117.8)	10.53 ± 3.473 (102.8)	10.42 ± 2.911 (97.22)
	70%	OLSE	4.214 ± 14.61 (223.87)	4.075 ± 70.24 (4943)	3.478 ± 1.970 (10.02)	3.230 ± 1.114 (6.213)	3.120 ± 0.816 (5.159)	3.065 ± 0.669 (4.712)
		MLE	2.798 ± 1.387 (5.156)	2.754 ± 1.092 (4.269)	2.739 ± 0.838 (3.728)	2.717 ± 0.583 (3.288)	2.706 ± 0.478 (3.140)	2.703 ± 0.409 (3.067)
	80%	OLSE	7.089 ± 45.34 (2093)	6.098 ± 23.75 (589.94)	5.300 ± 6.151 (56.33)	4.734 ± 2.256 (19.03)	4.538 ± 1.556 (14.94)	4.434 ± 1.279 (13.42)
		MLE	3.912 ± 2.154 (13.12)	3.819 ± 1.671 (10.74)	3.779 ± 1.255 (9.298)	3.744 ± 0.864 (8.278)	3.721 ± 0.697 (7.887)	3.721 ± 0.603 (7.769)
β_T=1	50%	OLSE	2.456 ± 4.437 (21.81)	2.393 ± 1.600 (4.501)	2.343 ± 0.984 (2.771)	2.270 ± 0.549 (1.915)	2.233 ± 0.426 (1.700)	2.209 ± 0.350 (1.585)
		MLE	1.993 ± 0.647 (1.405)	1.993 ± 0.507 (1.243)	1.994 ± 0.383 (1.136)	2.000 ± 0.268 (1.071)	1.998 ± 0.219 (1.045)	2.000 ± 0.189 (1.036)
	60%	OLSE	3.726 ± 36.74 (1357)	3.382 ± 7.577 (63.09)	3.314 ± 24.20 (590.9)	3.087 ± 1.347 (6.173)	3.012 ± 0.984 (5.018)	2.972 ± 0.780 (4.498)
		MLE	2.685 ± 1.629 (5.493)	2.625 ± 1.014 (3.669)	2.592 ± 0.691 (3.013)	2.574 ± 0.456 (2.684)	2.566 ± 0.364 (2.585)	2.565 ± 0.312 (2.546)
	70%	OLSE	1.689 ± 0.581 (0.813)	1.660 ± 0.472 (0.659)	1.627 ± 0.351 (0.517)	1.588 ± 0.232 (0.400)	1.572 ± 0.184 (0.361)	1.562 ± 0.158 (0.340)
		MLE	1.469 ± 0.373 (0.360)	1.474 ± 0.306 (0.319)	1.478 ± 0.238 (0.285)	1.480 ± 0.167 (0.258)	1.481 ± 0.137 (0.250)	1.482 ± 0.119 (0.246)
	80%	OLSE	1.974 ± 0.934 (1.822)	1.939 ± 0.682 (1.345)	1.900 ± 0.512 (1.072)	1.847 ± 0.326 (0.824)	1.828 ± 0.262 (0.754)	1.810 ± 0.218 (0.703)
		MLE	1.668 ± 0.451 (0.650)	1.676 ± 0.372 (0.595)	1.679 ± 0.286 (0.543)	1.683 ± 0.200 (0.507)	1.687 ± 0.164 (0.499)	1.686 ± 0.142 (0.491)
β_T=2	50%	OLSE	1.211 ± 0.192 (0.081)	1.203 ± 0.156 (0.066)	1.195 ± 0.120 (0.053)	1.184 ± 0.083 (0.041)	1.179 ± 0.067 (0.037)	1.177 ± 0.058 (0.035)
		MLE	1.141 ± 0.155 (0.044)	1.144 ± 0.127 (0.037)	1.149 ± 0.099 (0.032)	1.151 ± 0.069 (0.028)	1.152 ± 0.056 (0.026)	1.153 ± 0.048 (0.026)
	60%	OLSE	1.280 ± 0.230 (0.131)	1.275 ± 0.189 (0.111)	1.266 ± 0.145 (0.092)	1.254 ± 0.101 (0.075)	1.247 ± 0.082 (0.068)	1.245 ± 0.069 (0.065)
		MLE	1.194 ± 0.173 (0.068)	1.201 ± 0.141 (0.060)	1.206 ± 0.109 (0.054)	1.210 ± 0.077 (0.050)	1.210 ± 0.064 (0.048)	1.212 ± 0.054 (0.048)
	70%	OLSE	1.375 ± 0.308 (0.235)	1.374 ± 0.245 (0.200)	1.367 ± 0.188 (0.170)	1.356 ± 0.133 (0.145)	1.348 ± 0.106 (0.132)	1.343 ± 0.092 (0.126)
		MLE	1.272 ± 0.205 (0.116)	1.279 ± 0.167 (0.106)	1.287 ± 0.128 (0.098)	1.293 ± 0.090 (0.094)	1.294 ± 0.073 (0.092)	1.295 ± 0.064 (0.091)
	80%	OLSE	1.512 ± 0.496 (0.507)	1.527 ± 0.392 (0.432)	1.531 ± 0.312 (0.380)	1.524 ± 0.211 (0.319)	1.517 ± 0.165 (0.295)	1.509 ± 0.141 (0.279)
		MLE	1.403 ± 0.290 (0.246)	1.411 ± 0.228 (0.220)	1.419 ± 0.172 (0.205)	1.426 ± 0.119 (0.196)	1.429 ± 0.097 (0.193)	1.430 ± 0.084 (0.192)
β_T=3	50%	OLSE	1.189 ± 0.162 (0.062)	1.190 ± 0.133 (0.054)	1.188 ± 0.104 (0.046)	1.180 ± 0.073 (0.038)	1.176 ± 0.059 (0.035)	1.174 ± 0.051 (0.033)
		MLE	1.132 ± 0.127 (0.034)	1.139 ± 0.103 (0.030)	1.145 ± 0.079 (0.027)	1.148 ± 0.056 (0.025)	1.150 ± 0.045 (0.024)	1.151 ± 0.040 (0.024)
	60%	OLSE	1.253 ± 0.226 (0.115)	1.262 ± 0.182 (0.102)	1.265 ± 0.144 (0.091)	1.261 ± 0.102 (0.079)	1.258 ± 0.083 (0.073)	1.254 ± 0.071 (0.069)
		MLE	1.193 ± 0.161 (0.063)	1.201 ± 0.129 (0.057)	1.208 ± 0.098 (0.053)	1.215 ± 0.068 (0.051)	1.216 ± 0.056 (0.050)	1.217 ± 0.048 (0.050)
	70%	OLSE	1.109 ± 0.117 (0.025)	1.105 ± 0.095 (0.020)	1.101 ± 0.073 (0.016)	1.095 ± 0.052 (0.012)	1.092 ± 0.042 (0.010)	1.090 ± 0.036 (0.009)
		MLE	1.068 ± 0.101 (0.015)	1.072 ± 0.083 (0.012)	1.075 ± 0.064 (0.010)	1.077 ± 0.045 (0.008)	1.078 ± 0.037 (0.007)	1.078 ± 0.032 (0.007)
	80%	OLSE	1.142 ± 0.134 (0.038)	1.141 ± 0.109 (0.032)	1.136 ± 0.084 (0.026)	1.130 ± 0.059 (0.020)	1.127 ± 0.048 (0.018)	1.124 ± 0.042 (0.017)
		MLE	1.094 ± 0.111 (0.021)	1.100 ± 0.090 (0.018)	1.104 ± 0.070 (0.016)	1.107 ± 0.049 (0.014)	1.108 ± 0.040 (0.013)	1.108 ± 0.035 (0.013)
β_T=5	50%	OLSE	1.082 ± 0.087 (0.014)	1.082 ± 0.071 (0.012)	1.081 ± 0.056 (0.010)	1.078 ± 0.039 (0.008)	1.076 ± 0.032 (0.007)	1.075 ± 0.027 (0.006)
		MLE	1.052 ± 0.074 (0.008)	1.056 ± 0.061 (0.007)	1.060 ± 0.047 (0.006)	1.063 ± 0.033 (0.005)	1.064 ± 0.027 (0.005)	1.064 ± 0.023 (0.005)
	60%	OLSE	1.109 ± 0.113 (0.025)	1.115 ± 0.091 (0.022)	1.114 ± 0.070 (0.018)	1.113 ± 0.051 (0.015)	1.111 ± 0.041 (0.014)	1.110 ± 0.035 (0.013)
		MLE	1.076 ± 0.090 (0.014)	1.084 ± 0.073 (0.012)	1.087 ± 0.055 (0.011)	1.092 ± 0.039 (0.010)	1.093 ± 0.032 (0.010)	1.094 ± 0.028 (0.010)
	70%	OLSE	1.048 ± 0.067 (0.007)	1.046 ± 0.055 (0.005)	1.044 ± 0.042 (0.004)	1.041 ± 0.030 (0.003)	1.040 ± 0.024 (0.002)	1.039 ± 0.021 (0.002)
		MLE	1.026 ± 0.061 (0.004)	1.029 ± 0.050 (0.003)	1.031 ± 0.038 (0.002)	1.032 ± 0.027 (0.002)	1.033 ± 0.022 (0.002)	1.033 ± 0.019 (0.001)
	80%	OLSE	1.063 ± 0.075 (0.010)	1.061 ± 0.061 (0.008)	1.059 ± 0.047 (0.006)	1.056 ± 0.033 (0.004)	1.054 ± 0.027 (0.004)	1.054 ± 0.023 (0.003)
		MLE	1.037 ± 0.066 (0.006)	1.040 ± 0.054 (0.005)	1.043 ± 0.042 (0.004)	1.045 ± 0.029 (0.003)	1.045 ± 0.024 (0.003)	1.046 ± 0.021 (0.003)
β_T=8	50%	OLSE	1.039 ± 0.053 (0.004)	1.039 ± 0.043 (0.003)	1.038 ± 0.034 (0.003)	1.036 ± 0.024 (0.002)	1.035 ± 0.019 (0.002)	1.034 ± 0.017 (0.001)
		MLE	1.021 ± 0.048 (0.003)	1.024 ± 0.039 (0.002)	1.026 ± 0.030 (0.002)	1.028 ± 0.021 (0.001)	1.029 ± 0.017 (0.001)	1.029 ± 0.015 (0.001)
	60%	OLSE	1.051 ± 0.065 (0.007)	1.054 ± 0.053 (0.006)	1.054 ± 0.041 (0.005)	1.053 ± 0.029 (0.004)	1.052 ± 0.024 (0.003)	1.051 ± 0.020 (0.003)
		MLE	1.031 ± 0.056 (0.004)	1.035 ± 0.045 (0.003)	1.039 ± 0.035 (0.003)	1.042 ± 0.024 (0.002)	1.042 ± 0.020 (0.002)	1.043 ± 0.017 (0.002)
	70%	OLSE	1.023 ± 0.041 (0.002)	1.022 ± 0.034 (0.002)	1.020 ± 0.026 (0.001)	1.019 ± 0.018 (0.001)	1.018 ± 0.015 (0.001)	1.018 ± 0.013 (0.000)
		MLE	1.010 ± 0.039 (0.002)	1.011 ± 0.032 (0.001)	1.013 ± 0.024 (0.001)	1.014 ± 0.017 (0.000)	1.014 ± 0.014 (0.000)	1.014 ± 0.012 (0.000)
	80%	OLSE	1.029 ± 0.046 (0.003)	1.029 ± 0.037 (0.002)	1.027 ± 0.029 (0.002)	1.026 ± 0.021 (0.001)	1.025 ± 0.017 (0.001)	1.024 ± 0.014 (0.001)
		MLE	1.014 ± 0.043 (0.002)	1.017 ± 0.034 (0.001)	1.018 ± 0.027 (0.001)	1.020 ± 0.019 (0.001)	1.020 ± 0.015 (0.001)	1.020 ± 0.013 (0.001)

3.4 Summary

In this chapter, the properties of the OLS estimators of the Weibull parameters were examined through both analytical methods and Monte Carlo simulation experiments. The important findings are summarized as follows.

Theoretical Findings

1) The OLS estimators of α and β are biased and may not have minimum variance among all linear estimators.

2) A sensible selection for y_i is to use the expected values of the order statistics of the reduced variable $Z = (X - \ln\alpha)/(1/\beta)$. The values can be calculated by Equation (3-8), and the corresponding estimates for failure probability F can be calculated by the relationship $Y = \ln[-\ln(1-F)]$.

3) The Weibull distribution, denoted by $Wei(\alpha, \beta)$, is related to the extreme value distribution, denoted by $Exm(\mu, \sigma)$, with $\mu = \ln\alpha$ and $\sigma = 1/\beta$. The transformation to the extreme-value distribution, which is of location-scale type, helps to ease the analytical deductions. The BLUEs for μ and σ are well-established; however, as the relationships $\mu = \ln\alpha$ and $\sigma = 1/\beta$ are both nonlinear, the BLUEs for α and β cannot be easily obtained.

4) Same as the MLE of α and β, the LSE of α and β have two pivotal functions $\hat{\beta}/\beta$ and $\hat{\beta}\ln(\hat{\alpha}/\alpha)$ whose distributions are independent of α and β.

Simulation Findings

1) For complete data, the relative bias of the OLS estimated β is typically within 5% and is inconsistent with the sample size. The relative bias reaches smallest between $n = 6$ and $n = 7$. During $10 < n < 30$, the relative bias is like a constant and remains at around 4%. The standard deviation and MSE are typically much larger than the bias, indicating that OLSE has a low efficiency. OLSE outperforms MLE for estimating β for small samples, while MLE performs better for estimating α, especially when β_T is small (although both estimators of α have large bias when β_T is small).

2) For multiply censored data, the bias of the OLS estimated β is inconsistent with either n or c. The bias reaches smallest at different combinations of n and c, e.g., $c = 30\%$ and $n = 150 - 200$, $c = 40\%$ and $n = 100 - 150$, $c = 50\%$ and $n = 80 - 100$, $c = 60\%$ and $n = 50 - 60$, and $c = 70\%$ and $n = 20 - 30$. The bias is significant for small samples with very low censoring levels ($c \le 20\%$) or large samples with very high censoring levels ($c \ge 70\%$). For estimating α, the results are generally unsatisfactory at $\beta_T = 0.5$. MLE always outperforms OLSE for estimating α. OLSE outperforms MLE for estimating β as long as the censoring level is not very low, i.e., $c > 20\%$.

3) For both complete data and censored data, the standard deviation and MSE of $\hat{\alpha}$ and $\hat{\beta}$ of the OLSE generally decrease with the increase of sample size. However, the bias is inconsistent with the sample size. This means for the OLSE method, the increase of sample size may not generate better estimates.

Chapter 4

Modifications on the OLSE Method

This chapter presents some modifications on the OLSE method with the aim of providing better estimates for the Weibull parameters. The importance of using LSE together with WPP is emphasized. Discussions on the plotting positions in the cases of complete data and censored data, respectively, are presented. The expected plotting positions or its approximations are recommended. A comparison between two LSE methods, LS *Y* on *X* and LS *X* on *Y*, is presented. The simulation results show that the two methods outperform each other at different conditions.

4.1 Introduction

In the previous chapter, the properties of the OLS estimators have been carefully examined via both analytical method and experimental method. It was found that the OLS estimators of the Weibull parameters, especially for the shape parameter, are biased and have large variance for certain sample sizes or censoring levels. There are many possibilities to improve the OLSE method, as can be seen in the following of this thesis. This chapter presents a few small modifications without change the least squares regression technique used in the OLSE method.

In the following, Section 4.2 describes the advantages of using OLSE with WPP instead of using it merely as a simple analytical method. Section 4.3 examines the selection of the *Y*-axis plotting positions on parameter estimation. It will show that the Bernard estimator and the Herd-Johnson estimator used in OLSE for complete data and censored data, respectively, can be replaced by other estimators to achieve better parameter estimators under certain circumstances. Section 4.4 presents another modification on OLSE, which is to reverse the dependent variable and the independent variable in the least squares regression. In OLSE, $X = \ln T$ is the independent variable and $Y = \ln[-\ln(1-F(t))]$ is the dependent variable. This is in good agreement with WPP which plots t along the *X*-axis and F along the *Y*-axis. However, from the viewpoint of a controlled experiment design, it is more appropriate to set $X = \ln[-\ln(1-F(t))]$ as the independent variable and $Y = \ln T$ as the dependent variable because t is the measured values or output from the experiment and the values of F are estimated by some non-parametric estimators which are independent of t. The comparisons between the two methods are presented in details. Some of the work presented in this chapter has been published in Zhang et al. (2005, 2007).

4.2 Modification 1: Always Use LSE with WPP

Parameter estimation methods for the Weibull distribution are commonly divided into two groups: graphical methods and analytical methods. In Chapter 2, the WPP method and the LSE method are described as two types of estimation methods: WPP is a graphical estimation method, and LSE belongs to the group of analytical estimation methods. In practice, however, these two methods are frequently used together. Theoretically, LSE and WPP are both based on the linearized Weibull CDF, i.e., Equation (2-1). By combining LSE with WPP, it is basically to use the least squares regression technique to generate the straight line on the probability plot instead of by eye. The advantages of the combination over the two individual methods are obvious: 1) compared to WPP, it avoids the subjectivity by using eye-fitting so as to improve the estimation efficiency; and 2) compared to LSE, it gives a graphical presentation which can serve as model validation and outlier identification, in addition to parameter estimation.

Application Procedure of LSE with WPP

For a random Weibull sample denoted by $t_1, t_2, \cdots, t_i, \cdots, t_n$ $(i = 1,2,\cdots,n)$, and in the case of censored data, let $t_{f,1}, t_{f,2}, \cdots, t_{f,j}, \cdots, t_{f,r}$ $(j = 1,2,\cdots,r)$ denote the failures in this sample, the following procedure shows how to apply LSE with WPP to estimate the Weibull parameters:

Step 1: Rank the failure times, i.e., t_i (for complete sample) or $t_{f,j}$ (for censored sample), from smallest to largest.

Step 2: Calculate the estimated values of failure probability, i.e., $\hat{F}_{(i)}$ (for complete sample) or $\hat{F}_{f,(j)}$ (for censored sample), at each failure data point.

Step 3: Generate the Weibull probability plot: plot $t_{(i)}$ vs. $\hat{F}_{(i)}$ (for complete sample), or $t_{f,(j)}$ vs. $\hat{F}_{f,(j)}$ (for censored data) on Weibull probability paper. If the Weibull distribution fits, the data points should appear to be on a straight line.

Step 4: Generate a straight line for the data points on WPP using the least squares regression technique.

Step 5: Estimate α and β with Equation (2-12).

If the Weibull probability paper is not available, ***Step 3*** can be modified as plotting $\ln t_{(i)}$ vs. $\ln\left[-\ln(1-\hat{F}_{(i)})\right]$ (for complete sample), or $\ln t_{f,(j)}$ vs.

$\ln\left[-\ln(1-\hat{F}_{f,(j)})\right]$ (for censored data) on linear-linear paper. This can be carried out in spreadsheet like MS Excel.

4.3 Modification 2: Estimation of *F*(*t*) (Plotting Positions)

WPP, LSE and other linear regression estimation methods discussed in this thesis all require the estimated value of failure probability F at each failure time. Weibull researchers have agreed the importance of the estimation of F, commonly known as the *Y*-axis plotting positions, on parameter estimation. Much work has been done on this topic, as briefly described in Section 1.3.1. Among the existing estimators of F, most are simple non-parametric estimators that can be used for complete data. The estimation of F in the case of multiply censored data is less discussed.

The definition of $F(t)$ is the probability that a random variable T takes on a value less than or equal to a real number, e.g., t_0. For the Weibull distribution, we have

$$F(t_0) = P(T \le t_0) = 1 - \exp\left[-\left(\frac{t_0}{\alpha}\right)^{\beta}\right] \tag{4-1}$$

From Equation (4-1), the value of $F(t_0)$ depends on t_0, α and β. t_0 is a failure observation which is known, but α and β are unknown parameters of the Weibull distribution, hence the value of $F(t_0)$ can only be estimated. The estimation of $F(t_0)$ is frequently called the determination of *Y*-axis plotting positions for the Weibull probability plot. This is not a unique problem for the Weibull probability plotting, for example, some discussions on the similar problem can be found for the normal probability plotting, see, e.g., Looney & Gulledge (1985). As is well known, $\hat{F} = (i - 1/4)/(n + 3/8)$ is used for the normal distribution.

The estimation of $F(t_0)$, or the selection of the *Y*-axis plotting positions for WPP, is such a hot topic that a large portion of literature about WPP and LSE examined this problem. Different estimators of $F(t_0)$ have been proposed, to be applied to complete data and censored data, respectively. Most of the existing estimators are expressed by the functions of order number and sample size. Unlike the situation for the normal probability plotting, where $\hat{F} = (i - 1/4)/(n + 3/8)$ is used as a standard formula for calculating $\hat{F}$ and there is rarely an alternative, currently there is no fixed method for the estimation of the Weibull F, especially for censored data. The discussion is still ongoing.

In the following, Section 4.3.1 summarizes the common methods for calculating $\hat{F}$ for complete Weibull samples into different groups and the results are presented in a table for easy reference. The related work is described and the research gaps are pointed out. Similar work is presented in Section 4.3.2, in the case of censored data. Then, Section 4.3.3 and Section 4.3.4 present the Monte Carlo experiment study of the different methods for calculating $\hat{F}$, in the cases of complete data and censored data, respectively, and the results will suggest which method is best under certain circumstances. The simulation results are presented in figures for the convenience of comparison.

4.3.1 Estimation of *F* for Complete Data

More than eight non-parametric estimators for calculating $\hat{F}_{(i)}$ have been proposed and compared in the literature. The general form of these estimators can be expressed by

$$\hat{F}_{(i)} = \frac{i - c_1}{n + c_2} \tag{4-2}$$

where c_1, c_2 are two real numbers.

Table 4-1 gives a summary of these estimators. As can be seen from the table, the existing non-parametric estimators are divided into five categories:

1) **Mean rank plotting positions:** the Weibull estimator (Weibull, 1939).
2) **Median rank plotting positions:** the Bernard estimator (Bernard & Bosi-Levenbach, 1953) and the Filliben estimator (Filliben, 1975).
3) **Expected plotting positions:** the Ross estimator (Ross, 1994b) and the Drap-Kos estimator (Drapella & Kosznik, 1999).
4) **'Optimal' plotting positions:** the estimators vary with sample sizes (Wu & Lu, 2004; Tiryakioglu & Hudak, 2007).
5) **Others:** the Hazen estimator (Hazen, 1930), the Blom estimator (Blom, 1958), etc.

Table 4-1: Summary of the common estimators for *F* applied to complete data.

General Form	Estimators of *F*			Category
$\hat{F} = (i - c_0)/(n + 1 - 2c_0)$ where c_0 is a constant.	Weibull estimator (Weibull, 1939)	$i/(n+1)$	$c_0 = 0$	Mean rank plotting position
	Hazen estimator (Hazen, 1930)	$(i-0.5)/n$	$c_0 = 0.5$	Other
	Bernard estimator (Bernard & Bosi-Levenbach, 1953)	$(i-0.3)/(n+0.4)$	$c_0 = 0.3$	Median rank plotting position
	Filliben estimator (Filliben, 1975)	$(i-0.3175)/(n+0.365)$	$c_0 = 0.3175$	
$\hat{F} = (i - c_1)/(n + c_2)$ where c_1, c_2 are two constants and $c_2 \neq 1 + 2c_1$.	Blom estimator (Blom, 1958)	$(i-0.5)/(n+0.25)$	$c_1 = 0.5, c_2 = 0.25$	Other
	Ross estimator (Ross, 1994)	$(i-0.44)/(n+0.25)$	$c_1 = 0.44, c_2 = 0.25$	Expected plotting position
	Drap-Kos estimator (Drapella & Kosznik, 1999) $\hat{y}_i = E(Y_{(i)}) = i\binom{n}{i} \cdot \sum_{k=0}^{i-1}\left\{(-1)^k \binom{i-1}{k} \times \frac{-\gamma - \ln(n-i+k+1)}{n-i+k+1}\right\}$ and $\hat{F}_i = 1 - \exp(-\exp(y_i))$			
$\hat{F} = (i - c_1)/(n + c_2)$ where c_1, c_2 are varying with the sample size.	$(i - 0.68)/(n + 0.82)$ at $n = 10$ $(i - 0.62)/(n + 0.92)$ at $n = 15$ $(i - 0.68)/(n + 0.1)$ at $n = 20$ … (Wu & Lu, 2004; Wu et al., 2006; Tiryakioglu & Hudak, 2007)			'Optimal' plotting position

The theoretical backgrounds of the Weibull estimator, the Bernard estimator and the Hazen estimator have been described in Section 2.2. The mean rank plotting positions and the median rank plotting positions are most frequently used. The estimators in these two categories satisfy the following general form,

$$\hat{F}_{(i)} = \frac{i - c_0}{n - 2c_0 + 1} \tag{4-3}$$

where c_0 is a real number. The Hazen estimator is a special case which also satisfies this equation with $c_0 = 0.5$.

Fothergill (1990), with Monte Carlo simulations, compared the LSE methods with the Bernard estimator, the Weibull estimator and the Hazen estimator on estimating Weibull parameters for samples of size 3 to 20. The author concluded that when the Bernard estimator is used, the LS estimators of α and β are nearly unbiased, while the Weibull estimator results in underestimated β and the Hazen estimator results in overestimated β. It was also showed that the Bernard estimator is a very good approximation to the exact median rank values. Cacciari & Montanari (1991) extended Fothergill's work and added the Blom estimator and the Filliben estimator in the comparison via Monte Carlo simulations. The authors concluded that the Bernard estimator and the Filliben estimator are clearly better than the Weibull estimator and the Blom estimator on parameter estimation and should be preferred for small samples. Their results also showed that when the Bernard or the Filliben estimator is used, the LS estimators of α and β are not consistent, i.e., the accuracy improves as the sample size increases; while when the Blom estimator is used, the LS estimators of α and β are consistent.

The methods in the third category can be found in the early literature such as Weibull (1967) and White (1969); however, it is not as popular as the Weibull or the Bernard estimator nowadays. Ross (1994b) examined the method and gave it the name expected plotting positions. The idea is to first calculate the expected values of $Y_{(i)}$, and then calculate the values for $\hat{F}_{(i)}$ by $\hat{F}_{(i)} = 1 - \exp\left[-\exp(E(Y_{(i)}))\right]$. Weibull (1967) said that the Bernard estimator, though generally acceptable, will be biased, and the correct plotting positions are calculated in this way. Section 3.2.2 and Section 3.2.3 have presented the analytical deduction on $E(Y_{(i)})$ and the theoretical justification on this plotting position. Drapella & Kosznik (1999) also suggested the calculation of $\hat{F}_{(i)}$ through $E(Y_{(i)})$, and the formulas are given by

Drap-Kos estimator

$$\hat{y}_i = E(Y_{(i)}) = i\binom{n}{i} \cdot \sum_{k=0}^{i-1}\left\{(-1)^k \binom{i-1}{k} \times \frac{-\gamma - \ln(n-i+k+1)}{n-i+k+1}\right\} \tag{4-4}$$

and

$$\hat{F}_{(i)} = 1 - \exp\left[-\exp(\hat{y}_i)\right] \tag{4-5}$$

Equation (4-4) is similar to Equation (3-8) but the deduction is not provided in Drapella & Kosznik (1999). With Monte Carlo simulations, the authors concluded that, with their formulas used in LSE, the bias of the LS estimators is greatly reduced, while the MSE of the estimators are slightly increased. Equation (4-4) has been cited many times in recent years, see, e.g., Xie et al. (2000), Yang & Xie (2003), Hung (2004) and Lu et al. (2004).

The disadvantage of the expected plotting positions is obvious, i.e., the complexity in calculating the values of $E(Y_{(i)})$, especially when the sample size is large. Ross (1994b) proposed a simple approximation formula which satisfies the general form $(i - c_1)/(n + c_2)$ for the expected plotting positions via numerical methods. The formula is given by

Ross estimator $$\hat{F}_{(i)} = \frac{i - 0.44}{n + 0.25} \quad (4\text{-}6)$$

Ross compared this estimator with the Bernard estimator and the Weibull estimator in view of plotting and parameter estimation, respectively. It was concluded that the new estimator, when used in LSE, outperforms the others and generates nearly unbiased LS shape parameter estimator. This simple approximation formula, unfortunately, has not received much attention.

The estimators for estimating F in the fourth category also belong to the simple form in Equation (4-2); however, the values of c_1 and c_2 are not fixed but depend on the sample size and are determined via the Monte Carlo method based on certain objectives which make the estimators 'optimal'. The objective used to determine c_1 and c_2 in the work of Wu & Lu (2004) is to maximize the probability that $\hat{\beta}/\beta$ fall into the interval $[0.9, 1.1]$, and the objective in Wu et al. (2006) is to minimize the bias of $\hat{\beta}$, i.e., to make $\hat{\beta}/\beta$ closest to 1. The values of c_1 and c_2 were determined for selected sample sizes and tabulated in the two papers. The authors concluded that there is no distinct relationship existing between the values of c_1, c_2 and the sample size. A similar research can also be found in Tiryakioglu & Hudak (2007). Obviously, this type of method has great limitations on applications, because one can not know the optimal values of c_1 and c_2 for those sample sizes that are not shown in the authors' work.

The discussions regarding the estimation of failure probabilities for the Weibull distribution have not received much agreement. The Bernard estimator is used in the OLSE method for complete data and it is probably the most recognized estimator, followed by the Hazen estimator and the Weibull estimator. The expected plotting positions have good theoretical backgrounds and were noticed by some researchers in the last decade. Ross' approximation formula for the expected plotting positions in Equation (4-6) may have a good potential for its simplicity and accuracy. It is carefully examined, together with other popular plotting positions for LSE, on parameter estimation via Monte Carlo simulations in Section 4.3.3. The fourth category of the plotting positions is not further discussed due to the application inconvenience.

4.3.2 Estimation of *F* for Censored Data

For a censored sample, LSE uses only the failure times to conduct regression analysis and WPP plots only failure data points. How to make use of the information provided by the part of censored data in a sample is the key problem in the LSE procedure and it will greatly affect the parameter estimation results. Obviously, ignoring censored data or treating them as failures will cause unreliable estimates because the information provided by censored data is lost or misused.

As a common practice, the influence of censoring is reflected in the estimation of F at each failure data point. Therefore, the estimation of F for censored data is more complicated and more important than that for complete data.

The literature on estimating F for censored data is not as much as that for complete data. Nelson (2004) described the WPP procedure including the calculation of $\hat{F}_{f,(j)}$ in the cases of different types of censored data. The Herd-Johnson estimator (Herd, 1960; Johnson, 1964) in Equation (2-6) is recommended for calculating $\hat{F}_{f,(j)}$. The theoretical background of the Herd-Johnson estimator has been presented in Section 2.2. Other methods have also been proposed; however, not as popular as the Herd-Johnson method. Table 4-2 summarizes the existing methods for calculating $\hat{F}_{f,(j)}$ for censored data. The references are listed and the characteristics of each method are pointed out.

As can be seen from the table, the existing methods on estimating F for censored data are divided into two categories:

1) **Without calculating the MFON:** the Kaplan-Meier (KM) estimator (Kaplan & Meier, 1958), the Herd-Johnson (HJ) estimator (Herd, 1960; Johnson, 1964) and the Zimmer estimator (Skinner et al., 2001; Hossain & Zimmer, 2003).

2) **First calculate the MFON, denoted by $m_{f,(j)}$, and then use $m_{f,(j)}$ in the Bernard estimator (or other non-parametric estimators like Hazen or Weibull) to calculate $\hat{F}_{f,(j)}$:** the modified Johnson (JM) method (Keats et al., 2000), the age sensitive method (ASM) of Hastings & Bartlett (1997), the exponential age sensitive method (EASM) of Campean (2000) and the refined rank regression method (RRRM) of Wang (2001, 2004).

In the following, the methods in both categories are briefly described.

Table 4-2: Summary of the common estimators for F applied to censored data.

General Form	Estimators of F or $MFON$	Category	Note
$\begin{cases} \hat{R}_{f,(j)} = \left(\dfrac{n-I_j+d_1}{n-I_j+d_2}\right)\cdot \hat{R}_{f,(j-1)} \\ \hat{F}_{f,(j)} = 1-\hat{R}_{f,(j)} \end{cases}$ where d_1, d_2 are two constants.	Kaplan-Meier (KM) estimator (Kaplan & Meier, 1958) $\begin{cases} \hat{R}_{f,(j)} = \left(\dfrac{n-I_j}{n+1-I_j}\right)\cdot \hat{R}_{f,(j-1)} \\ \hat{F}_{f,(j)} = 1-\hat{R}_{f,(j)} \end{cases}$	Calculate F without calculating the MFON.	1) Simplest and oldest 2) $\hat{F}_{f,(j)} = 1$ when $I_j = n$, which is unrealistic 3) Insensitive to the exact censoring time
	Herd-Johnson (HJ) estimator (Herd, 1960; Johnson, 1964) $\begin{cases} \hat{R}_{f,(j)} = \left(\dfrac{n+1-I_j}{n+2-I_j}\right)\cdot \hat{R}_{f,(j-1)} \\ \hat{F}_{f,(j)} = 1-\hat{R}_{f,(j)} \end{cases}$		1) Widely used 2) $\hat{F}_{fj} < 1$ when $I_j = n$ 3) Insensitive to the exact censoring time
	Zimmer estimator (Skinner et al., 2001; Hossain & Zimmer, 2003) $\begin{cases} \hat{R}_{f,(j)} = \left(\dfrac{n+0.5-I_j}{n+1.5-I_j}\right)\cdot \hat{R}_{f,(j-1)} \\ \hat{F}_{f,(j)} = 1-\hat{R}_{f,(j)} \end{cases}$		1) $\hat{F}_{fj} < 1$ when $I_j = n$ 2) Midpoint of KM and HJ 3) Insensitive to the exact censoring time
	Modified Johnson's method (JM) (Keats et al., 2000) $\begin{cases} \Delta_j = \dfrac{n+1-m_{f,(j-1)}}{n-I_j+2} = \dfrac{n+1-m_{f,(j-1)}}{I_j^{*}+1} \\ m_{f,(j)} = m_{f,(j-1)} + \Delta_j \\ \hat{F}_{f,(j)} = \dfrac{m_{f,(j)}-0.3}{n-0.4} \end{cases}$	First calculate the MFON, then use the Bernard estimator to calculate F.	1) Appears in many reliability textbooks 2) Under the Herd's (1960) assumption: censoring occur concurrently with a failure 3) Under equal probability assumption 4) Insensitive to the exact censoring time

Table 4-2 Continued.

General Form	Estimators of F or $MFON$	Category	Note
	Age sensitive method (ASM) (Hastings & Bartlett, 1997) $\begin{cases} \Delta_j = \dfrac{n+1-m_{f,(j-1)}}{I_j^*+1} - \dfrac{I_j^*}{I_j^*+1} \cdot \sum_{i\in\omega_i} \dfrac{\alpha_k}{I_j^*+\overline{\alpha}_k} \\ m_{f,(j)} = m_{f,(j-1)} + \Delta_j \\ \hat{F}_{f,(j)} = \dfrac{m_{f,(j)}-0.3}{n-0.4} \end{cases}$, where $\alpha_k = (t_{c,(k)} - t_{f,(j-1)}) / (t_{f,(j)} - t_{f,(j-1)})$ and $\overline{\alpha}_k = 1-\alpha_k$		1) Removed the Herd's (1960) assumption 2) Partially relaxed the equal probability assumption 3) Sensitive to the exact censoring time
	Exponential age sensitive method (EASM) (Campean, 2000) $\begin{cases} \Delta_j = \sum_{i\in\omega_j}(t_{f,(j)} - t_{c,(k)}) \cdot \hat{h}_j + (I_j^* - j) \cdot \hat{h}_j \cdot (t_{f,(j)} - t_{f,(j-1)}) + 1 \\ m_{f,(j)} = m_{f,(j-1)} + \Delta_j \\ \hat{F}_{f,(j)} = \dfrac{m_{f,(j)}-0.3}{n-0.4} \end{cases}$	First calculate the MFON, then use the Bernard estimator to calculate F.	1) Removed the Herd's (1960) assumption 2) Partially relaxed the equal probability assumption 3) Sensitive to the exact censoring time
	Refined rank regression method (RRR) (Wang, 2001, 2004) $\begin{cases} m_{f,(j)} = \sum_{i=1}^{I_i} \left[\delta_i + (1-\delta_i) \cdot \dfrac{F(t_{f,(j)}) - F(t_i)}{1-F(t_i)} \right] \\ \hat{F}_{f,(j)} = \dfrac{m_{f,(j)}-0.3}{n-0.4} \end{cases}$ where $\delta_i = \begin{cases} 0 & \text{if } t_i \text{ is a censoring time} \\ 1 & \text{if } t_i \text{ is a failure time} \end{cases}$		1) Removed the Herd's (1960) assumption 2) Removed the equal probability assumption 3) Under the assumption that censoring time and failure time are from the same distribution 4) Iterative procedure

The KM estimator is the oldest non-parametric estimator for F applied to censored data. Its formula is given by

KM estimator
$$\begin{cases} \hat{R}_{f,(j)} = \left(\dfrac{n - I_j}{n + 1 - I_j} \right) \cdot \hat{R}_{f,(j-1)} \\ \hat{F}_{f,(j)} = 1 - \hat{R}_{f,(j)} \end{cases} \tag{4-7}$$

where the definition of I_j is given in Section 2.2, i.e., the event number of the j^{th} failure in the sample. From Equation (4-7), if the last observation in a sample is a failure, we have $I_j = n$, and hence the failure probability is always equal to 1 for this failure data point. This is obviously unrealistic for censored data and it tends to underestimate the failures in the tail of the distribution; therefore, the KM estimator is not recommended.

The HJ estimator overcomes the shortcoming of the KM estimator and is widely used for censored data. The formula of the HJ estimator is given in Equation (2-6).

Besides the KM estimator and the HJ estimator, Skinner et al. (2001) and Hossain & Zimmer (2003) proposed a similar estimator, named the Zimmer estimator, which is expressed by

Zimmer estimator
$$\begin{cases} \hat{R}_{f,(j)} = \left(\dfrac{n + 0.5 - I_j}{n + 1.5 - I_j} \right) \cdot \hat{R}_{f,(j-1)} \\ \hat{F}_{f,(j)} = 1 - \hat{R}_{f,(j)} \end{cases} \tag{4-8}$$

The authors compared it with the HJ estimator on estimating the Weibull parameters in the cases of Type II censored samples and selected patterns of multiply censored samples via Monte Carlo simulations. It was concluded that in view of both bias and

MSE of the estimators, the HJ method is generally better than the Zimmer method for estimating β while the Zimmer method is better for estimating α.

The JM estimator belongs to the second category, but it has a close relationship with the HJ estimator. The formula of the JM estimator is

JM estimator

$$\begin{cases} \Delta_j = \dfrac{n+1-m_{f,(j-1)}}{n-I_j+2} = \dfrac{n+1-m_{f,(j-1)}}{I_j^*+1} \\ m_{f,(j)} = m_{f,(j-1)} + \Delta_j \\ \hat{F}_{f,(j)} = \dfrac{m_{f,(j)}-0.3}{n-0.4} \end{cases} \tag{4-9}$$

where I_j^* is the reverse rank of I_j, i.e., $I_j^* = n+1-I_j$. Δ_j is the increment between $m_{f,(j-1)}$ and $m_{f,(j)}$. At $j=1$, $m_{f,(0)}=0$. If the first observation is a failure, $m_{f,(1)}=1$.

In Equation (4-9), $m_{f,(j)}$ is used in the Bernard estimator for calculating $\hat{F}_{f,(j)}$, but if it is used in the Weibull estimator, i.e., $\hat{F}_{f,(j)} = m_{f,(j)}/(n+1)$, the JM estimator and the HJ estimator become the same.

The methods of KM, HJ, Zimmer and JM are insensitive to the exact censoring times. The JM estimator and the HJ estimator are derived based on two assumptions: one assumption is suggested by Herd (1960) that assumes a censoring event occurs concurrently with a failure event, and the second assumption assumes that a censored unit, if allowed to continue in service, has equal probability to fail in any of the subsequent intervals of two consecutive failure times.

Hastings & Bartlett (1997) proposed a so-called age sensitive method to take the censoring times into account for calculating $\hat{F}_{f,(j)}$. The method uses the proportion of the interval length between event times to estimate the probability that a censored unit would fail in the current interval and in any of the subsequent intervals of two consecutive failure times. The exact censoring times are used in the calculation. For a multiply censored sample, plot the failure times and censoring times along the time axis. Assume that the k^{th} censoring time $t_{c,(k)}$ lies in the interval of two consecutive failure times $[t_{f,(j-1)}, t_{f,(j)})$. Let $\alpha_k = (t_{c,(k)} - t_{f,(j-1)})/(t_{f,(j)} - t_{f,(j-1)})$ and $\overline{\alpha}_k = 1 - \alpha_k$, the formula of the Hastings & Bartlett's ASM estimator is given by

ASM estimator

$$\begin{cases} \Delta_j = \dfrac{n+1-m_{f,(j-1)}}{I_j^*+1} - \dfrac{I_j^*}{I_j^*+1} \cdot \sum\limits_{l\in\omega_j} \dfrac{\alpha_k}{I_j^*+\overline{\alpha}_k} \\ m_{f,(j)} = m_{f,(j-1)} + \Delta_j \\ \hat{F}_{f,(j)} = \dfrac{m_{f,(j)}-0.3}{n-0.4} \end{cases} \tag{4-10}$$

where ω_j denotes the collection of all k that satisfy $t_{c,(k)} \in [t_{f,(j-1)}, t_{f,(j)})$.

Hastings & Bartlett (1997) compared the ASM with the JM method using a numerical example and showed that their method is sensitive to the censoring time while JM is not. However, the average performance of the method over the JM method was not examined (Campean, 2000). Theoretically, compared to the JM method, the ASM removed the Herd's assumption and relaxed, or partially removed, the equal probability assumption.

Campean (2000) proposed another age sensitive method called exponential age sensitive method. The method is based on the assumption that the hazard rate, denoted by h_j, for each time interval of two consecutive failures is constant within the interval. The author stated that this constant failure rate assumption offers a more robust criterion for age sensitiveness than the simple proportional distance used by Hastings & Bartlett (1997).

The formula of the EASM is

$$\textbf{\textit{EASM estimator}} \begin{cases} \Delta_j = \sum_{l \in w_j} (t_{f,(j)} - t_{c,(k)}) \cdot \hat{h}_j + (I_j^* - j) \cdot \hat{h}_j \cdot (t_{f,(j)} - t_{f,(j-i)}) + 1 \\ m_{f,(j)} = m_{f,(j-1)} + \Delta_j \\ \hat{F}_{f,(j)} = \dfrac{m_{f,(j)} - 0.3}{n - 0.4} \end{cases} \tag{4-11}$$

Campean (2000) provided two methods for estimating h_j, one is the maximum likelihood estimation and the other is called the Bayesian smoothed piecewise estimation method which, according to the author, can offer a smooth and robust estimation for the hazard rate. A simulation study was conducted to compare the JM, ASM and EASM (with the Bayesian smoothed piecewise estimator for hazard rate) on the estimation of Weibull parameters. The results clearly showed the advantages of the EASM at the censoring level of 12.5%. It is also surprised to see that the performance of all methods improves with the increase of censoring level.

All the methods described above are non-parametric methods, i.e., the calculation of $\hat{F}_{f,(j)}$ does not involve the two Weibull parameters α and β. Wang (2001, 2004) proposed a parametric approach to calculate the MFON and $\hat{F}_{f,(j)}$, which is also an age sensitive method. Wang's formula for MFON is based on the Weibull CDF and the definition of conditional probability. The method is named refined rank regression method by the author. The formula is given by

$$\textbf{\textit{RRRM estimator}} \begin{cases} m_{f,(j)} = \sum_{i=1}^{I_j} \left[\delta_i + (1 - \delta_i) \cdot \dfrac{F(t_{f,(j)}) - F(t_i)}{1 - F(t_i)} \right] \\ \hat{F}_{f,(j)} = \dfrac{m_{f,(j)} - 0.3}{n - 0.4} \end{cases} \tag{4-12}$$

where δ_i is the censoring indicator, and

$$\delta_i = \begin{cases} 0 & \text{if } t_i \text{ is a censoring time} \\ 1 & \text{if } t_i \text{ is a failure time} \end{cases}$$

From Equation (4-12), the calculation of $m_{f,(j)}$ is not straightforward because $F(t_{f,(j)})$ and $F(t_i)$ are unknown as α and β are unknown. To solve this problem, Wang proposed an iterative procedure which combines the calculation of $m_{f,(j)}$ and the parameter estimation for α and β. The procedure needs initial estimates of α and β that can be obtained from the LSE method with the JM estimator. The application procedure of the RRRM, according to Wang (2004), is as follows.

Step 0: Find distribution parameters using standard LS method as the initial estimates.

Step 1: With the initial parameter estimates, calculate $m_{f,(j)}$ and $\hat{F}_{f,(j)}$ using Equation (4-12).

Step 2: Update the estimates of the distribution parameters through a revised LS regression using the new values of $\hat{F}_{f,(j)}$.

Step 3: Return and repeat the process from step 1 until an acceptable convergence is reached on the parameter estimates.

An advantage of the RRRM is that it removes both the Herd's assumption and the equal probability assumption. However, the calculation is obviously more complicated compared to other methods. With Monte Carlo simulations, the author compared the RRRM and the JM method on the goodness-of-fit in view of plotting. It was concluded that the RRRM generates a better fit for the Weibull distributed data.

In summary, the KM estimator has a big problem and should not be used. The HJ estimator or the JM estimator is probably the most widely used estimator. If the Weibull estimator is used in JM, i.e., $\hat{F}_{f,(j)} = m_{f,(j)} / (n+1)$, the JM estimator and the HJ estimator are same. The ASM and EASM both remove the Herd (1960) assumption, i.e., censoring occur concurrently with a failure event, and use the exact censoring time in calculating MFON. In theory, the EASM makes some improvements over the ASM; however, the calculation becomes much more complicated. The RRRM is the only parametric method and it has a good statistical foundation. The application, however, needs iterations and hence is inconvenient without the aid of a computer. Although computation is usually not a big problem nowadays, there are still situations where the trade-off between computation complexity and estimation accuracy is of interest.

In the following, selected methods for calculating $\hat{F}_{(i)}$ and $\hat{F}_{f,(j)}$ are compared via Monte Carlo simulations and the results will provide suggestions on their usage.

4.3.3 Simulation Study on Plotting Positions for Complete Data

A Monte Carlo experiment was carried out to find the best plotting position, among those described in Section 4.3.1, used in the LSE method to estimate the two Weibull parameters for complete data. Table 4-3 lists the experiment factors and their values. Five plotting positions were examined in this experiment including the Bernard estimator in Equation (2-4), the Weibull estimator in Equation (2-2), the Hazen estimator in Equation (2-5), the Ross estimator in Equation (4-6) and the Drap-Kos estimator in Equation (4-4). The comparisons focus on the small to medium sized samples because it is known that OLSE performs not very well under such conditions.

Table 4-3: Setting of experiment factors. The experiment is to compare different plotting positions used in LSE for complete data on parameter estimation.

Factors	Values
α_T	1
β_T	1 (for $\hat{\beta}$) and 0.5, 1, 5 (for $\hat{\alpha}$)
n	3 – 30
M	10000
Methods	Bernard, Weibull, Hazen, Ross, Drap-Kos

For a randomly generated Weibull sample, all the five methods were used to calculate the values of y_i, and then these y_i were used in Equation (2-12) to generate the LS estimates of α and β. This procedure was repeated for 10000 times in each combination of α_T, β_T and n. Finally, the mean and MSE of $\hat{\alpha}$ and $\hat{\beta}$ for each method were calculated as the comparison criteria.

The comparison results are presented in figures instead of tables so that the performance of the methods can be easily compared. The results for $\hat{\beta}$ are presented in Figure 4-1 and Figure 4-2. The mean and MSE of the estimators are separately presented. Based on the first pivotal function $\hat{\beta}/\beta$, the results for $\hat{\beta}_{1,1}$ can represent the results for $\hat{\beta}$ given any β_T. The results for $\hat{\alpha}$ are presented in Figure 4-3 – Figure 4-8. Since $\hat{\alpha}/\alpha$ is not a pivotal function, different values of β_T (0.5, 1 and 5) were considered. The following conclusions can be observed.

<u>Simulation Results for Estimators of β (Figure 4-1 and Figure 4-2)</u>

1) ***Bias of $\hat{\beta}$ (refer to Figure 4-1)**:* Unfortunately, none of the methods always performs best at all sample sizes investigated. Also, none of them are unbiased. When the sample sizes are very small, say $n = 3, 4$, the Weibull estimator is the best; however, it is the worst among the five from $n = 5$

onwards. The Bernard estimator performs best during $n = 6 - 8$ where the bias is almost 0. From $n = 9$ onwards, the Ross estimator and the Drap-Kos estimator are the best ones. The Hazen estimator generates highly overestimated $\hat{\beta}$ when $n \le 10$, but it performs very close to the Ross estimator and the Drap-Kos estimator when $n \ge 20$. The Bernard estimator results in underestimated $\hat{\beta}$ when $n \ge 7$ and the bias is close to a constant. The bias of $\hat{\beta}$ generated by the Ross estimator and the Drap-Kos estimator almost disappears when $n \ge 12$.

2) ***MSE of $\hat{\beta}$ (refer to Figure 4-2)**:* Same as bias, none of the method has smallest MSE at all sample sizes investigated. But the MSE of $\hat{\beta}$ generated by the Hazen estimator is always largest among the five. When $n \le 10$, the MSE of $\hat{\beta}$ generated by the Weibull estimator is significantly smaller than that of the other estimators, especially at $n = 3, 4$. The Bernard estimator performs the second best, followed by the Ross estimator and the Drap-Kos estimator, and finally the Hazen estimator. When $n > 10$, however, the MSE of $\hat{\beta}$ generated by all the methods are close, and that of the Bernard estimator is slightly smaller than that of the others.

3) ***Comparison between the Ross estimator and the Drap-Kos estimator:*** In view of both bias and MSE, the two estimators perform closely for all the sample sizes examined. This result indicates that the Ross estimator is a good approximation for the exact expected plotting positions.

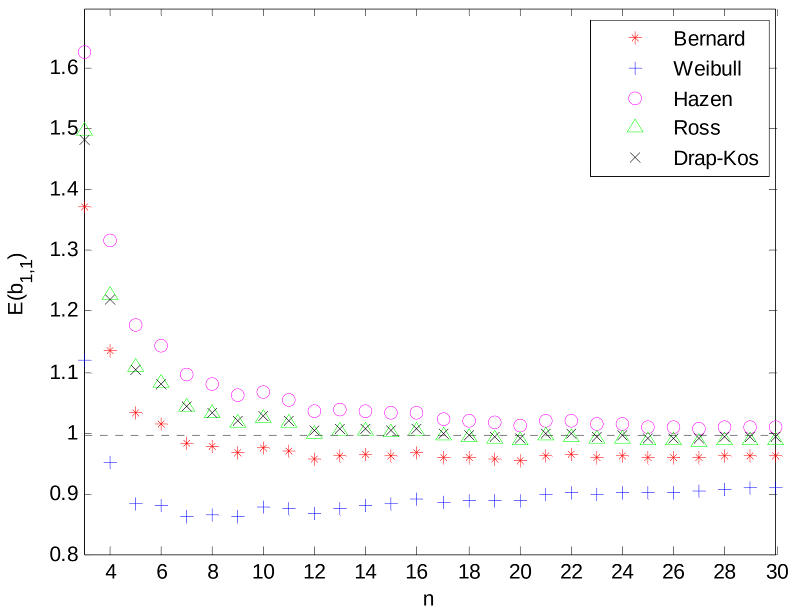


Figure 4-1: Comparison of the shape parameter estimators for complete data, obtained by LSE with different plotting positions used, at different n: the values of $E(\hat{\beta}_{1,1})$.

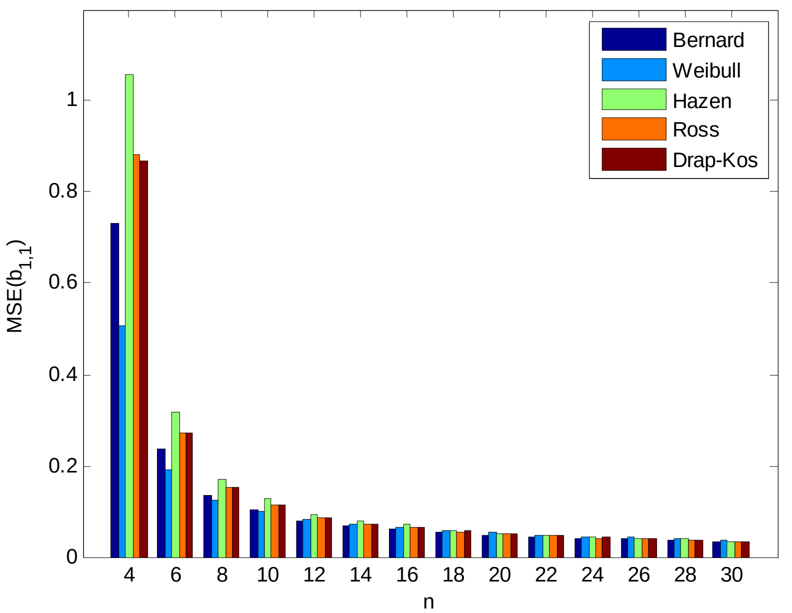


Figure 4-2: Comparison of the shape parameter estimators for complete data, obtained by LSE with different plotting positions used, at different n: the values of $MSE(\hat{\beta}_{1,1})$.

Simulation Results for Estimators of α (Figure 4-3 – Figure 4-8)

1) ***Bias of $\hat{\alpha}$ (refer to Figures 4-3, 4-5 and 4-7)***: Comparing the three figures, it can be seen that, although generally the bias of $\hat{\alpha}$ of all methods decreases as β_T increases, the trends of $\hat{\alpha}$ as a function of n vary with β_T for all the methods. The trends of $\hat{\alpha}$ of Hazen and Bernard at $\beta_T = 5$ are dramatically different from the trends of them at $\beta_T = 0.5$ and $\beta_T = 1$. The estimators of α of all methods are roughly consistent at $\beta_T = 0.5$, but inconsistent at $\beta_T = 1$ and $\beta_T = 5$. At $\beta_T = 0.5$ and $\beta_T = 1$, the Hazen estimator outperforms the others at all sample sizes investigated, followed by the Bernard estimator, and the Weibull estimator performs worst most of the time. All methods result in highly overestimated $\hat{\alpha}$. The bias of $\hat{\alpha}$ of all methods is larger than 10% at $\beta_T = 0.5$, and at $\beta_T = 1$, the bias of $\hat{\alpha}$ of Hazen is within 10% and typically within 5%. At $\beta_T = 5$, however, $\hat{\alpha}$ of Hazen is underestimated when $n \le 6$. At $n < 5$, Bernard becomes the best one. From $n = 5$ onwards, Hazen returns to the best, followed by Bernard, and then Ross, Drap-Kos and finally Weibull. The bias of $\hat{\alpha}$ of Hazen is typically 0.3% and that of Bernard is typically 0.6%.

2) ***MSE of $\hat{\alpha}$ (refer to Figures 4-4, 4-6 and 4-8)***: The difference in the MSE of $\hat{\alpha}$ of all methods decreases with the increase of n and β_T. The difference is significant only at $\beta_T = 0.5, 1$ and $n \le 10$. At all β_T, the MSE of $\hat{\alpha}$ of Hazen is smaller than that of the others, especially at $\beta_T = 0.5$ and $n \le 10$. The MSE of $\hat{\alpha}$ of Weibull and Drap-Kos are always the largest.

3) ***Both bias and MSE of $\hat{\alpha}$:*** Considering both bias and MSE, Hazen outperforms the others in most of the times except when $\beta_T = 5$ and $n < 5$ (Bernard has a smaller bias). Especially when $5 \le n \le 10$, the bias and MSE of $\hat{\alpha}$ of Hazen are significantly smaller than that of the other methods. On the other hand, Weibull is generally inferior to others in view of both bias and MSE of $\hat{\alpha}$.

4) ***Comparison between the Ross estimator and the Drap-Kos estimator:*** The difference between the two is larger for estimating α than for estimating β. The Ross estimator performs slightly better than Drap-Kos for estimating α in view of both bias and MSE of $\hat{\alpha}$.

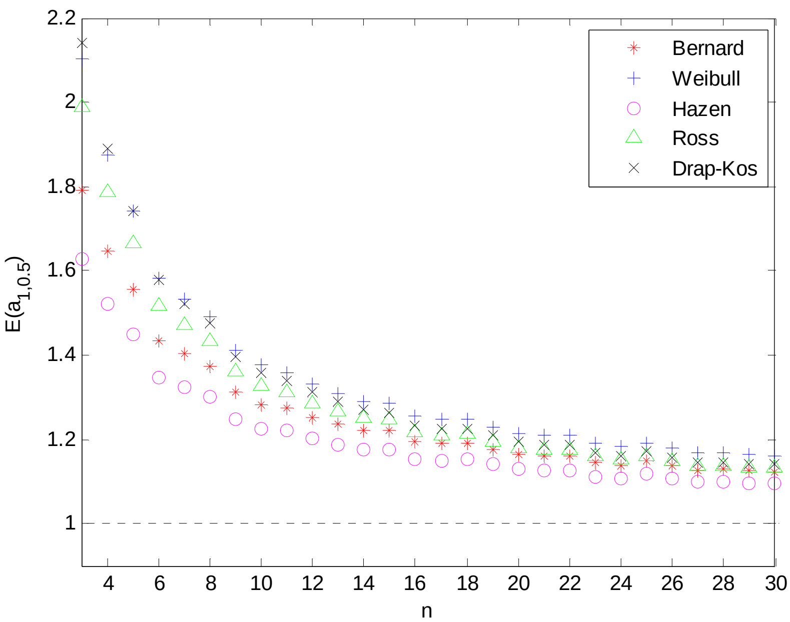


Figure 4-3: Comparison of the scale parameter estimators for complete data, obtained by LSE with different plotting positions used, at different *n* and β_T : the values of $E(\hat{\alpha}_{1,0.5})$.

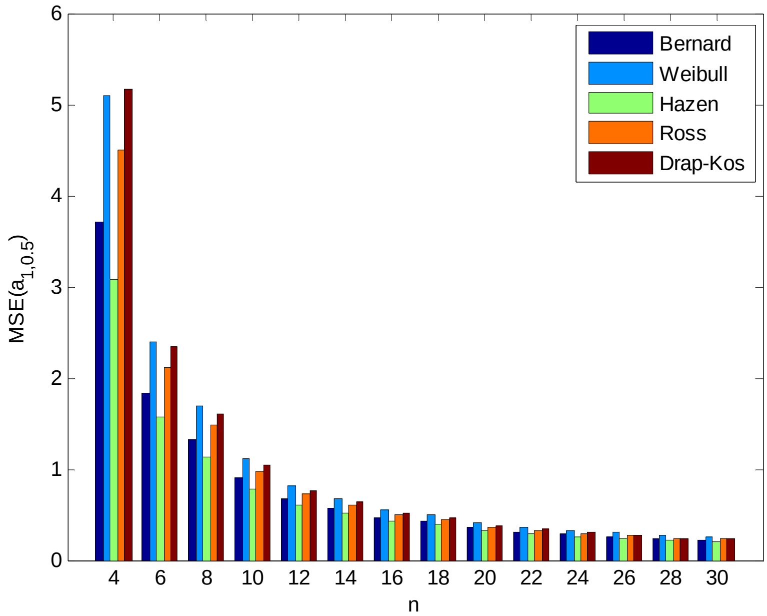


Figure 4-4: Comparison of the scale parameter estimators for complete data, obtained by LSE with different plotting positions used, at different *n* and β_T : the values of $MSE(\hat{\alpha}_{1,0.5})$.

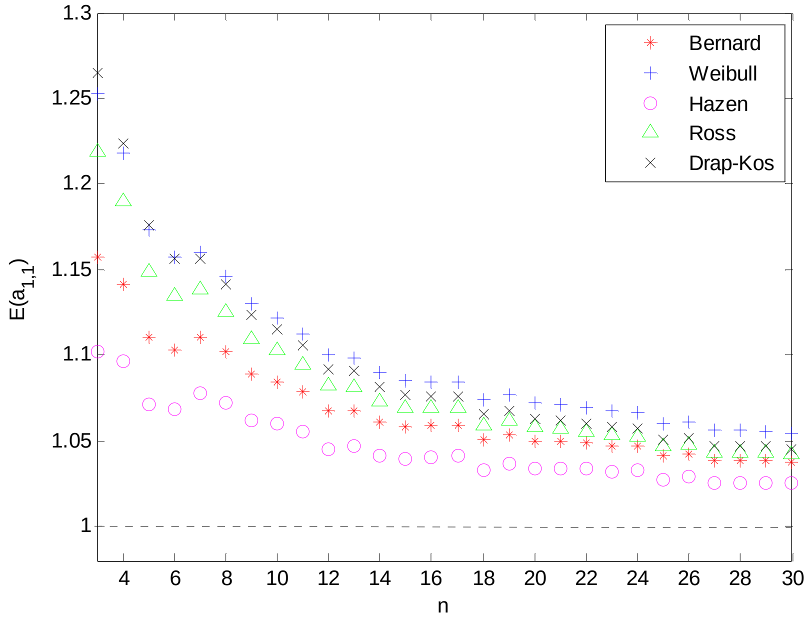


Figure 4-5: Comparison of the scale parameter estimators for complete data, obtained by LSE with different plotting positions used, at different *n* and β_T : the values of $E(\hat{\alpha}_{1,1})$.

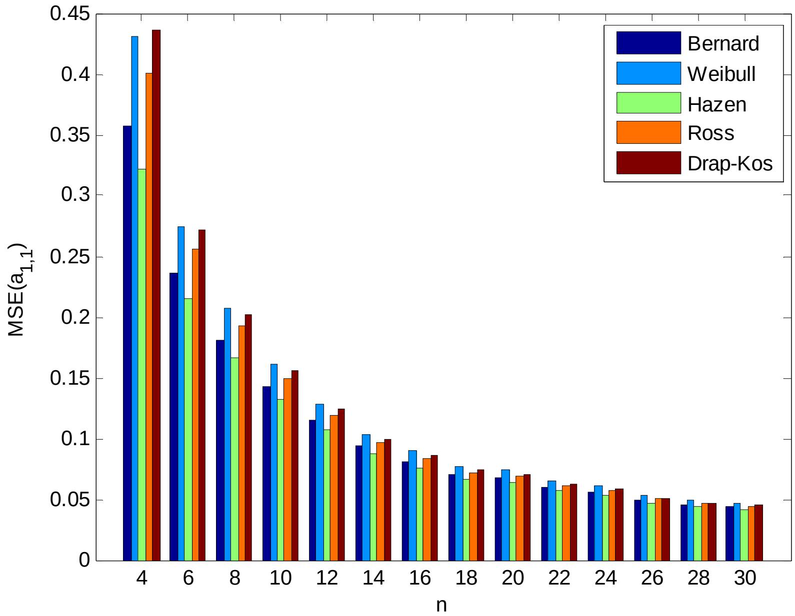


Figure 4-6: Comparison of the scale parameter estimators for complete data, obtained by LSE with different plotting positions used, at different *n* and β_T : the values of $MSE(\hat{\alpha}_{1,1})$.

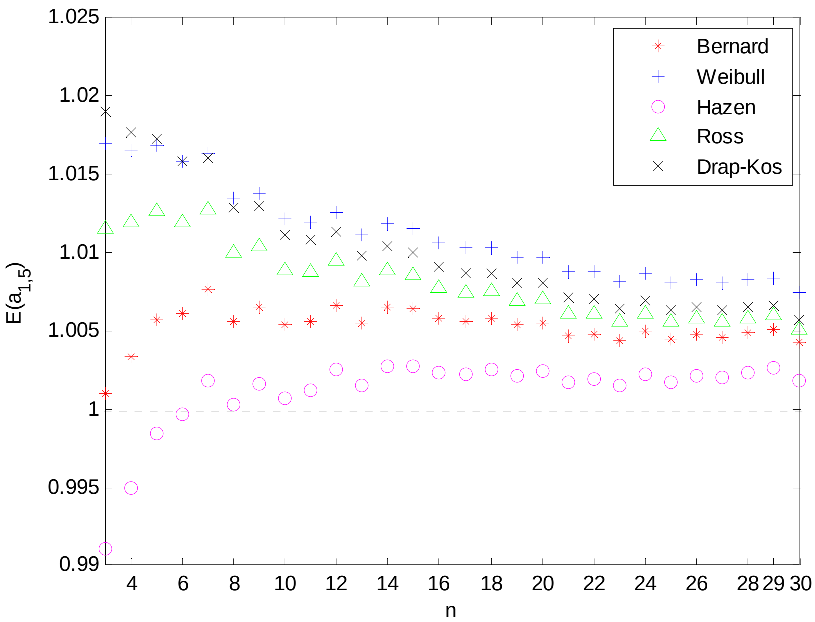


Figure 4-7: Comparison of the scale parameter estimators for complete data, obtained by LSE with different plotting positions used, at different *n* and β_T : the values of $E(\hat{\alpha}_{1,5})$.

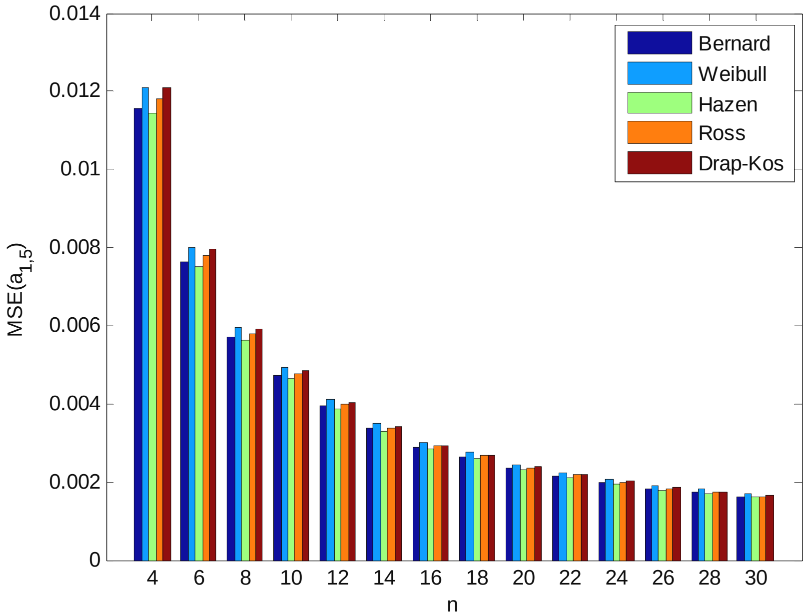


Figure 4-8: Comparison of the scale parameter estimators for complete data, obtained by LSE with different plotting positions used, at different *n* and β_T : the values of $MSE(\hat{\alpha}_{1,5})$.

4.3.4 Simulation Study on Plotting Positions for Censored Data

The objective of this Monte Carlo experiment is to find the best plotting position applied to multiply censored data among those descried in Section 4.3.2, used in the LSE method for estimating the Weibull parameters.

Table 4-4 lists the experiment factors and their values. The plotting positions examined in this experiment include the HJ estimator in Equation (2-6), the JM estimator in Equation (4-9), the ASM estimator in Equation (4-10) and the RRRM estimator in Equation (4-12). The EASM estimator is not considered because its computation complexity may greatly limit its application. The KM estimator and the Zimmer estimator are not considered because former work has shown that they are not clearly better than the HJ estimator. The mean and MSE of $\hat{\alpha}$ and $\hat{\beta}$, obtained by each method, were calculated as comparison criteria.

Table 4-4: Setting of experiment factors. The experiment is to compare different plotting positions used in LSE for censored data on parameter estimation.

Factors	**Values**
α_T	1
β_T	1 (for $\hat{\beta}$) and 0.5, 1, 5 (for $\hat{\alpha}$)
n	10, 20, 30, 50, 80, 100
c	10%, 30%, 50%, 70%
M	10000
Methods	HJ, JM, ASM, RRRM

The simulation results are presented in Figure 4-9 – Figure 4-24. The following conclusions can be observed.

<u>Simulation Results for Estimators of β (Figure 4-9 – Figure 4-16)</u>

1) ***General observations:*** The results at low censoring levels (10%, 30%) and high censoring levels (50%, 70%) are quite different. None of the methods outperforms the others at all combinations of the experiment factors in view of both bias and MSE of $\hat{\beta}$.

2) ***Bias of*** $\hat{\beta}$ ***(refer to Figure 4-9, Figure 4-11, Figure 4-13 and Figure 4-15)**:* The HJ estimated $\hat{\beta}$ presents different trends as a function of the sample size from the other three methods. At low censoring levels (10%, 30%), the bias of $\hat{\beta}$ of HJ is significantly larger than that of the other methods. JM, ASM and RRRM perform similarly at $c = 10\%$ (the bias is typically within 4%) and $c = 30\%$ (the bias is typically within 2%). The bias of the three methods is within 1% at the combinations of $c = 30\%$ and $30 \le n \le 80$. On the other hand, at high censoring levels (50%, 70%), however, the bias of $\hat{\beta}$ of HJ is the smallest in most cases. Especially at $c = 70\%$, the bias of $\hat{\beta}$ of HJ is within 1% at $n = 20, 30$ and within 4% at $n = 50, 80, 100$, which is

significantly smaller than that of the other methods. The difference among JM, ASM and RRRM at high censoring levels is large for small samples. When $n \leq 30$, JM performs best, followed by ASM, and RRRM performs worst; however, RRRM performs slightly better than JM and ASM for larger sample sizes, e.g., $n = 80, 100$.

3) ***MSE of $\hat{\beta}$ (refer to Figure 4-10, Figure 4-12, Figure 4-14 and Figure 4-16):*** HJ is the best at most conditions except that when $c = 10\%$ and $n \geq 20$, the MSE of $\hat{\beta}$ of HJ is slightly larger than that of the other methods. Among JM, ASM and RRRM, regardless the censoring levels, the MSE of $\hat{\beta}$ of JM is always smallest, followed by ASM and finally RRRM.

4) ***Both bias and MSE of $\hat{\beta}$***: HJ outperforms the others at high censoring levels (50%, 70%) in view of both bias and MSE. JM, ASM and RRRM are better for low censoring levels (10%, 30%) and the difference between them is small.

5) ***Consistency****:* At low censoring levels (10%, 30%), the bias of $\hat{\beta}$ of HJ decreases as the sample size increases and this is not true for the other methods. At high censoring levels (50%, 70%), all the estimators are inconsistent with the sample size. Moreover, all the estimators are inconsistent with the censoring level.

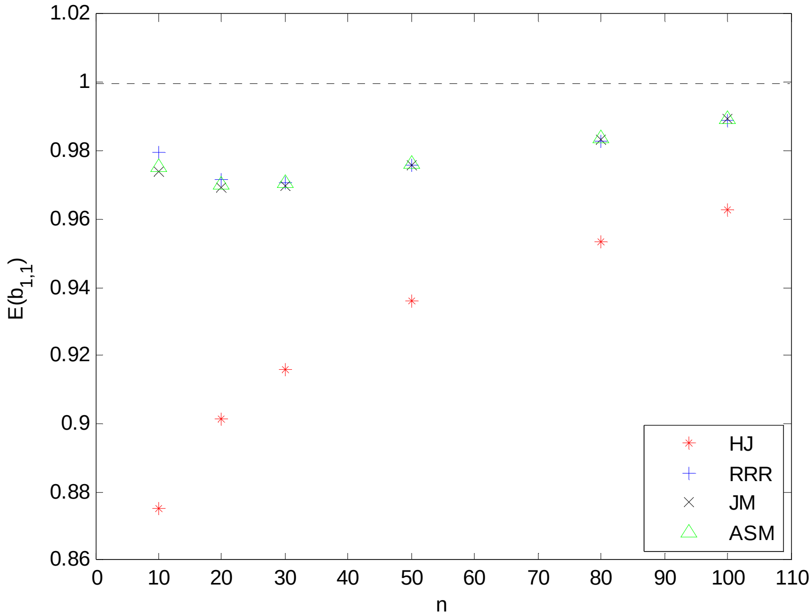


Figure 4-9: Comparison of the shape parameter estimators for censored data, obtained by LSE with different plotting positions used, at different *n*: the values of $E(\hat{\beta}_{1,1})$ at *c* = 10%.

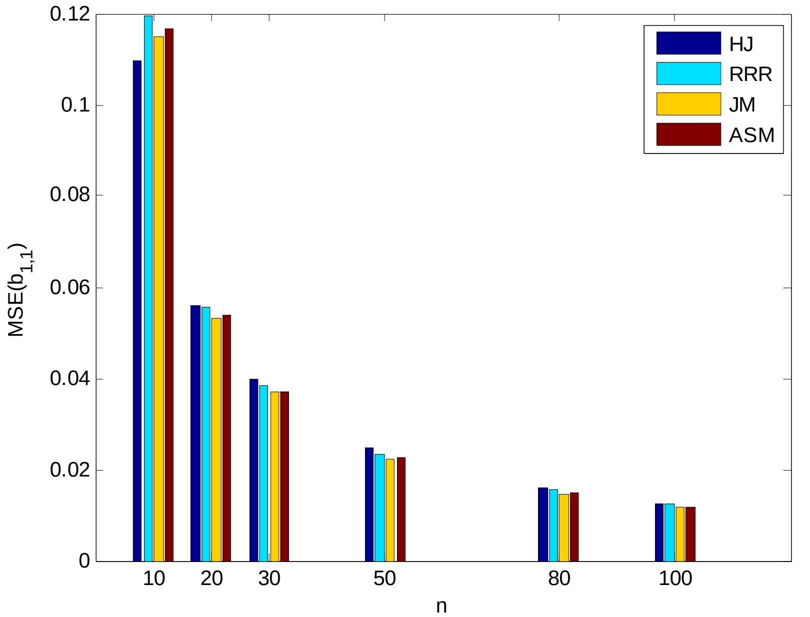


Figure 4-10: Comparison of the shape parameter estimators for censored data, obtained by LSE with different plotting positions used, at different *n*: the values of $MSE(\hat{\beta}_{1,1})$ at *c* = 10%.

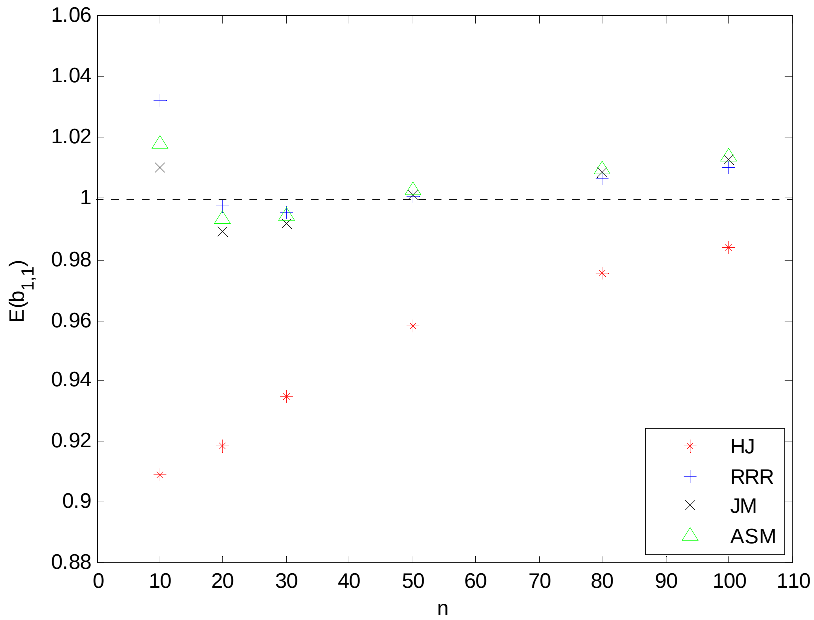


Figure 4-11: Comparison of the shape parameter estimators for censored data, obtained by LSE with different plotting positions used, at different *n*: the values of $E(\hat{\beta}_{1,1})$ at *c* = 30%.

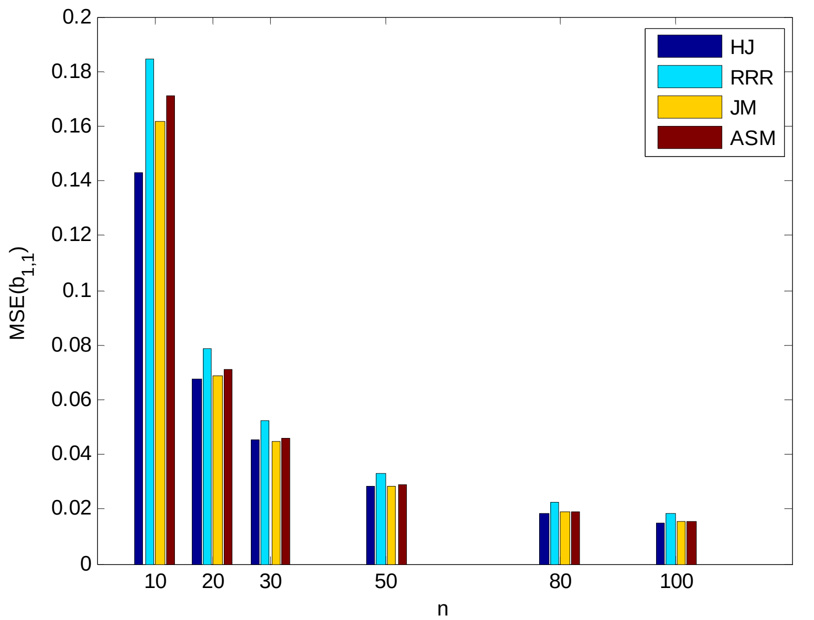


Figure 4-12: Comparison of the shape parameter estimators for censored data, obtained by LSE with different plotting positions used, at different *n*: the values of $MSE(\hat{\beta}_{1,1})$ at *c* = 30%.

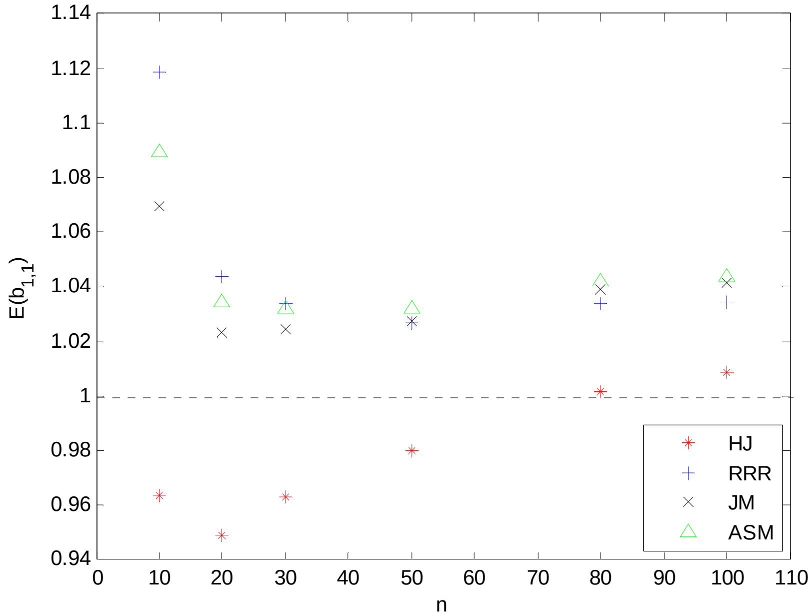


Figure 4-13: Comparison of the shape parameter estimators for censored data, obtained by LSE with different plotting positions used, at different *n*: the values of $E(\hat{\beta}_{1,1})$ at *c* = 50%.

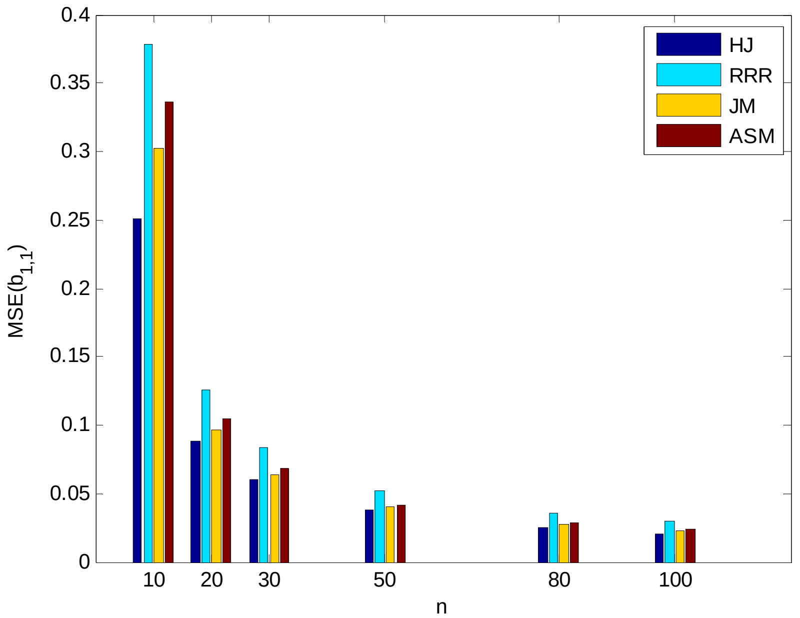


Figure 4-14: Comparison of the shape parameter estimators for censored data, obtained by LSE with different plotting positions used, at different *n*: the values of $MSE(\hat{\beta}_{1,1})$ at *c* = 50%.

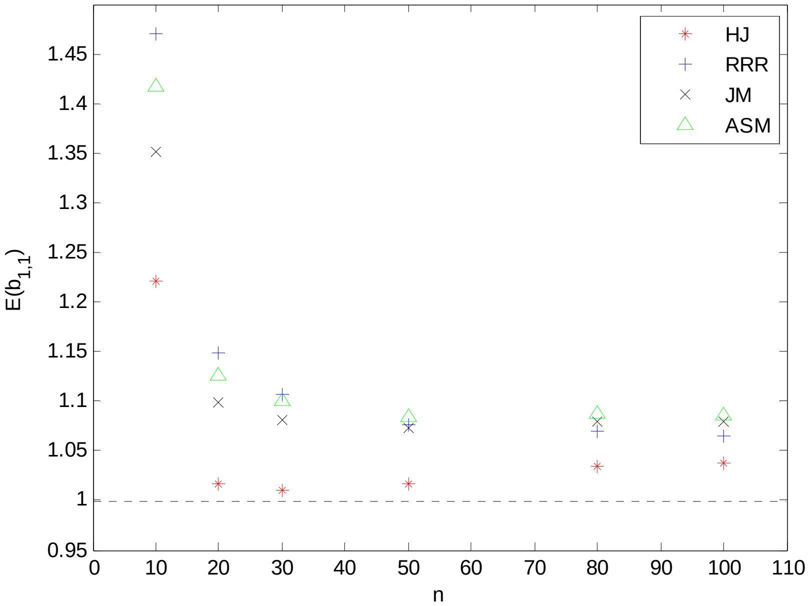


Figure 4-15: Comparison of the shape parameter estimators for censored data, obtained by LSE with different plotting positions used, at different n: the values of $E(\hat{\beta}_{1,1})$ at $c = 70\%$.

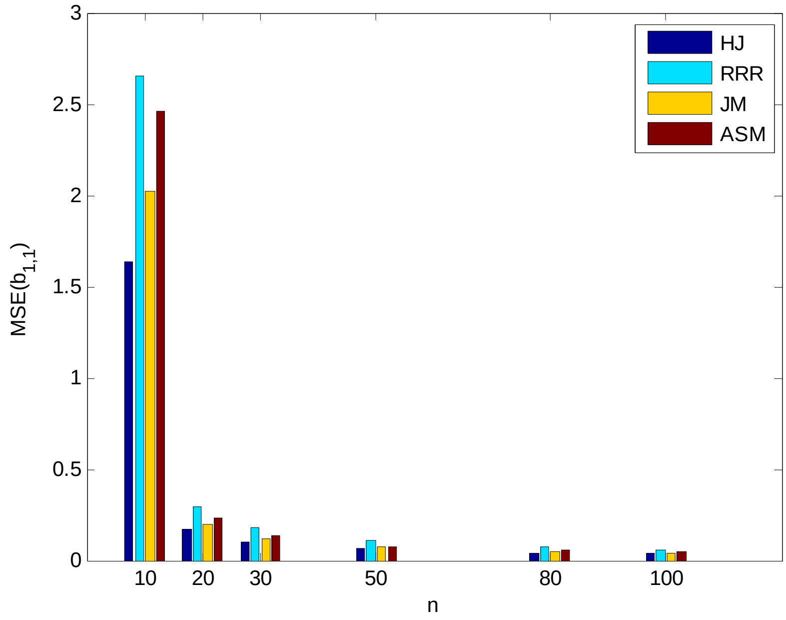


Figure 4-16: Comparison of the shape parameter estimators for censored data, obtained by LSE with different plotting positions used, at different n: the values of $MSE(\hat{\beta}_{1,1})$ at $c = 70\%$.

Simulation Results for Estimators of α (Figure 4-17 – Figure 4-22)

1) ***Bias of $\hat{\alpha}$ (refer to Figure 4-17, Figure 4-19 and Figure 4-21):*** $\hat{\alpha}$ of the RRRM is very unstable. The bias is extremely large at $\beta_T = 0.5$ and $c = 50\%, 70\%$. The bias of $\hat{\alpha}$ of the RRRM is the largest among all methods at high censoring levels (50%, 70%) at most times; at low censoring levels (10%), however, the bias of $\hat{\alpha}$ of the RRRM is in the middle. The bias of $\hat{\alpha}$ of HJ is the largest at $c = 10\%$. JM always performs best for estimating α in view of bias.

2) ***MSE of $\hat{\alpha}$ (refer to Figure 4-18, Figure 4-20 and Figure 4-22):*** Similar to the results for bias, the MSE of $\hat{\alpha}$ of the RRRM is extremely large at high censoring levels (50%, 70%). Among the other three methods, JM always has the smallest MSE, ASM is better than HJ at low censoring levels, and HJ is better than ASM at high censoring levels.

3) ***Both bias and MSE:*** Combining both bias and MSE, JM is the best for estimating α and the RRRM should be used with caution.

Figure 4-17: Comparison of the scale parameter estimators for censored data, obtained by LSE with different plotting positions used, at different *n* and *c*: the values of E($\hat{a}_{1,0.5}$).

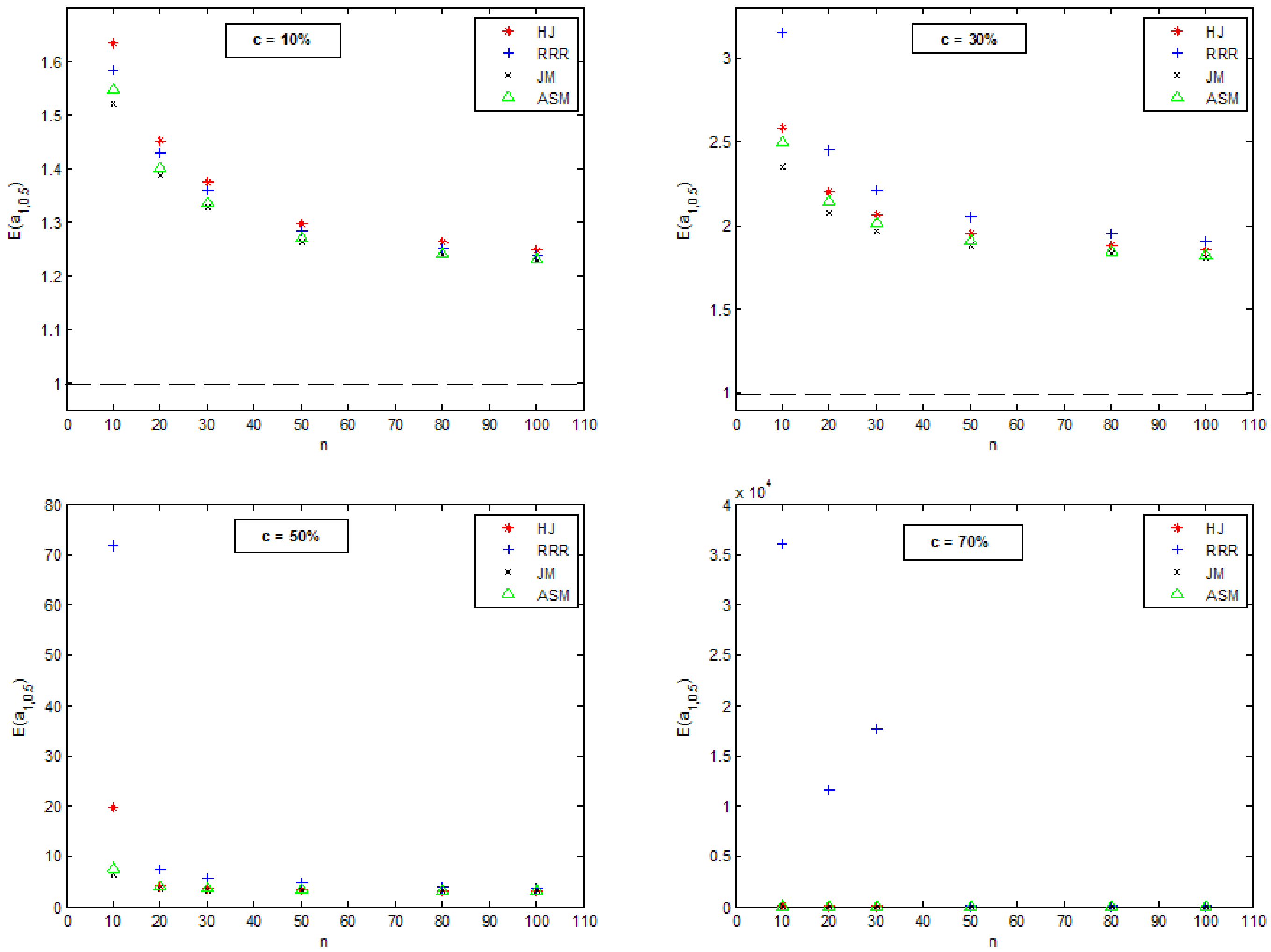


Figure 4-18: Comparison of the scale parameter estimators for censored data, obtained by LSE with different plotting positions used, at different *n* and *c*: the values of MSE($\hat{a}_{1,0.5}$).

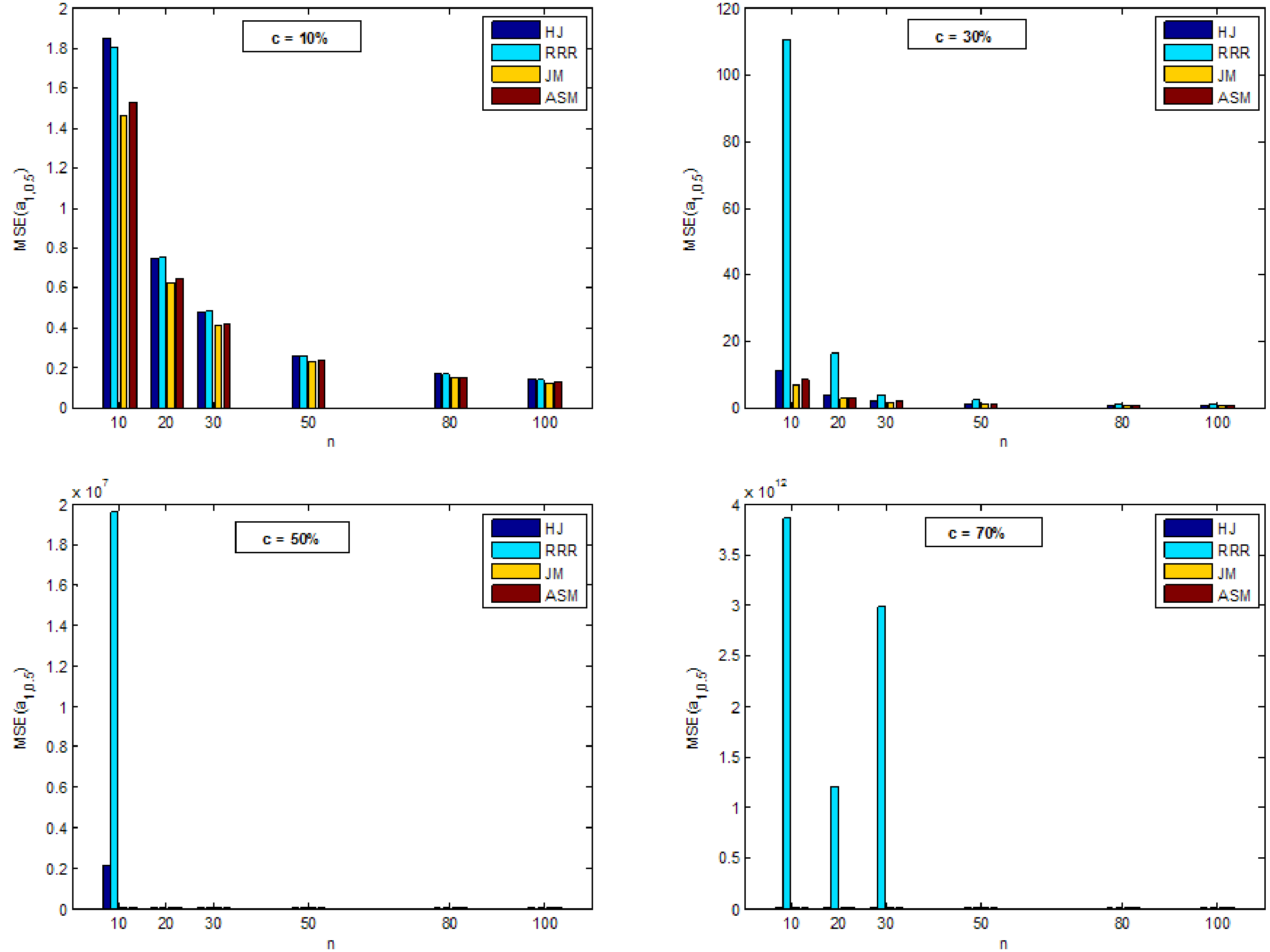


Figure 4-19: Comparison of the scale parameter estimators for censored data, obtained by LSE with different plotting positions used, at different *n* and *c*: the values of $E(\hat{a}_{1,1})$.

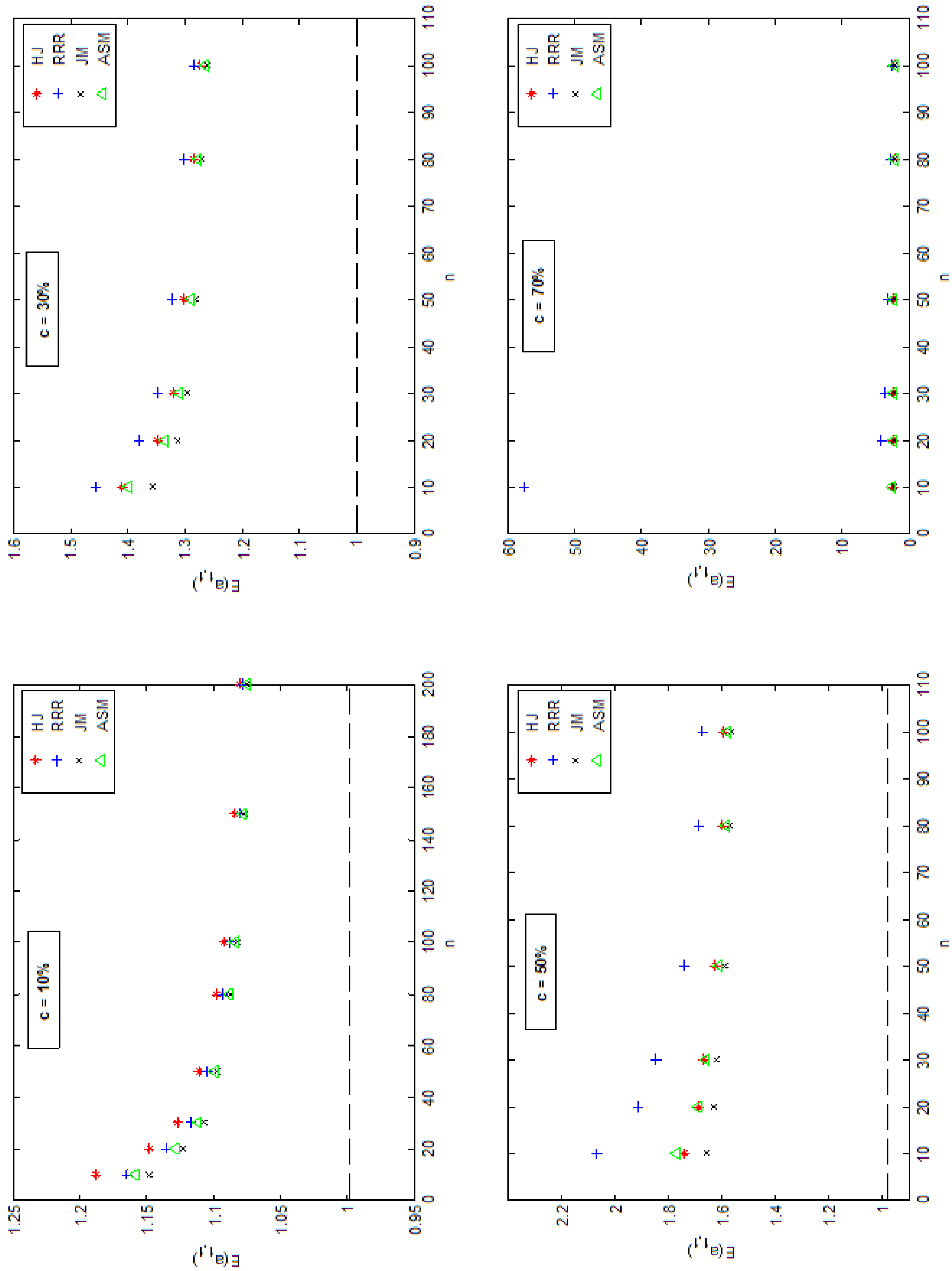


Figure 4-20: Comparison of the scale parameter estimators for censored data, obtained by LSE with different plotting positions used, at different *n* and *c*: the values of MSE($\hat{a}_{1,1}$).

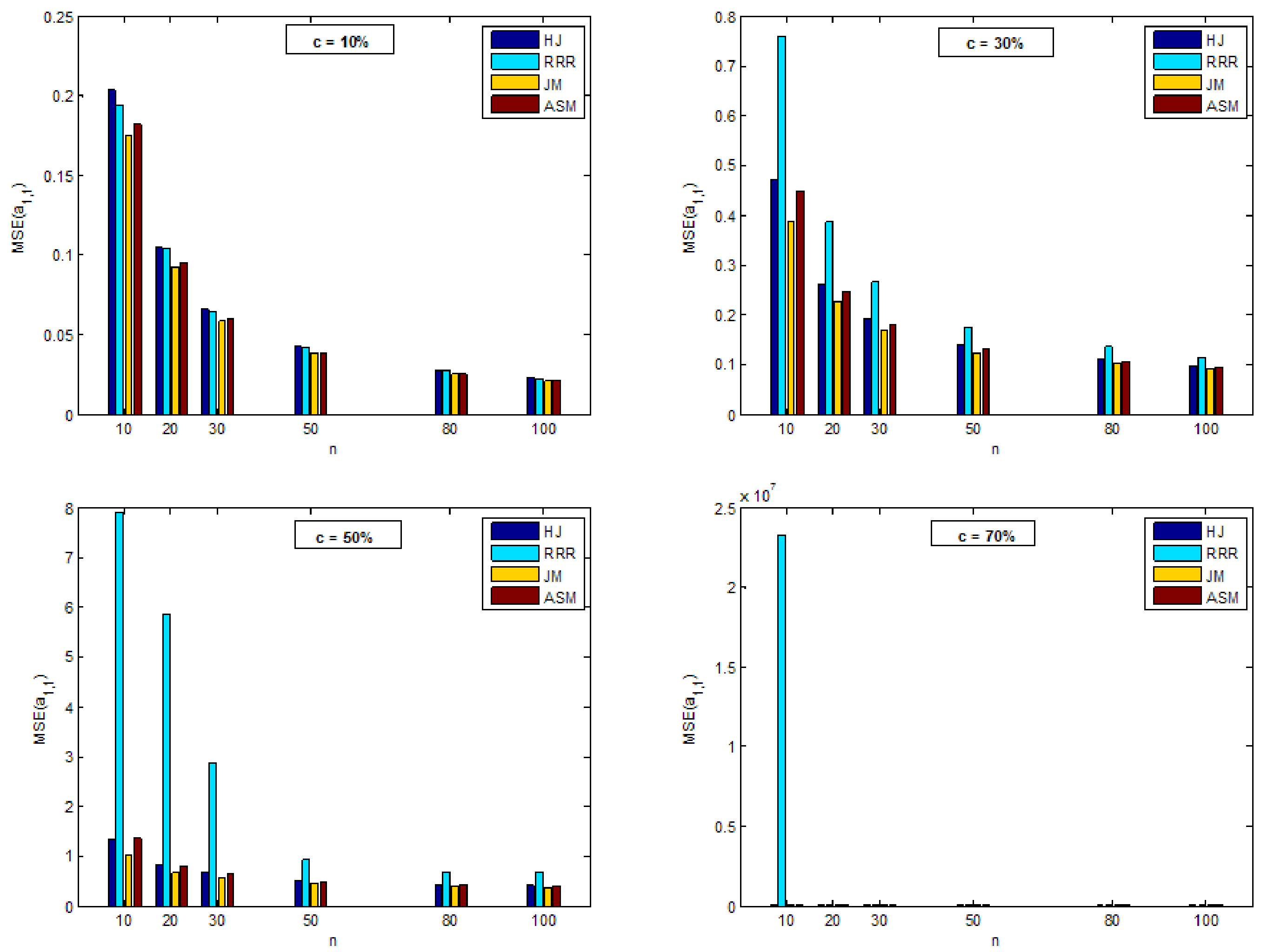
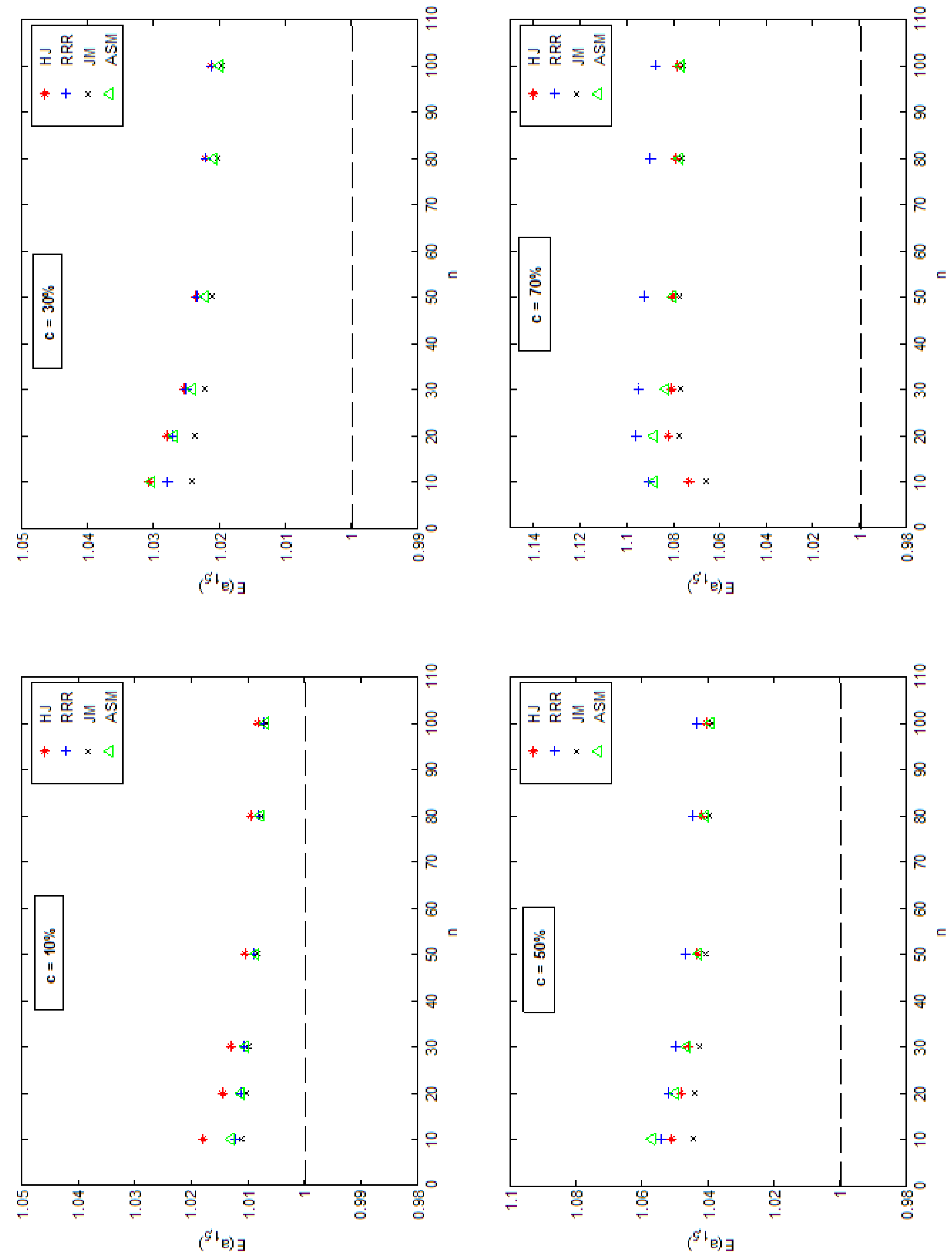


Figure 4-21: Comparison of the scale parameter estimators for censored data, obtained by LSE with different plotting positions used, at different *n* and *c*: the values of E($\hat{a}_{1,5}$).



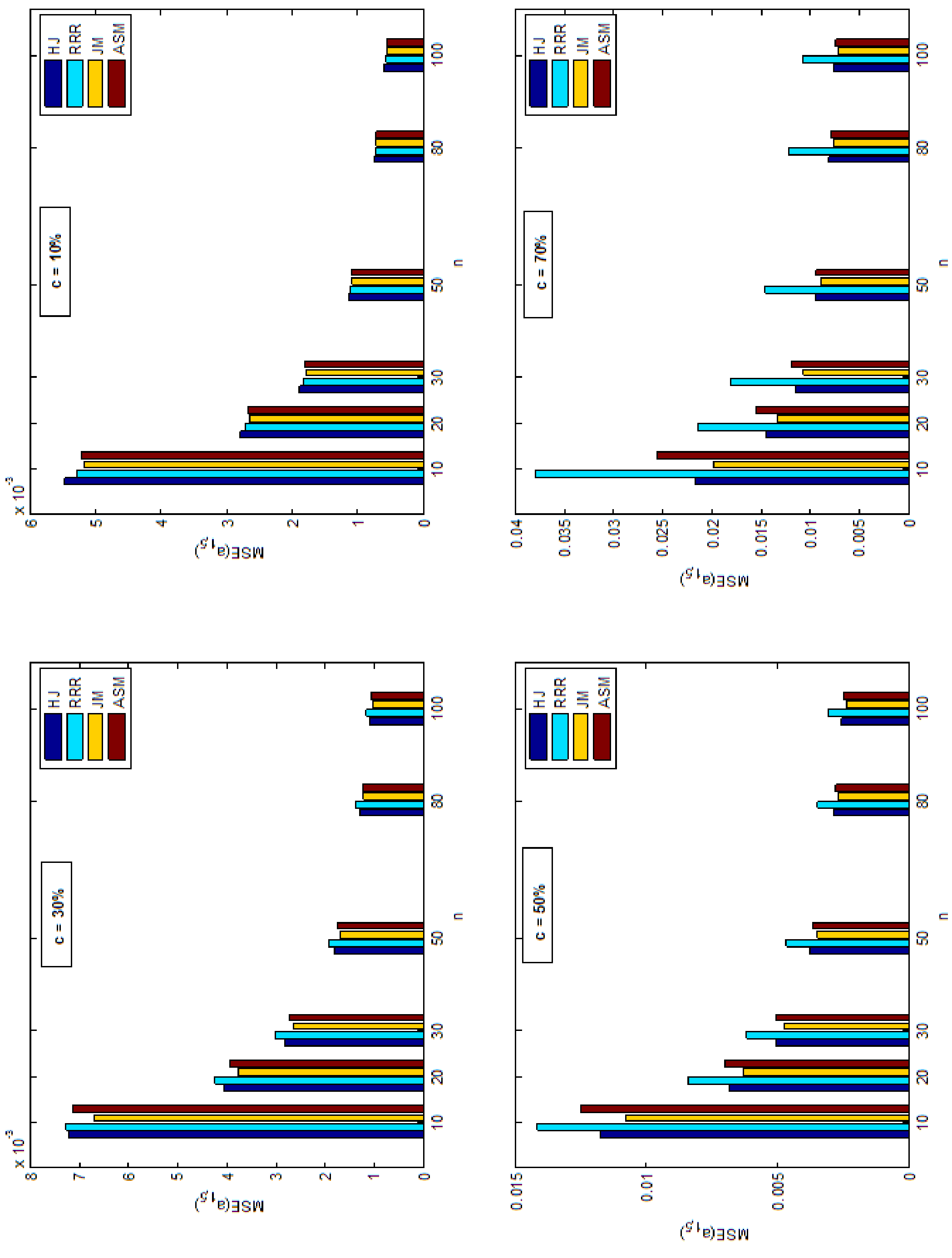


Figure 4-22: Comparison of the scale parameter estimators for censored data, obtained by LSE with different plotting positions used, at different *n* and *c*: the values of MSE($\hat{a}_{1,5}$).

4.3.5 Summary of Results

The following conclusions are made by combining the results for α and β.

For Complete Data

1) For estimating β, the Bernard estimator performs very well when $n < 10$. When $n \geq 10$, the Ross estimator or the Drap-Kos estimator is preferred because the resulted $\hat{\beta}$ is nearly unbiased. However, the Ross estimator or the Drap-Kos estimator cannot improve the efficiency of estimation.
2) For estimating α, the Hazen estimator is best especially at small β_T and small n. The Bernard estimator is the second best, followed by the Ross estimator and the Drap-Kos estimator.
3) The Ross estimator is a good approximation to the Drap-Kos estimator (i.e., the exact expected plotting positions). The two methods perform similar for estimating β, and the Ross estimator even performs slightly better for estimating α.

For Censored Data

1) For estimating β, JM, ASM and RRRM are good for samples with low censoring levels, say $c < 50\%$. Considering the application simplicity, JM is recommended to be used. HJ should be preferred for samples with high censoring levels, say $c \geq 50\%$.
2) For estimating α, JM is recommended for all censoring levels and sample sizes. RRRM should be used with caution because it can generate extremely large bias and MSE.

4.4 Modification 3: LS *Y* on *X* vs. LS *X* on *Y*

As mentioned in Section 2.3, the conventional setting of the independent and dependent variables in the LSE method is that $X = \ln T$ and $Y = \ln[-\ln(1-F)]$, which is consistent with the WPP where the *X*-axis is t and *Y*-axis is F. This method is named LS Y on X in this study. Some researchers (see, e.g., Abernethy, 2000) argued that it is more appropriate to set $Y = \ln T$ and $X = \ln[-\ln(1-F)]$ because *t* is the measured value or output from the experiment, and F is estimated by some non-parametric method and is independent of T. The replacement of the setting for *X* and *Y* has the same effect as reversing the regression direction, and by doing this, another method named LS *X* on *Y* is proposed. Abernethy (2000) compared the two methods on parameter estimation via Monte Carlo simulations and suggested LS *X* on *Y* to be used. However, the author's experiment examined only a few sample sizes and only complete data.

Nowadays, LS Y on X is the default method for LSE used by most Weibull researchers and practitioners. However, it was found in the early literature that quite a few Weibull researchers including Weibull (1967), White (1969) and Mann et al. (1974) used LS *X* on *Y*. This motivated us to conduct a careful comparison between these two methods. As the OLSE method cannot provide unbiased estimators of α and β, the two methods must perform differently.

In the following, Section 4.4.1 presents the theoretical background and the estimating equations for LS *Y* on *X* and LS *X* on *Y*, respectively. In Section 4.4.2, the two methods are examined as two regression models by analytical methods. Some results are found for the ratio of the MS_{Error} of the two models and suggestions are given on when to use which method in view of the goodness of model. Finally, Section 4.4.3 presents the Monte Carlo experiment that compares the two methods on parameter estimation.

4.4.1 Estimating Equations of LS *Y* on *X* and LS *X* on *Y*

Let $X = \ln T$ and $Y = \ln[-\ln(1-F)]$ for both methods. The calculation for x_i and y_i for complete samples and censored samples, respectively, can be found in Section 2.3.

Estimators of LS Y on X

If the Bernard estimator or the HJ estimator is used for estimating F, LS *Y* on *X* is the OLSE method. Therefore, the estimating equation of the LS *Y* on *X* method is given by Equation (2-12). Here it is rewritten as

$$\begin{cases} \hat{\beta}_{LS-YX} = \dfrac{\sum_{i=1}^{r}\left[(x_i-\bar{x})(y_i-\bar{y})\right]}{\sum_{i=1}^{r}(x_i-\bar{x})^2} = \dfrac{r\sum_{i=1}^{r}x_iy_i - \sum_{i=1}^{r}x_i\cdot\sum_{i=1}^{r}y_i}{r\sum_{i=1}^{r}x_i^2-\left(\sum_{i=1}^{r}x_i\right)^2} \\ \hat{\alpha}_{LS-YX} = \exp\left(-\dfrac{\bar{y}-\hat{\beta}_{LS-YX}\bar{x}}{\hat{\beta}_{LS-YX}}\right) = \exp\left(-\dfrac{\sum_{i=1}^{r}y_i-\hat{\beta}_{LS-YX}\sum_{i=1}^{r}x_i}{r\hat{\beta}_{LS-YX}}\right) \end{cases} \quad (4\text{-}13)$$

where $\hat{\alpha}_{LS-YX}$ and $\hat{\beta}_{LS-yx}$ denote the estimators of α and β of the LS *Y* on *X* method. The equation is applicable for both complete and censored data. $r = n$ for a complete sample, and $r < n$ for a censored sample.

Estimators of LS X on Y

Rewrite Equation (2-7) as

$$X = A' + B'Y \tag{4-14}$$

where $A' = \ln\alpha$ and $B' = 1/\beta$. Thus the estimation of α and β can be transferred to the estimation of the regression coefficients for a simple linear regression model of the form $X = A' + B'Y + e'$, where e' is the error term.

The objective function of the LS *X* on *Y* method is

$$\min S' = \sum_{i=1}^{r}\left[x_i - (A' + B'y_i)\right]^2 \tag{4-15}$$

The estimating equations can be easily obtained as

$$\begin{cases} \hat{B}' = \dfrac{\sum_{i=1}^{r}\left[(x_i - \bar{x})(y_i - \bar{y})\right]}{\sum_{i=1}^{r}(y_i - \bar{y})^2} = \dfrac{r\sum_{i=1}^{r}x_i y_i - \sum_{i=1}^{r}x_i \cdot \sum_{i=1}^{r}y_i}{r\sum_{i=1}^{r}y_i^2 - \left(\sum_{i=1}^{r}y_i\right)^2} \\ \hat{A}' = \bar{x} - \hat{B}'\bar{y} = \dfrac{\sum_{i=1}^{r}x_i - \hat{B}'\sum_{i=1}^{r}y_i}{r} \end{cases} \tag{4-16}$$

Thus the estimators of α and β can be obtained by

$$\begin{cases} \hat{\beta}_{LS-XY} = \dfrac{\sum_{i=1}^{r}(y_i - \bar{y})^2}{\sum_{i=1}^{r}\left[(x_i - \bar{x})(y_i - \bar{y})\right]} = \dfrac{r\sum_{i=1}^{r}y_i^2 - \left(\sum_{i=1}^{r}y_i\right)^2}{r\sum_{i=1}^{r}x_i y_i - \sum_{i=1}^{r}x_i \cdot \sum_{i=1}^{r}y_i} \\ \hat{\alpha}_{LS-XY} = \exp\left(\dfrac{\hat{\beta}_{LS-XY}\bar{x} - \bar{y}}{\hat{\beta}_{LS-XY}}\right) = \exp\left(\dfrac{\hat{\beta}_{LS-XY}\sum_{i=1}^{r}x_i - \sum_{i=1}^{r}y_i}{r\hat{\beta}_{LS-XY}}\right) \end{cases} \tag{4-17}$$

where $\hat{\alpha}_{LS-XY}$ and $\hat{\beta}_{LS-XY}$ denote the estimators of α and β of the LS *X* on *Y* method. The equation is applicable for both complete and censored data. $r = n$ for a complete sample, and $r < n$ for a censored sample.

4.4.2 Analytical Examination of the Two Methods

Let $$S_x = \sum_{i=1}^{r}(x_i - \overline{x})^2,\ S_y = \sum_{i=1}^{r}(y_i - \overline{y})^2, S_{xy} = \sum_{i=1}^{r}(x_i - \overline{x})(y_i - \overline{y}) \quad (4\text{-}18)$$

Thus for the LS *Y* on *X* method, the estimators of A and B can be expressed by

$$\hat{B} = S_{xy} / S_x,\ \hat{A} = \overline{y} - \hat{B}\overline{x} \quad (4\text{-}19)$$

Similarly, for the LS X on Y method, the LS estimators of A' and B' can be obtained as

$$\hat{B}' = S_{xy} / S_y,\ \hat{A}' = \overline{x} - \hat{B}'\overline{y} \quad (4\text{-}20)$$

The common model statistics for a linear regression model include R^2 (i.e., coefficient of determination) and MS_{Error}. R^2 is frequently used to measure the goodness-of-fit. MS_{Error} is the variance of error, and a smaller MS_{Error} normally means a better model. To compare the models of LS *Y* on *X* and LS *X* on *Y*, the ratios of their R^2 and MS_{Error} are derived. The results are given below.

The definition of R^2 is

$$R^2 = 1 - \frac{SS_{Model}}{SS_{Total}} \quad (4\text{-}21)$$

Thus the ratio of the R^2 of the two models is

$$\frac{R^2_{LS-YX}}{R^2_{LS-XY}} = \frac{1 - \dfrac{S_{xy}^2 \big/ S_x}{S_y}}{1 - \dfrac{S_{xy}^2 \big/ S_y}{S_x}} = \frac{1 - \dfrac{S_{xy}^2}{S_x S_y}}{1 - \dfrac{S_{xy}^2}{S_x S_y}} = 1 \quad (4\text{-}22)$$

Therefore, the two methods generate same values of R^2.

The ratio of the MS_{Error} of the two models is given by

$$\frac{MS_{error(LS-YX)}}{MS_{error(LS-XY)}} = \frac{\dfrac{S_y - S_{xy}^2 / S_x}{(r-2)}}{\dfrac{S_x - S_{xy}^2 / S_y}{(r-2)}} = \frac{S_y - S_{xy}^2 / S_x}{S_x - S_{xy}^2 / S_y} = \frac{S_y}{S_x} \tag{4-23}$$

Since

$$\begin{cases} S_x = \sum_{i=1}^{r} (x_i - \bar{x})^2 = Var(X) \\ S_y = \sum_{i=1}^{r} (y_i - \bar{y})^2 = Var(Y) \end{cases} \tag{4-24}$$

Equation (4-24) becomes

$$\frac{MS_{error(LS-YX)}}{MS_{error(LS-XY)}} = \frac{Var(Y)}{Var(X)} \tag{4-25}$$

It can be obtained, either from $Y = A + BX$ of the LS Y on X method or $X = A' + B'Y$ of the LS X on Y method, that

$$Var(Y) = \beta^2 Var(X) \tag{4-26}$$

Thus finally, the ratio of the MS_{Error} of the two models is

$$\frac{MS_{error(LS-YX)}}{MS_{error(LS-XY)}} = \beta^2 \tag{4-27}$$

Based on this equation, if $\beta < 1$, we have $MS_{error(LS-YX)} < MS_{error(LS-XY)}$; if $\beta = 1$, $MS_{error(LS-YX)} = MS_{error(LS-XY)}$; and if $\beta > 1$, $MS_{error(LS-YX)} > MS_{error(LS-XY)}$.

In summary, the analytical examinations on the two methods show that LS Y on X and LS X on Y generate same values of R^2, which means the two methods perform similarly in view of the goodness-of-fit. However, the examination of MS_{error} suggests that LS Y on X be used when $\beta < 1$ and LS X on Y be used when $\beta > 1$.

4.4.3 Simulation Study of the Two Methods

A Monte Carlo experiment was conducted to compare the performance of LS Y on X and LS X on Y on parameter estimation for complete and multiply censored samples, respectively. The conventional methods for estimating F used in OLSE, i.e., the Bernard estimator for complete data and the HJ estimator for censored data, are used for both methods. The simulation conditions are summarized in Table 4-5.

Table 4-5: Setting of experiment factors. The experiment is to compare the estimators of LS *Y* on *X* and LS *X* on *Y*.

Factors	Values
α_T	1
β_T	0.5, 1, 2, 3, 5
n	5 – 20, 22, …, 28, 30, 35, …, 45, 50, 80, 100 (complete data) 10, 20, …, 90, 100, 150, 200 (censored data)
c	10%, 20%, …, 70%, 80%
M	10000
Methods	LS Y on X, LS X on Y

For a randomly generated Weibull sample, the two methods were used to generate the LS estimates of α and β simultaneously. This procedure was repeated for 10000 times in each combination of α_T, β_T, n and c. Finally, the mean and MSE of $\hat{\alpha}$ and $\hat{\beta}$ for each method were calculated as the comparison criteria.

4.4.3.1 Comparison Results for Complete Data

The simulation results for the shape parameter estimators are shown in Figure 4-23 and Table 4-6, and the results for the scale parameter estimators are shown in Figure 4-24 and Table 4-7. The bias of the estimators can be easily compared using Figure 4-23 and Figure 4-24. Table 4-6 and Table 4-7 tabulate the mean and MSE of the estimators at selected simulation conditions. Not all the simulation results are tabulated; however, the omitted results will not affect the following conclusions.

<u>Simulation Results for Estimators of β (Figure 4-23 and Table 4-6)</u>

1) ***General observations:*** In view of both bias and MSE of $\hat{\beta}$, LS *Y* on *X* outperforms LS *X* on *Y* when $n \le 10$. On the other hand, from $n = 11$ onwards, LS *X* on *Y* outperforms LS *Y* on *X* for estimating β in view of bias but the values of MSE of the two estimators are close.

2) ***Bias of $\hat{\beta}$ of LS X on Y:*** The relative bias of $\hat{\beta}$ of LS *X* on *Y* is larger than 5% when $n < 10$, but it drops fast from $n = 5$ to 20, and the bias becomes significantly smaller than that of LS *Y* on *X* at $n > 20$. The estimator of β of LS *X* on *Y* is nearly unbiased when $n \ge 25$ and the bias reaches 0 at about $n = 40$.

3) ***Bias of $\hat{\beta}$ of LS Y on X:*** For LS *Y* on *X*, the bias of $\hat{\beta}$ reaches 0 between $n = 6$ and $n = 7$. During $10 < n < 30$, the relative bias is like a constant and remains at 4% or so.

4) ***Consistency of $\hat{\beta}$:*** $\hat{\beta}$ of both methods are inconsistent with the sample size.

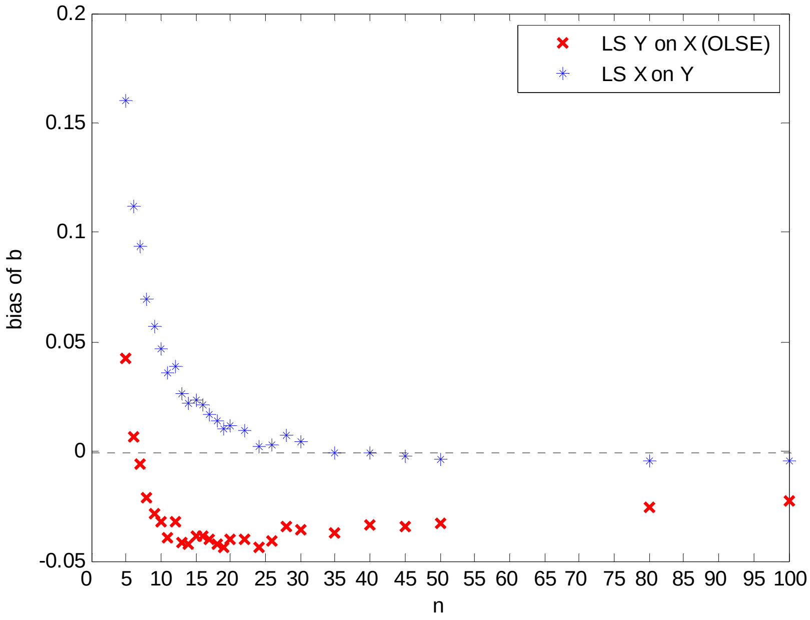


Figure 4-23: Bias of $\hat{\beta}_{1,1}$, obtained by LS *Y* on *X* and LS *X* on *Y*, at different *n*.

Table 4-6: Simulation results of $\hat{\beta}_{1,1}$ for complete data, generated by LS *Y* on *X* and LS *X* on *Y*, at different *n*: the values of $E(\hat{\beta}_{1,1})$ and $MSE(\hat{\beta}_{1,1})$ (in parentheses).

Method	*n* 5	6	7	8	9	10	11	12	14	16
LS Y on X	1.042 (0.342)	1.007 (0.228)	0.994 (0.180)	0.979 (0.143)	0.972 (0.117)	0.968 (0.109)	0.961 (0.091)	0.968 (0.084)	0.958 (0.070)	0.962 (0.061)
LS X on Y	1.160 (0.461)	1.112 (0.297)	1.093 (0.234)	1.070 (0.180)	1.057 (0.144)	1.047 (0.132)	1.036 (0.106)	1.039 (0.100)	1.022 (0.078)	1.021 (0.067)
	18	**20**	**24**	**26**	**28**	**30**	**40**	**50**	**80**	**100**
LS Y on X	0.958 (0.053)	0.960 (0.050)	0.956 (0.043)	0.959 (0.038)	0.966 (0.036)	0.965 (0.034)	0.966 (0.026)	0.967 (0.021)	0.974 (0.014)	0.977 (0.011)
LS X on Y	1.014 (0.057)	1.012 (0.053)	1.002 (0.044)	1.003 (0.039)	1.008 (0.037)	1.004 (0.035)	1.000 (0.025)	0.997 (0.020)	0.996 (0.013)	0.995 (0.011)

Simulation Results for Estimators of α (Figure 4-24 and Table 4-7)

1) ***General observations:*** In view of both bias and MSE, the method of LS *X* on *Y* always outperforms LS *Y* on *X* for estimating α .

2) ***Bias of*** $\hat{\alpha}$ ***:*** The bias of $\hat{\alpha}$ of LS *X* on *Y* is significantly smaller than that of LS *Y* on *X* at $\beta_T \le 2$. The differences are small when $\beta_T > 2$, and both estimators of α are nearly unbiased and have small MSE at $\beta_T = 5$.

3) ***Consistency of*** $\hat{\alpha}$ **:** For LS *Y* on *X*, the bias of $\hat{\alpha}$ decreases as *n* and β_T increase. However, $\hat{\alpha}$ of LS *X* on *Y* is inconsistent with β_T as the estimator is unbiased at $\beta_T = 2$.

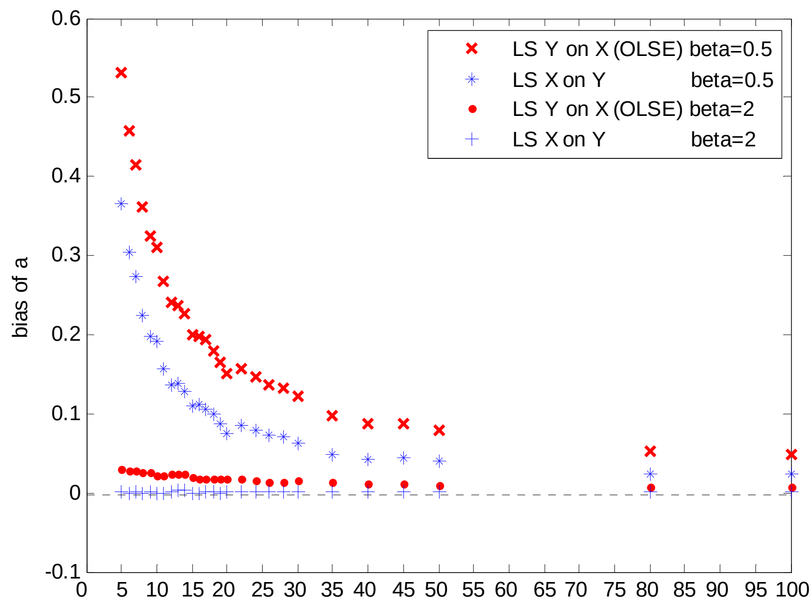


Figure 4-24: Bias of $\hat{\alpha}_{1,\beta_T}$, obtained by LS *Y* on *X* and LS *X* on *Y*, at different *n* and β_T.

Table 4-7: Simulation results of $\hat{\alpha}_{1,\beta_T}$ for complete data, generated by LS *Y* on *X* and LS *X* on *Y*, at different *n* and β_T: the values of $E(\hat{\alpha}_{1,\beta_T})$ and $MSE(\hat{\alpha}_{1,\beta_T})$ (in parentheses).

		n													
	Method	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**	**20**	**30**	**40**	**50**	**80**	**100**
β_T=0.5	LS Y on X	1.532	1.457	1.416	1.362	1.325	1.311	1.241	1.199	1.150	1.122	1.088	1.079	1.052	1.049
		(2.571)	(1.918)	(1.563)	(1.286)	(1.045)	(0.944)	(0.672)	(0.499)	(0.337)	(0.221)	(0.147)	(0.117)	(0.071)	(0.057)
	LS X on Y	1.367	1.305	1.273	1.225	1.198	1.192	1.136	1.109	1.075	1.063	1.041	1.040	1.024	1.024
		(1.945)	(1.431)	(1.177)	(0.943)	(0.769)	(0.709)	(0.502)	(0.388)	(0.269)	(0.179)	(0.126)	(0.101)	(0.063)	(0.051)
β_T=1	LS Y on X	1.113	1.109	1.106	1.087	1.092	1.082	1.077	1.056	1.051	1.036	1.032	1.026	1.020	1.017
		(0.270)	(0.235)	(0.205)	(0.171)	(0.154)	(0.141)	(0.116)	(0.089)	(0.066)	(0.043)	(0.033)	(0.026)	(0.016)	(0.012)
	LS X on Y	1.052	1.048	1.048	1.033	1.039	1.031	1.032	1.016	1.017	1.010	1.010	1.007	1.006	1.006
		(0.235)	(0.201)	(0.173)	(0.147)	(0.131)	(0.120)	(0.101)	(0.078)	(0.058)	(0.039)	(0.030)	(0.024)	(0.014)	(0.012)
β_T=1.5	LS Y on X	1.056	1.052	1.045	1.050	1.044	1.039	1.036	1.033	1.026	1.021	1.018	1.016	1.010	1.010
		(0.109)	(0.091)	(0.079)	(0.070)	(0.062)	(0.056)	(0.047)	(0.037)	(0.028)	(0.018)	(0.014)	(0.011)	(0.007)	(0.005)
	LS X on Y	1.016	1.014	1.009	1.014	1.011	1.007	1.006	1.007	1.004	1.003	1.004	1.003	1.001	1.003
		(0.099)	(0.083)	(0.071)	(0.062)	(0.056)	(0.051)	(0.042)	(0.034)	(0.026)	(0.017)	(0.013)	(0.010)	(0.006)	(0.005)
β_T=2	LS Y on X	1.030	1.028	1.029	1.025	1.026	1.022	1.023	1.019	1.018	1.015	1.012	1.010	1.008	1.007
		(0.059)	(0.049)	(0.044)	(0.038)	(0.033)	(0.030)	(0.026)	(0.020)	(0.015)	(0.010)	(0.008)	(0.006)	(0.004)	(0.003)
	LS X on Y	1.001	1.000	1.001	1.000	1.001	0.998	1.002	1.000	1.002	1.002	1.001	1.001	1.001	1.001
		(0.055)	(0.046)	(0.040)	(0.035)	(0.031)	(0.028)	(0.024)	(0.018)	(0.014)	(0.009)	(0.007)	(0.006)	(0.004)	(0.003)
β_T=3	LS Y on X	1.013	1.017	1.014	1.014	1.014	1.013	1.013	1.012	1.012	1.009	1.007	1.006	1.005	1.004
		(0.026)	(0.021)	(0.019)	(0.016)	(0.014)	(0.013)	(0.011)	(0.009)	(0.007)	(0.005)	(0.003)	(0.003)	(0.002)	(0.001)
	LS X on Y	0.994	0.999	0.996	0.997	0.997	0.998	0.998	0.999	1.001	1.000	1.000	1.000	1.001	1.000
		(0.025)	(0.020)	(0.018)	(0.016)	(0.014)	(0.012)	(0.010)	(0.008)	(0.006)	(0.004)	(0.003)	(0.003)	(0.002)	(0.001)
β_T=5	LS Y on X	1.005	1.006	1.005	1.006	1.006	1.007	1.005	1.005	1.006	1.004	1.004	1.004	1.003	1.002
		(0.009)	(0.008)	(0.007)	(0.006)	(0.005)	(0.005)	(0.004)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
	LS X on Y	0.993	0.994	0.994	0.995	0.996	0.997	0.997	0.998	0.999	0.999	0.999	1.000	1.000	1.000
		(0.009)	(0.008)	(0.006)	(0.006)	(0.005)	(0.005)	(0.004)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)

4.4.3.2 Comparison Results for Censored Data

The comparison results for multiply censored data are shown in Table 4-8 for the shape parameter estimators and Table 4-9 for the scale parameter estimators. The following conclusions can be observed from the two tables.

Simulation Results for Estimators of β (Table 4-8)

1) ***Bias of $\hat{\beta}$:*** The mean values of $\hat{\beta}$ of LS *X* on *Y* are always larger than that of LS *Y* on *X*. In view of the bias of $\hat{\beta}$, LS *X* on *Y* is clearly better at low censoring levels (10% – 40%), where the bias of $\hat{\beta}$ of LS *X* on *Y* is typically 8% – 9% smaller than that of LS *Y* on *X* at $n = 10$, and 2% – 6% smaller at $n > 10$. On the other hand, LS *Y* on *X* outperforms LS *X* on *Y* at high censoring levels (60% – 80%), where the bias of $\hat{\beta}$ of LS *Y* on *X* is typically 5% – 9% less than that of LS *X* on *Y*. At $c = 50\%$, LS *X* on *Y* is better when $n < 50$ and LS *Y* on *X* is better when $n \geq 50$. The difference between the bias of $\hat{\beta}$ of the two methods is significant at small sample sizes ($n = 10 - 20$). The bias of $\hat{\beta}$ of LS *Y* on *X* is close to 0 at the combination of, e.g., $c = 60\%$ and $n = 50 - 60$, and the bias of $\hat{\beta}$ of LS *X* on *Y* is close to 0 at the combination of, e.g., $c = 30\%$ and $n = 60 - 80$.

2) ***MSE of $\hat{\beta}$:*** The MSE of $\hat{\beta}$ of the two methods are close at most of the times, except when the sample size is very small. The MSE of $\hat{\beta}$ of LS *Y* on *X* is much smaller than that of LS *X* on *Y* at $n = 10$.

3) $\hat{\beta}$ of both method are inconsistent with n for a specific c and inconsistent with c for a specific n.

Table 4-8: Simulation results of $\hat{\beta}_{1,1}$ for multiply censored data, generated by LS *Y* on *X* and LS *X* on *Y*, at different *n* and *c*: the values of $E(\hat{\beta}_{1,1})$ and $MSE(\hat{\beta}_{1,1})$ (in parentheses).

		n											
c	Method	10	20	30	40	50	60	70	80	90	100	150	200
10%	LS Y on X	0.878	0.902	0.917	0.928	0.938	0.943	0.949	0.955	0.957	0.962	0.972	0.979
		(0.112)	(0.056)	(0.039)	(0.031)	(0.025)	(0.021)	(0.018)	(0.016)	(0.014)	(0.013)	(0.009)	(0.006)
	LS X on Y	0.953	0.954	0.959	0.964	0.969	0.971	0.975	0.978	0.979	0.982	0.987	0.992
		(0.119)	(0.052)	(0.034)	(0.027)	(0.021)	(0.017)	(0.016)	(0.014)	(0.012)	(0.011)	(0.007)	(0.006)
20%	LS Y on X	0.895	0.908	0.927	0.938	0.948	0.954	0.961	0.963	0.968	0.973	0.983	0.990
		(0.124)	(0.060)	(0.042)	(0.032)	(0.026)	(0.022)	(0.019)	(0.017)	(0.015)	(0.014)	(0.009)	(0.007)
	LS X on Y	0.973	0.963	0.973	0.976	0.981	0.985	0.988	0.988	0.992	0.995	1.000	1.004
		(0.137)	(0.057)	(0.038)	(0.029)	(0.023)	(0.020)	(0.017)	(0.015)	(0.013)	(0.012)	(0.008)	(0.006)
30%	LS Y on X	0.906	0.917	0.935	0.946	0.957	0.965	0.970	0.972	0.979	0.982	0.995	1.003
		(0.140)	(0.065)	(0.046)	(0.035)	(0.029)	(0.024)	(0.021)	(0.019)	(0.016)	(0.015)	(0.010)	(0.008)
	LS X on Y	0.989	0.976	0.983	0.987	0.994	0.998	1.000	1.001	1.006	1.007	1.015	1.019
		(0.158)	(0.064)	(0.043)	(0.032)	(0.026)	(0.022)	(0.019)	(0.017)	(0.015)	(0.014)	(0.010)	(0.008)
40%	LS Y on X	0.931	0.930	0.945	0.961	0.970	0.975	0.983	0.988	0.989	0.994	1.008	1.015
		(0.179)	(0.078)	(0.050)	(0.041)	(0.032)	(0.028)	(0.024)	(0.021)	(0.019)	(0.017)	(0.012)	(0.010)
	LS X on Y	1.017	0.994	0.997	1.007	1.010	1.012	1.017	1.019	1.018	1.021	1.030	1.033
		(0.214)	(0.080)	(0.049)	(0.039)	(0.030)	(0.026)	(0.023)	(0.019)	(0.018)	(0.017)	(0.012)	(0.010)
50%	LS Y on X	0.974	0.944	0.953	0.967	0.983	0.989	0.994	0.999	1.002	1.006	1.022	1.030
		(0.251)	(0.086)	(0.059)	(0.046)	(0.038)	(0.032)	(0.029)	(0.026)	(0.023)	(0.021)	(0.015)	(0.012)
	LS X on Y	1.067	1.012	1.011	1.017	1.029	1.030	1.032	1.035	1.036	1.038	1.047	1.051
		(0.314)	(0.093)	(0.060)	(0.046)	(0.038)	(0.032)	(0.029)	(0.025)	(0.023)	(0.021)	(0.015)	(0.013)
60%	LS Y on X	1.025	0.971	0.976	0.992	0.999	1.001	1.010	1.014	1.020	1.023	1.036	1.046
		(0.415)	(0.116)	(0.072)	(0.057)	(0.046)	(0.039)	(0.035)	(0.031)	(0.028)	(0.026)	(0.020)	(0.016)
	LS X on Y	1.119	1.047	1.039	1.049	1.050	1.047	1.053	1.054	1.059	1.059	1.066	1.071
		(0.509)	(0.135)	(0.078)	(0.061)	(0.049)	(0.040)	(0.036)	(0.033)	(0.029)	(0.028)	(0.021)	(0.018)
70%	LS Y on X	1.180	1.013	1.005	1.008	1.014	1.020	1.029	1.029	1.033	1.036	1.053	1.063
		(0.906)	(0.174)	(0.103)	(0.076)	(0.061)	(0.053)	(0.048)	(0.042)	(0.038)	(0.035)	(0.028)	(0.023)
	LS X on Y	1.266	1.095	1.076	1.071	1.073	1.074	1.079	1.076	1.079	1.080	1.089	1.094
		(1.060)	(0.218)	(0.120)	(0.085)	(0.067)	(0.057)	(0.053)	(0.046)	(0.042)	(0.038)	(0.030)	(0.026)
80%	LS Y on X	1.625	1.111	1.057	1.043	1.053	1.046	1.053	1.053	1.057	1.059	1.074	1.083
		(2.747)	(0.398)	(0.171)	(0.114)	(0.100)	(0.081)	(0.072)	(0.064)	(0.059)	(0.054)	(0.042)	(0.036)
	LS X on Y	1.625	1.200	1.136	1.116	1.121	1.110	1.114	1.111	1.113	1.113	1.119	1.123
		(2.747)	(0.517)	(0.215)	(0.138)	(0.118)	(0.094)	(0.083)	(0.072)	(0.066)	(0.061)	(0.047)	(0.041)

Simulation Results for Estimators of α (Table 4-9)

1) In general, in view of both bias and MSE, the method of LS *X* on *Y* always outperforms the method of LS *Y* on *X* for estimating α .

2) At $\beta_T = 0.5$, both methods perform unstable and generate extremely large bias and MSE at some conditions. The bias and MSE of $\hat{\alpha}$ of LS *X* on *Y* at most times are significantly smaller than those of LS *Y* on *X* at $\beta_T = 0.5$ and $\beta_T = 1$. The difference becomes larger as *c* increases. At $\beta_T = 5$, the two methods perform closely and LS *X* on *Y* is slightly better in view of both

bias and MSE. The bias of $\hat{\alpha}$ of both methods drops greatly as β_T increases.

3) The bias and MSE of $\hat{\alpha}$ of both methods are much larger than those of $\hat{\beta}$ when β_T is small.

Table 4-9: Simulation results of $\hat{\alpha}_{1,\beta_T}$ for multiply censored data, generated by LS *Y* on *X* and LS *X* on *Y*, at different *n*, β_T and *c*: the values of $E(\hat{\alpha}_{1,\beta_T})$ and $MSE(\hat{\alpha}_{1,\beta_T})$ (in parentheses)†.

			n													
		Method	**10**		**20**		**30**		**50**		**100**		**150**		**200**	
β_T=0.5	10%	LS Y on X	1.650	(1.987)	1.446	(0.764)	1.372	(0.460)	1.303	(0.267)	1.250	(0.140)	1.226	(0.100)	1.217	(0.083)
		LS X on Y	1.450	(1.181)	1.318	(0.504)	1.277	(0.324)	1.238	(0.200)	1.211	(0.113)	1.198	(0.083)	1.194	(0.071)
	20%	LS Y on X	2.008	(5.193)	1.756	(1.516)	1.659	(0.966)	1.576	(0.613)	1.502	(0.374)	1.473	(0.299)	1.458	(0.265)
		LS X on Y	1.713	(1.914)	1.567	(0.905)	1.517	(0.629)	1.478	(0.437)	1.444	(0.296)	1.431	(0.250)	1.423	(0.228)
	30%	LS Y on X	2.548	(23.24)	2.194	(3.390)	2.076	(2.191)	1.959	(1.438)	1.856	(0.942)	1.817	(0.796)	1.793	(0.723)
		LS X on Y	2.072	(3.379)	1.899	(1.711)	1.854	(1.303)	1.806	(0.979)	1.764	(0.738)	1.749	(0.663)	1.739	(0.623)
	40%	LS Y on X	3.561	(1289)	2.899	(10.10)	2.720	(8.430)	2.534	(3.686)	2.385	(2.406)	2.320	(1.997)	2.288	(1.841)
		LS X on Y	2.622	(7.226)	2.396	(3.568)	2.333	(2.785)	2.275	(2.195)	2.230	(1.787)	2.208	(1.638)	2.199	(1.570)
	50%	LS Y on X	5.083	(1119)	4.214	(223.9)	4.075	(4943)	3.478	(10.02)	3.230	(6.213)	3.120	(5.159)	3.065	(4.712)
		LS X on Y	3.491	(19.80)	3.156	(8.513)	3.058	(6.438)	3.005	(5.219)	2.949	(4.369)	2.917	(4.052)	2.902	(3.895)
	60%	LS Y on X	18.99	(1778543)	7.089	(2093)	6.098	(589.9)	5.300	(56.33)	4.734	(19.03)	4.538	(14.94)	4.434	(13.42)
		LS X on Y	5.639	(674.4)	4.509	(28.50)	4.326	(18.04)	4.241	(13.97)	4.160	(11.55)	4.119	(10.74)	4.098	(10.35)
	70%	LS Y on X	3660	x	453.6	x	13.177	(15694)	9.739	(1187)	8.086	(138.5)	7.574	(66.33)	7.284	(49.42)
		LS X on Y	327.1	x	8.750	x	7.063	(184.3)	6.695	(51.56)	6.524	(37.33)	6.469	(34.28)	6.434	(32.68)
	80%	LS Y on X	x	x	38376	x	591.5	x	56.16	(3303913)	22.94	(299856)	17.12	(1295)	16.02	(707.9)
		LS X on Y	x	x	768.6	x	30.57	x	14.43	(1326)	12.85	(257.7)	12.54	(176.4)	12.47	(162.2)
β_T=1	10%	LS Y on X	1.194	(0.209)	1.146	(0.101)	1.127	(0.069)	1.109	(0.042)	1.091	(0.023)	1.085	(0.017)	1.081	(0.014)
		LS X on Y	1.125	(0.159)	1.099	(0.078)	1.090	(0.054)	1.083	(0.034)	1.075	(0.019)	1.073	(0.014)	1.071	(0.012)
	20%	LS Y on X	1.288	(0.297)	1.233	(0.154)	1.213	(0.110)	1.195	(0.076)	1.174	(0.049)	1.165	(0.039)	1.161	(0.035)
		LS X on Y	1.202	(0.211)	1.174	(0.111)	1.166	(0.081)	1.161	(0.058)	1.153	(0.040)	1.150	(0.033)	1.149	(0.030)
	30%	LS Y on X	1.397	(0.451)	1.352	(0.264)	1.328	(0.197)	1.302	(0.142)	1.277	(0.101)	1.267	(0.087)	1.260	(0.079)
		LS X on Y	1.293	(0.297)	1.274	(0.177)	1.264	(0.137)	1.257	(0.106)	1.249	(0.082)	1.245	(0.074)	1.243	(0.069)
	40%	LS Y on X	1.553	(0.766)	1.496	(0.450)	1.467	(0.347)	1.440	(0.266)	1.410	(0.202)	1.396	(0.178)	1.389	(0.168)
		LS X on Y	1.417	(0.459)	1.392	(0.289)	1.383	(0.234)	1.378	(0.195)	1.371	(0.163)	1.367	(0.152)	1.365	(0.147)
	50%	LS Y on X	1.746	(1.670)	1.689	(0.813)	1.660	(0.659)	1.627	(0.517)	1.588	(0.400)	1.572	(0.361)	1.562	(0.340)
		LS X on Y	1.572	(0.764)	1.547	(0.491)	1.543	(0.421)	1.539	(0.365)	1.532	(0.320)	1.529	(0.305)	1.526	(0.296)
	60%	LS Y on X	2.042	(11.32)	1.974	(1.822)	1.939	(1.345)	1.900	(1.072)	1.847	(0.824)	1.828	(0.754)	1.810	(0.703)
		LS X on Y	1.798	(1.870)	1.766	(0.917)	1.764	(0.797)	1.764	(0.709)	1.758	(0.636)	1.759	(0.617)	1.753	(0.599)
	70%	LS Y on X	2.684	(556.1)	2.456	(21.81)	2.393	(4.501)	2.343	(2.771)	2.270	(1.915)	2.233	(1.700)	2.209	(1.585)
		LS X on Y	2.226	(44.48)	2.108	(2.081)	2.103	(1.695)	2.107	(1.496)	2.113	(1.371)	2.110	(1.321)	2.109	(1.297)
	80%	LS Y on X	29.78	(8858266)	3.726	(1357)	3.382	(63.09)	3.314	(590.9)	3.087	(6.173)	3.012	(5.018)	2.972	(4.498)
		LS X on Y	29.78	(8858266)	2.817	(34.28)	2.737	(5.779)	2.727	(4.032)	2.746	(3.504)	2.748	(3.350)	2.750	(3.286)
β_T=5	10%	LS Y on X	1.018	(0.005)	1.015	(0.003)	1.013	(0.002)	1.011	(0.001)	1.009	(0.001)	1.008	(0.000)	1.007	(0.000)
		LS X on Y	1.007	(0.005)	1.007	(0.003)	1.007	(0.002)	1.006	(0.001)	1.006	(0.001)	1.006	(0.000)	1.005	(0.000)
	20%	LS Y on X	1.024	(0.006)	1.020	(0.003)	1.019	(0.002)	1.017	(0.001)	1.014	(0.001)	1.013	(0.001)	1.013	(0.000)
		LS X on Y	1.013	(0.006)	1.012	(0.003)	1.012	(0.002)	1.012	(0.001)	1.011	(0.001)	1.011	(0.000)	1.011	(0.000)
	30%	LS Y on X	1.031	(0.007)	1.028	(0.004)	1.026	(0.003)	1.024	(0.002)	1.021	(0.001)	1.020	(0.001)	1.019	(0.001)
		LS X on Y	1.018	(0.006)	1.018	(0.003)	1.018	(0.002)	1.018	(0.002)	1.018	(0.001)	1.017	(0.001)	1.017	(0.001)
	40%	LS Y on X	1.039	(0.009)	1.037	(0.005)	1.035	(0.004)	1.032	(0.003)	1.030	(0.002)	1.029	(0.001)	1.028	(0.001)
		LS X on Y	1.025	(0.008)	1.026	(0.004)	1.026	(0.003)	1.026	(0.002)	1.026	(0.001)	1.026	(0.001)	1.025	(0.001)
	50%	LS Y on X	1.049	(0.011)	1.048	(0.007)	1.046	(0.005)	1.044	(0.004)	1.041	(0.003)	1.040	(0.002)	1.039	(0.002)
		LS X on Y	1.034	(0.010)	1.035	(0.005)	1.036	(0.004)	1.036	(0.003)	1.036	(0.002)	1.036	(0.002)	1.036	(0.002)
	60%	LS Y on X	1.062	(0.015)	1.063	(0.010)	1.061	(0.008)	1.059	(0.006)	1.056	(0.004)	1.054	(0.004)	1.054	(0.003)
		LS X on Y	1.046	(0.013)	1.048	(0.007)	1.049	(0.006)	1.050	(0.004)	1.050	(0.003)	1.049	(0.003)	1.049	(0.003)
	70%	LS Y on X	1.074	(0.022)	1.082	(0.014)	1.082	(0.012)	1.081	(0.010)	1.078	(0.008)	1.076	(0.007)	1.075	(0.006)
		LS X on Y	1.060	(0.019)	1.065	(0.011)	1.067	(0.009)	1.068	(0.007)	1.070	(0.006)	1.069	(0.006)	1.069	(0.005)
	80%	LS Y on X	1.081	(0.035)	1.109	(0.025)	1.115	(0.022)	1.114	(0.018)	1.113	(0.015)	1.111	(0.014)	1.110	(0.013)
		LS X on Y	1.081	(0.035)	1.090	(0.019)	1.096	(0.016)	1.098	(0.013)	1.101	(0.012)	1.101	(0.012)	1.101	(0.011)

† There are some "x" in the table which denote the omitted results as they are extremely large values.

4.4.3.3 Summary of Results

The following conclusions are made combining the results for α and β.

For Complete Data

1) LS Y on X is recommended for estimating β for very small samples, say $n \le 10$. LS X on Y is recommended for estimating β for medium to large samples, especially for $n \ge 30$.
2) LS X on Y is recommended for estimating α.

For Censored Data

1) LS Y on X is recommended for estimating β for samples with high censoring levels, say $c \ge 50\%$. LS X on Y is recommended for estimating β for samples with low censoring levels, say $c < 50\%$.
2) LS X on Y is recommended for estimating α.

4.5 Summary

This chapter presents several modifications or refinements on the OLSE method. Firstly, it was emphasized to use LSE with WPP in order to have a graphical presentation. Besides LSE, all the linear regression estimation methods should be used with WPP as the graphical presentations are always useful for practitioners.

Two problems intrinsic to OLSE were examined. One is the determination of Y-axis plotting positions. The existing plotting positions in the cases of complete data and censored data, respectively, were summarized and analyzed in different groups. Via intensive Monte Carlo experiments, selected plotting positions with the focus on those proposed in recent years and have not received much attention, were compared on the estimation of two Weibull parameters. The results showed that the Ross estimator is a promising one for complete data. For censored data, HJ should be preferred for samples with high censoring levels while JM is good for samples with low censoring levels. However, it should be noted that none of the existing estimators outperforms the others for all the cases.

Another intrinsic problem of the LSE method is the direction of regression. Two methods, i.e., LS Y on X and LS X on Y were compared for both complete data and censored data. In terms of model statistics, it was found that LS X on Y should be used when $\beta > 1$ and LS Y on X should be used when $\beta < 1$. In terms of parameter estimation, the simulation results have provided suggestions on when to use which method, as listed in Section 4.4.3.3.

Chapter 5

Bias Correction Methods for the Shape Parameter Estimator of OLSE

This chapter presents the bias correction methods for the OLS estimated Weibull shape parameter in the cases of complete data and censored data. Several bias correcting formulas are proposed which can be used in the end of the OLS estimation procedure to correct the bias of the shape parameter estimator. The proposed methods are easy to use and can effectively reduce the bias.

5.1 Introduction

Bias is often an important issue of the estimator in the sense that it tells us whether the estimator is an accurate estimate value of the population value. As one of the most commonly used criteria to compare different estimation methods, the issue of bias has raised the attention of Weibull researchers. In the 1990s, many researchers pointed out that the estimators of the MLE method are significantly biased when the sample size is small, among them, Ross (1994a) mentioned that 'the frequently use of small-size samples of life tests, e.g., $n = 5$, where n is the sample size, can give significant support to the investigation of unbiasing procedures'. Indeed, several bias correction methods for the MLE have been proposed. Jacquelin (1993) modified the estimating equation of the MLE method by adding two parameters which are calculated as the functions of failure probability F_i. The method is named generalized MLE and is claimed to directly provide unbiased estimates without the aid of unbiasing factors.

Ross (1994a, 1996), in a different approach as Jacquelin's, proposed simple models of unbiasing factors for the MLE of the shape parameter, applied to complete data and censored data, respectively. The theoretical justification of the Ross' bias correction method is based on the first pivotal function of the MLE of the Weibull parameters, i.e., $\hat{\beta}/\beta$. With a similar theoretical background, Hirose (1999) provided another bias correcting model for the MLE of the shape parameter and it has a polynomial form. The unbiasing for the MLE of the scale parameter was also examined, and different formulas were provided at selected β values. Besides, Cacciari, Montanari, Mazzanti and Fothergill co-published a series of work (Cacciari et al., 1996; Montanari et al., 1997a, b, 1998) that compared several bias correction methods including the method of Engelhardt & Bain (1974), Jacquelin (1993), Ross (1994a, 1996), White's weighted least squares technique (White, 1969), etc., together with the conventional LSE method and the MLE method for both complete and censored data using the Monte Carlo method.

The values of bias of the Weibull parameter estimators can be obtained via the Monte Carlo method. While many researchers are keen on the bias correction for the MLE of the Weibull parameters, less has been discussed on the bias correction for the estimators of LSE. In fact, some researchers have pointed out that the OLSE of the shape parameter is less biased than that of the MLE for small samples, see, e.g., Ross (1999). This may hide the need for bias correction for this method; however, as shown in Chapter 3 that the OLSE of the shape parameter is biased and from Section 3.3.3, it can be observed that the OLS shape parameter estimator is not always satisfactory in view of bias, for example, the result for complete samples shows that during the sample sizes 11 – 30, there is always a relative bias of around 4%; and for censored sample, the relative bias is more than 10% at the combination of $c = 10\%$ and $n = 20$. As a result, simple bias correction methods for the OLSE will be helpful.

As shown in Section 4.3, the bias of the LSE of the shape parameter varies with the selection of the estimators for $F(t)$, for both complete and censored data. The simulation results presented in Section 4.3.3 have shown that the expected estimators including the Ross estimator (Ross, 1994b) and the Drap-Kos estimator (Drapella & Kosznik, 1999) can greatly reduce the bias of the LSE of the shape parameter in the case of complete data. This can be treated as one way to correct the bias for the LSE method. This chapter presents another kind of bias correction method which provides the unbiasing factors. The empirical bias correcting formulas are proposed and can be added to the end of the OLSE procedure to reduce bias.

This chapter is organized as follows. In Section 5.2, the theoretical justification for the existence of a single bias correcting formula for the OLSE of the shape parameter is presented. Section 5.3 presents the bias correction methods for the OLSE of the shape parameter applied to complete data. Firstly, the relationship between the bias of the OLS shape parameter estimator and sample size in the case of complete data is examined. Then, based on the relationship, the models of bias correcting factors are proposed and the model parameters are determined via numerical methods. Finally, the bias correcting formulas are presented as well as the application procedure. The proposed methods are named the modified Ross' method and the modified Hirose's method, respectively. Section 5.4 discusses the bias correction methods for the shape parameter estimator of LS X on Y applied to complete data and the shape parameter estimator of OLSE applied to multiply censored samples, respectively. A bias correcting formula is proposed for each condition. Some of the related work has been published in Zhang et al. (2006).

5.2 Theoretical Background of Bias Correction

The existence of the pivotal function $\hat{\beta}/\beta$, of the ML or LS estimated $\hat{\beta}$ of the Weibull distribution, makes the bias correction a simple job. Proof for the pivotal functions of the LSE is described in Section 3.2.5.

The pivotal function $\hat{\beta}/\beta$ says the following relationship,

$$E(\hat{\beta}/\beta) = E(\hat{\beta}_{1,1}) \tag{5-1}$$

or

$$E(\hat{\beta}) = \beta \cdot E(\hat{\beta}_{1,1}) \tag{5-2}$$

Now define an estimator $\hat{\beta}_U$ as

$$\hat{\beta}_U = \frac{\hat{\beta}}{E(\hat{\beta}_{1,1})} \tag{5-3}$$

Then the expected value of $\hat{\beta}_U$ is

$$E(\hat{\beta}_U) = \frac{E(\hat{\beta})}{E(\hat{\beta}_{1,1})} \tag{5-4}$$

Based on Equation (5-2), we have

$$E(\hat{\beta}_U) = \frac{\beta \cdot E(\hat{\beta}_{1,1})}{E(\hat{\beta}_{1,1})} = \beta \tag{5-5}$$

Therefore, $\hat{\beta}_U$ is the unbiased estimator of β.

The relationship between $\hat{\beta}$ and $\hat{\beta}_U$ can be expressed by an unbiasing factor U which satisfies

$$\hat{\beta}_U = U \cdot \hat{\beta} \tag{5-6}$$

Then from Equation (5-4), U can be determined as

$$U = \frac{1}{E(\hat{\beta}_{1,1})} \tag{5-7}$$

Since the values of $E(\hat{\beta}_{1,1})$ can be obtained via the Monte Carlo method, the values of the unbiasing factor U can also be determined.

As shown above, the bias correction for $\hat{\beta}$ is clearly independent of the true values of α and β, and a single formula, i.e., Equation (5-3), can work for any data set. This is not true for the scale parameter estimator $\hat{\alpha}$ because $\hat{\alpha}/\alpha$ is not a pivotal function. The bias correction for $\hat{\alpha}$ requires different formulas at different values of β. Since $\hat{\beta}$ is often of great importance, the bias correction for $\hat{\alpha}$ is not discussed in this chapter.

Without further examination, a traditional way of bias correction, e.g., in the case of complete data, is to tabulate the values of $E(\hat{\beta}_{1,1})$ or U at different sample sizes via the Monte Carlo method. The tabulation generates a reference table. Thus given a random data set, a look-up in the table using the sample size is needed to find the value of the unbiasing factor so that the unbiased estimate of the shape parameter can be calculated by Equation (5-6). It is noteworthy that Equation (5-6) and Equation (5-7) can also be applied to the LSE of the shape parameter for censored data. In the case of censored data, a reference table should show the values of the unbiasing factor at different combinations of sample sizes and censoring levels.

Obviously, the look-up method is inconvenient because it is troublesome or impossible to tabulate the unbiasing factors at all sample sizes or all combinations of sample sizes and censoring levels. A clearly better approach is to examine the pattern of the unbiasing factors and use analytical models.

5.3 Bias Correction for the OLSE of the Shape Parameter for Complete Data

As previously mentioned, the values of $E(\hat{\beta}_{1,1})$ of the OLS shape parameter estimator at different sample sizes can be obtained via the Monte Carlo method. For this purpose, a Monte Carlo simulation experiment was carried out. Table 5-1 lists the setting of experiment factors.

Table 5-1: Setting of experiment factors. The experiment is to examine the trends of the bias of the OLS and MLE estimated β for complete data as a function of sample size.

Factors	Values
α_T	1
β_T	1
n	3, 4, …, 19, 20, 22,…, 28, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100
M	50000
Methods	OLSE, MLE

50000 random samples were generated for each sample size and the parameter estimates were obtained from OLSE and MLE simultaneously. $E(\hat{\beta}_{1,1})$ is calculated by the average of the parameter estimates. Bias is calculated by the difference between $E(\hat{\beta}_{1,1})$ and 1. Figure 5-1 shows the bias of $\hat{\beta}_{1,1}$, obtained from both OLSE and MLE, at each sample size investigated. The simulation results can also be extracted from Section 3.3.3.1. From the figure we can see that, although the bias of the OLSE of the shape parameter is much smaller than that of the MLE of the shape parameter for small to medium sized samples (say $n \le 30$), OLSE still considerably overestimates the shape parameter for extremely small samples (say $n = 3$ and 4), and the bias keeps at around 4% during $10 < n \le 30$. Therefore, simple bias correction methods will be helpful for the OLS shape parameter estimator, especially for very small samples.

The shapes of the two curves in Figure 5-1 are similar and both have a hyperbolic shape. This suggests that the bias correcting models of the MLE may be used for the OLSE as well. Following this idea, the unbiasing formulas proposed by Ross (1994a, 1996) and Hirose (1999), respectively, for the MLE were modified for the OLSE and the proposed methods are named the modified Ross' method and the modified Hirose's method.

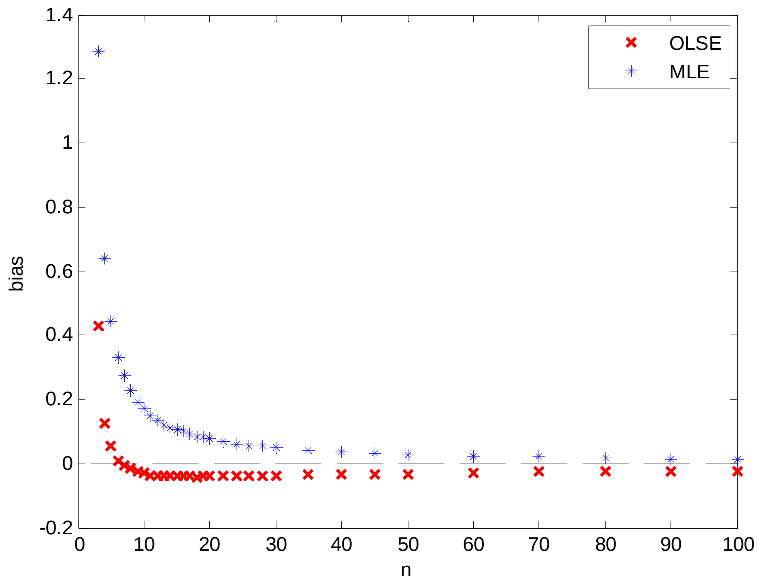


Figure 5-1: Bias of $\hat{\beta}_{1,1}$, obtained by OLSE and MLE, as a function of sample size.

5.3.1 Modified Ross' Bias Correction Method

Ross' Bias Correction Method for the MLE of the Shape Parameter

Ross (1994a) proposed an asymptotic difference function to model the bias of the ML estimated $\hat{\beta}_{1,1}$. The asymptotic difference function, denoted by $D(n)$, is defined as the difference between the expected value of the estimator from finite and infinite sample size. From the definition,

$$D(n) = E(\hat{\beta}_{1,1}) - 1 \tag{5-8}$$

$D(n)$ was then modelled as a power function of n with three parameters: the threshold parameter R, the power parameter P and the proportionality constant Q. Ross' bias correcting factor, denoted by U_R, is given by

$$U_R = \frac{1}{E(\hat{\beta}_{1,1})} = \frac{1}{1+D(n)} = \frac{1}{1+Q(n-R)^P} \tag{5-9}$$

With the values of $E(\hat{\beta}_{1,1})$ of the MLE derived via the Monte Carlo method at various sample sizes, the three parameters P, Q, R were determined by the author using both graphical and numerical methods. The results are $P = -1, Q = 1.32, R = 2$.

Finally, Ross' bias correcting formula for the MLE of the shape parameter was determined which has a very simple form, i.e.,

$$\beta_U = \hat{\beta} \cdot U_R = \hat{\beta} \cdot \frac{n-2}{n-0.68} \tag{5-10}$$

Ross concluded that the bias of the MLE of the shape parameter can be reduced to typically <0.3% for $n \geq 3$ if the proposed formula is used.

Modified Ross' Bias Correction Method for the OLSE of the Shape Parameter

Theoretically, to use Ross' asymptotic difference function and bias model for the LSE, the following assumptions have to be satisfied:

i. $\hat{\beta}/\beta$ is a pivotal function for the LS estimated shape parameter.

ii. The expected value of the LS shape parameter estimator approaches to an asymptotic value, i.e., the true value of β, when $n \to \infty$.

There is no doubt that both assumptions are true for the OLSE of the shape parameter. For assumption ***ii***, from Figure 5-1 it can be seen that, although the OLSE of the shape parameter is inconsistent (the bias reaches 0 when n is around 6 or 7), it still approaches to the true value when n becomes large. Since the two assumptions are satisfied, the modified Ross' bias correction method is proposed for the OLSE of the shape parameter, as presented in the following. It is mainly designed for small samples of size ≤ 20.

Ross' unbiasing factor U_R has three parameters P, Q, R. The condition $n - R > 0$ is set by the author. Thus, $U_R > 1$ when $Q < 0$ and $U_R < 1$ when $Q > 0$. It is impossible to have $U_R > 1$ for some sample sizes and $U_R < 1$ for other sample sizes because a single value of Q is required. Actually, the values of U_R are always less than 1 because the values of $E(\hat{\beta}_{1,1})$ obtained by MLE are always larger than 1, as can be seen from Figure 5-1. However, this is not applicable to the OLSE. From Figure 5-1 we can see that, the OLS shape parameter estimator needs a bias correcting factor whose values are less than 1 when $n < 7$, and larger than 1 when $n \geq 7$. Therefore, Ross' bias correcting factor U_R is not efficient for the OLSE. It can be improved by introducing a new parameter C_a. We name C_a the adjusting constant as it works to adjust the values of the unbiasing factor to be greater than 1 or less than 1.

The modified bias correcting factor for the OLSE of the shape parameter, denoted by U_{MR}, is proposed as

$$U_{MR} = \frac{1}{1+Q(n-R)^P} + C_a \tag{5-11}$$

The four parameters P, Q, R, C_a in U_{MR} were determined by using the unconstrained nonlinear optimization, e.g., Nelder-Mead direct search method (Nelder & Mead, 1965). The objective function is

$$\min \sum_{n_i} \left[1 - U_{MR}(n_i) \cdot E(\hat{\beta}_{1,1})_{n_i}\right]^2 = \sum_{n_i} \left[1 - \left(\left(1 + Q(n_i - R)^P\right)^{-1} + C_a\right) \cdot E(\hat{\beta}_{1,1})_{n_i}\right]^2 \tag{5-12}$$

where n_i denotes different sample sizes and $E(\hat{\beta}_{1,1})_{n_i}$ denotes the value of $E(\hat{\beta}_{1,1})$ at a specific n_i.

The values of $E(\hat{\beta}_{1,1})_{n_i}$ of the OLSE, obtained from the Monte Carlo experiment at $n_i = 3,4 \cdots, 19, 20, 22, \cdots, 28, 30, 35, \cdots, 45, 50, 60, \cdots, 90, 100$ (same as Table 5-1), were used to determine the values of P, Q, R, C_a. Different starting values for the parameters were tried in the Nelder-Mead direct search method. The calculation was executed by MATLAB 7 and the function *fminsearch* was used. The current result satisfies the termination criteria using OPTIONS.TolX of 1.000000e-001 and satisfies the convergence criteria using OPTIONS.TolFun of 1.000000e-008.

The values for the parameters were determined as

$$P = -2.1, Q = 1.4, R = 1.4, C_a = 0.05 \tag{5-13}$$

Substituting the values of P, Q, R, C_a into Equation (5-11), the bias correcting factor U_{MR} for the OLSE of the shape parameter is

$$U_{MR} = \frac{1}{1+1.4(n-1.4)^{-2.1}} + 0.05 \qquad (5\text{-}14)$$

Thus the bias correcting formula of the modified Ross' method for the OLSE of the shape parameter is

$$\hat{\beta}_U = \hat{\beta} \cdot U_{MR} = \hat{\beta} \cdot \left[\frac{1}{1+1.4(n-1.4)^{-2.1}} + 0.05 \right] \qquad (5\text{-}15)$$

Table 5-2 tabulates the values of $E(\hat{\beta}_{1,1})$, U_{MR} and $E_{U,MR}(\hat{\beta}_{1,1}) = E(\hat{\beta}_{1,1}) \cdot U_{MR}$ at selected sample sizes. As can be seen from the table, the differences between $E_{U,MR}(\hat{\beta}_{1,1})$ and 1 are less than the differences between $E(\hat{\beta}_{1,1})$ and 1 at all sample sizes. Especially at $n = 3$ and $n = 4$, the bias is significantly reduced. The bias is within 1% and typically within 0.5% during $n = 6 - 30$.

Table 5-2: Values of $E(\hat{\beta}_{1,1})$, U_{MR} and $E_{U,MR}(\hat{\beta}_{1,1})$ at selected sample sizes (the modified Ross' method for OLSE).

	n												
	3	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
$E(\hat{\beta}_{1,1})$	1.428	1.125	1.053	1.009	0.996	0.983	0.974	0.970	0.963	0.961	0.961	0.960	0.960
U_{MR}	0.707	0.892	0.963	0.996	1.014	1.024	1.031	1.035	1.038	1.040	1.042	1.043	1.044
$E_{U,MR}(\hat{\beta}_{1,1})$	1.010	1.003	1.014	1.005	1.010	1.006	1.003	1.004	1.000	1.000	1.001	1.001	1.002
	16	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**	**30**	**35**	**40**	**50**
$E(\hat{\beta}_{1,1})$	0.962	0.959	0.958	0.960	0.960	0.960	0.959	0.959	0.962	0.961	0.964	0.966	0.966
U_{MR}	1.045	1.046	1.046	1.047	1.047	1.048	1.048	1.048	1.049	1.049	1.049	1.049	1.050
$E_{U,MR}(\hat{\beta}_{1,1})$	1.006	1.003	1.002	1.005	1.005	1.006	1.005	1.006	1.009	1.008	1.012	1.013	1.014

To further check the proposed unbiasing formula in Equation (5-15) for a single Weibull sample, another Monte Carlo experiment was conducted. Normalized Weibull samples (i.e., $\alpha_T = \beta_T = 1$) of sizes 3 – 50 were randomly generated. For each sample, OLSE was used to estimate the parameters first, and then Equation (5-15) was applied to the OLS estimated shape parameter to generate the unbiased estimate. Both estimates of the shape parameter, i.e., with and without unbiasing, were recorded. 10000 iteration was used at each sample size and the average values of the estimates were calculated as $E(\hat{\beta}_{1,1})$ and $E(\hat{\beta}_{U_{1,1},MR})$, respectively.

The results are shown in Table 5-3. It can be observed from the table that the bias of $\hat{\beta}_{U_{1,1},MR}$ is significantly smaller than the bias of $\hat{\beta}_{1,1}$, especially at $n=3$. The bias of $\hat{\beta}_{1,1}$ is typically 4%, while the bias of $\hat{\beta}_{U_{1,1},MR}$ is typically within 1%.

Table 5-3: Simulation results of the modified Ross' method: the values of $E(\hat{\beta}_{1,1})$ and $E(\hat{\beta}_{U_{1,1},MR})$ at selected sample size‡.

	n												
	3	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
$E(\hat{\beta}_{1,1})$	1.436	1.131	1.051	1.013	0.991	0.975	0.976	0.969	0.963	0.965	0.965	0.963	0.961
$E(\hat{\beta}_{U_{1,1},MR})$	1.015	1.009	1.013	1.009	1.005	0.998	1.006	1.003	1.000	1.004	1.006	1.005	1.004
	16	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**	**30**	**35**	**40**	**50**
$E(\hat{\beta}_{1,1})$	0.961	0.956	0.958	0.961	0.962	0.960	0.959	0.962	0.961	0.964	0.965	0.964	0.968
$E(\hat{\beta}_{U_{1,1},MR})$	1.004	1.000	1.000	1.006	1.008	1.005	1.005	1.006	1.007	1.011	1.013	1.012	1.016

Figure 5-2 shows the histograms or the empirical PDFs of $\hat{\beta}_{U_{1,1},MR}$ at selected sample sizes: $n=5,10,20,30$. The estimates of the 10000 samples at each sample size from the experiment were used to generate the histograms. As can be seen, the distribution of $\hat{\beta}_{U_{1,1},MR}$ approaches to the normal distribution as the sample size increases. It can also be observed that the mean of the distribution is very close to 1.

‡ The values of $E(\hat{\beta}_{1,1})$ are slightly different compared to the values in Table 5-2 at same sample size because here the simulation iteration number is reduced to 10000.

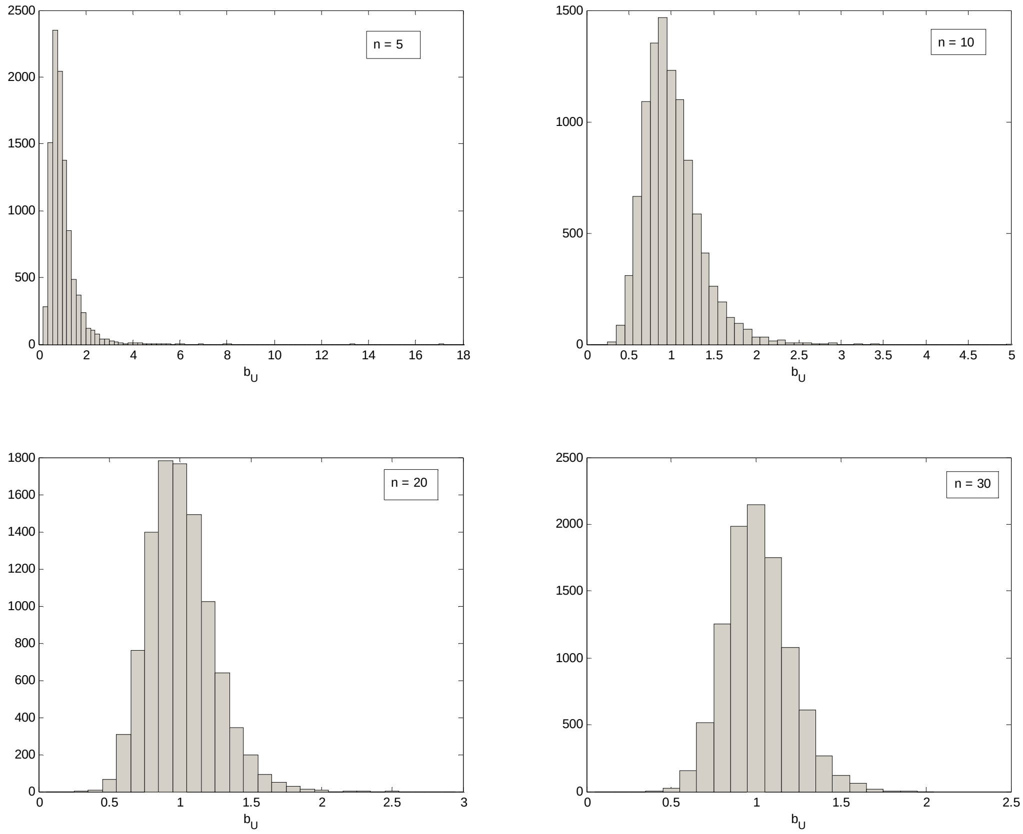


Figure 5-2: Histograms of $\hat{\beta}_{U_{1,1},MR}$ at selected sample sizes (the modified Ross' method for OLSE).

5.3.2 Modified Hirose's Bias Correction Method

<u>Hirose's Bias Correction Method for the MLE of the Shape Parameter</u>

Instead of modeling the unbiasing factor, Hirose (1999) proposed a function for modeling the bias of the MLE of the shape parameter, given by

$$B_n(\hat{\beta}_{1,1}) = k_0 + \frac{k_1}{n} + \frac{k_2}{n^2} + \cdots \frac{k_i}{n^i} + \cdots \tag{5-16}$$

where $B_n(\cdot)$ denotes the bias function as a function of n.

For simplicity, Hirose suggested using the approximation, i.e.,

$$B_n(\hat{\beta}_{1,1}) \approx k_0 + \frac{k_1}{n} + \frac{k_2}{n^2} + \frac{k_3}{n^3} + \frac{k_4}{n^4} \tag{5-17}$$

where k_0, k_1, k_2, k_3, k_4 are the model parameters.

Based on the values of $B_n(\hat{\beta}_{1,1})$ of the MLE obtained by the Monte Carlo method at various sample sizes, Hirose determined the values of the five parameters in Equation (5-17). The results are $k_0 = 0.0115, k_1 = 1.278, k_2 = 2.001, k_3 = 20.35$ and $k_4 = -49.68$. Thus the bias model is determined as

$$B_n(\hat{\beta}_{1,1}) \approx 0.0115 + \frac{1.278}{n} + \frac{2.001}{n^2} + \frac{20.35}{n^3} - \frac{49.68}{n^4} \tag{5-18}$$

Finally, Hirose's bias correcting formula for the MLE of the shape parameter is given by

$$\hat{\beta}_U = \frac{\hat{\beta}}{E(\hat{\beta}_{1,1})} = \frac{\hat{\beta}}{1 + B_n(\hat{\beta}_{1,1})} \approx \frac{\hat{\beta}}{1.0115 + \frac{1.278}{n} + \frac{2.001}{n^2} + \frac{20.35}{n^3} - \frac{49.68}{n^4}} \tag{5-19}$$

<u>Modified Hirose's Bias Correction Method for the OLSE of the Shape Parameter</u>

Hirose's bias model in Equation (5-17) uses the polynomial curve fitting technique. As previously mentioned, the trends of $E(\hat{\beta}_{1,1})$ vs. n of the MLE and the OLSE are similar, and both have a hyperbolic appearance (see Figure 5-1). Therefore, the Hirose's model can be applied to propose the unbiasing formula for the OLSE of the shape parameter.

Same as the modified Ross' method, we first determine the formula of the unbiasing factor U. Obviously, the trend of the unbiasing factor U vs. n should also have a hyperbolic shape. Therefore, the proposed model for the unbiasing factor of the modified Hirose's bias correction method, denoted by U_{MH}, is

$$U_{MH} \approx l_0 + \frac{l_1}{n} + \frac{l_2}{n^2} + \frac{l_3}{n^3} + \frac{l_4}{n^4} \tag{5-20}$$

where l_0, l_1, l_2, l_3, l_4 are the model parameters.

As in the modified Ross' method, the values of l_i were determined through the unconstrained nonlinear optimization technique (Nelder-Mead direct search method). The objective function is

$$\min \sum_{n_i} \left[1 - U_{MH}(n_i) \cdot E(\hat{\beta}_{1,1})_{n_i}\right]^2 = \sum_{n_i} \left[1 - \left(l_0 + \frac{l_1}{n_i} + \frac{l_2}{n_i^2} + \frac{l_3}{n_i^3} + \frac{l_4}{n_i^4}\right) \cdot E(\hat{\beta}_{1,1})_{n_i}\right]^2 \tag{5-21}$$

The calculation was executed by MATLAB 7 and the function *fminsearch* was used. Different starting values for l_i were tested in the Nelder-Mead direct search method. The current result satisfies the termination criteria using OPTIONS.TolX of 1.000000e-004 and satisfies the convergence criteria using OPTIONS.TolFun of 1.000000e-008.

The parameter values were determined as

$$l_0 = 1.0357, l_1 = 0.3082, l_2 = -3.6347, l_3 = 2.4386, l_4 = -10.0430 \tag{5-22}$$

Substituting the values into Equation (5-20), the formula of U_{MH} is

$$U_{MH} \approx 1.0357 + \frac{0.3082}{n} - \frac{3.6347}{n^2} + \frac{2.4386}{n^3} - \frac{10.0430}{n^4} \tag{5-23}$$

Thus the bias correcting formula of the modified Hirose's method for the OLSE of the shape parameter is

$$\hat{\beta}_U \approx \hat{\beta} \cdot \left(1.0357 + \frac{0.3082}{n} - \frac{3.6347}{n^2} + \frac{2.4386}{n^3} - \frac{10.0430}{n^4}\right) \tag{5-24}$$

Table 5-4 tabulates the values of $E(\hat{\beta}_{1,1})$, U_{MH} and $E_{U,MH}(\hat{\beta}_{1,1}) = E(\hat{\beta}_{1,1}) \cdot U_{MH}$ at selected sample sizes. As can be seen from the table, the differences between $E_{U,MH}(\hat{\beta}_{1,1})$ and 1 are much smaller than the differences between $E(\hat{\beta}_{1,1})$ and 1 at all sample sizes. Great improvements can be observed when $n = 3, 4$. In addition, comparing Table 5-4 with Table 5-2, we can see that $E_{U,MH}(\hat{\beta}_{1,1})$ is slightly better than $E_{U,MR}(\hat{\beta}_{1,1})$ in most cases.

Table 5-4: Values of $E(\hat{\beta}_{1,1})$, U_{MH} and $E_{U,MH}(\hat{\beta}_{1,1})$ at selected sample sizes (the modified Hirose's method for OLSE).

	n												
	3	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
$E(\hat{\beta}_{1,1})$	1.428	1.125	1.053	1.009	0.996	0.983	0.974	0.970	0.963	0.961	0.961	0.960	0.960
U_{MH}	0.701	0.884	0.955	0.990	1.008	1.020	1.027	1.032	1.035	1.037	1.039	1.040	1.041
$E_{U,MH}(\hat{\beta}_{1,1})$	1.001	0.995	1.006	0.998	1.005	1.002	1.000	1.001	0.997	0.996	0.998	0.998	0.999
	16	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**	**30**	**35**	**40**	**50**
$E(\hat{\beta}_{1,1})$	0.962	0.959	0.958	0.960	0.960	0.960	0.959	0.959	0.962	0.961	0.964	0.966	0.966
U_{MH}	1.041	1.042	1.042	1.042	1.042	1.042	1.042	1.042	1.042	1.042	1.042	1.041	1.040
$E_{U,MH}(\hat{\beta}_{1,1})$	1.002	0.999	0.998	1.001	1.000	1.001	1.000	1.000	1.002	1.002	1.005	1.005	1.005

The modified Hirose's unbiasing formula in Equation (5-24) was also checked for point estimation in the experiment described in the end of the modified Ross' method. Table 5-5 tabulates the expected values of the shape parameter estimates before and after correction at selected sample sizes. Figure 5-3 shows the histograms or the empirical PDFs of $\hat{\beta}_{U,MH}$ at selected sample sizes: $n = 5, 10, 20, 30$.

Table 5-5: Simulation results of the modified Hirose's method: the values of $E(\hat{\beta}_{1,1})$ and $E(\hat{\beta}_{U_{1,1},MH})$ at selected sample size‡.

	n												
	3	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
$E(\hat{\beta}_{1,1})$	1.436	1.131	1.051	1.013	0.991	0.975	0.976	0.969	0.963	0.965	0.965	0.963	0.961
$E(\hat{\beta}_{U_{1,1},MH})$	1.006	1.001	1.004	1.002	1.000	0.994	1.003	0.999	0.997	1.001	1.003	1.001	1.001
	16	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**	**30**	**35**	**40**	**50**
$E(\hat{\beta}_{1,1})$	0.961	0.956	0.958	0.961	0.962	0.960	0.959	0.962	0.961	0.964	0.965	0.964	0.968
$E(\hat{\beta}_{U_{1,1},MH})$	1.000	0.996	0.998	1.002	1.003	1.000	1.000	1.002	1.001	1.005	1.005	1.004	1.007

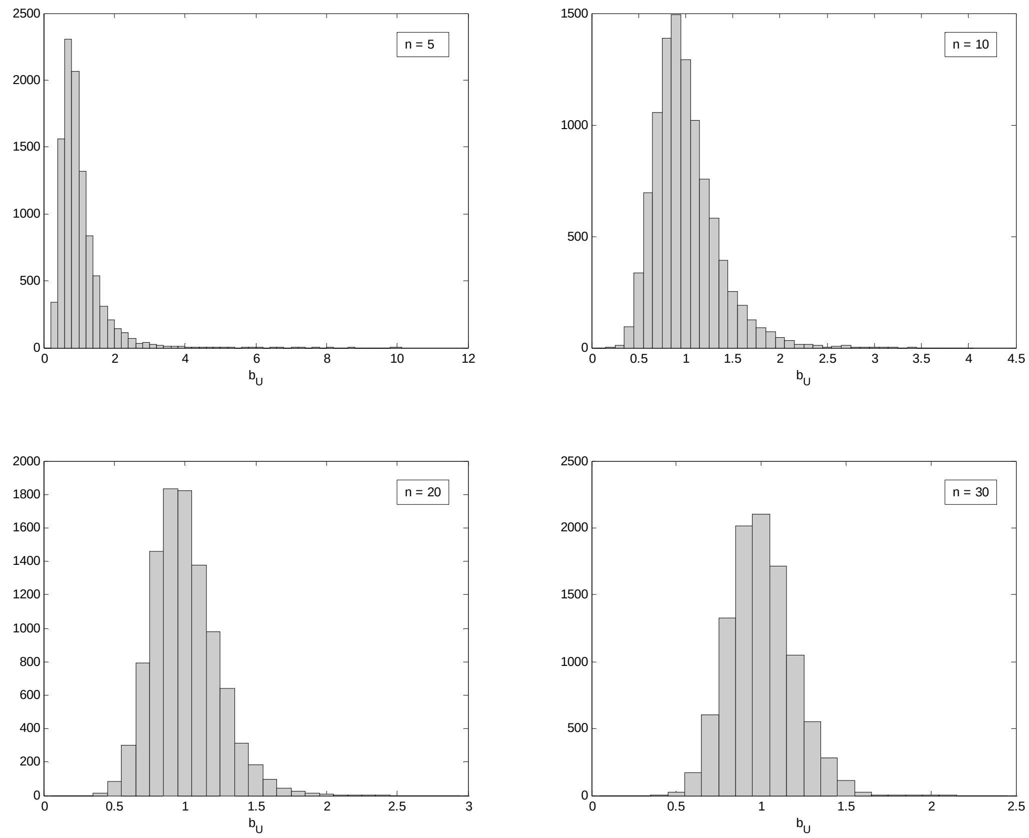


Figure 5-3: Histograms of $\hat{\beta}_{U_{1,1},MH}$ at selected sample sizes (the modified Hirose's method for OLSE).

It can be observed from Table 5-5 that $\hat{\beta}_{U_{1,1},MH}$ is significantly better than $\hat{\beta}_{1,1}$ in view of the bias. The bias of $\hat{\beta}_{U_{1,1},MH}$ is typically within 0.3%. Compared to Table 5-3, we can see that $\hat{\beta}_{U_{1,1},MH}$ of the modified Hirose's method is slightly more accurate than $\hat{\beta}_{U_{1,1},MR}$ of the modified Ross' method in most cases.

Figure 5-3 looks similar to Figure 5-2. As can be seen, the distribution of $\hat{\beta}_{U_{1,1},MH}$ approaches to the normal distribution as the sample size increases. The mean of the distribution is close to 1 at all sample sizes.

5.3.3 Application Procedure

Given a complete data set of size n, the procedure for obtaining an unbiased OLS estimate of the shape parameter is as follows:

Step 1: Rank the failure times t_i from smallest to largest and calculate the *Y*-axis plotting positions by $\hat{F}_{(i)} = (i-0.3)/(n+0.4)$, i.e., the Bernard estimator.

Step 2: Plot the ranked failure times $t_{(i)}$ against $\hat{F}_{(i)}$ on WPP. If the Weibull distribution fits, the data points should appear to be on a straight line.

Step 3: Estimate the shape parameter by the OLSE method using Equation (2-12).

Step 4: Calculate the unbiased estimate for the shape parameter by the modified Ross' unbiasing formula in Equation (5-15), or the modified Hirose's unbiasing formula in Equation (5-24).

5.3.4 A Numerical Example

Below is a randomly generated Weibull sample with $\alpha = 1000, \beta = 2$ and $n = 10$:
2230, 1057, 573.6, 617.5, 544, 940.5, 1672, 1427, 405.2, 698.9.

First calculate the estimate of the shape parameter by the OLSE method and the result is 1.923. If MLE is used, the shape parameter estimate is 1.963. Second, apply the proposed bias correcting formulas to the shape parameter estimate of the OLSE. If the modified Ross' method is used, we have

$$\hat{\beta}_{U,MR} = \hat{\beta} \cdot \left[\frac{1}{1+1.4(n-1.4)^{-2.1}} + 0.05 \right] = 1.923 \times \left[\frac{1}{1+1.4(10-1.4)^{-2.1}} + 0.05 \right] = 1.990$$

Otherwise, if the modified Hirose's method is used, the unbiased estimate is

$$\hat{\beta}_{U,MH} = \hat{\beta} \cdot \left(1.0357 + \frac{0.3082}{n} - \frac{3.6347}{n^2} + \frac{2.4386}{n^3} - \frac{10.0430}{n^4}\right)$$
$$= 2.238 \times \left(1.0357 + \frac{0.3082}{10} - \frac{3.6347}{10^2} + \frac{2.4386}{10^3} - \frac{10.0430}{10^4}\right) = 1.984$$

The unbiased estimates for the MLE can also be obtained using the original Ross' method and the Hirose's method. The results are 1.685 and 1.671, respectively.

In this example, the best estimate for the shape parameter is obtained using the modified Ross' method. For the MLE, the estimates after bias correction were worse than the original estimate. After all, the bias correction methods will work in the long run but may not work for a single sample.

5.4 Discussions on Bias Correction for the LSE in Other Circumstances

The bias correcting formulas presented in the previous section are specially designed for the OLSE of the shape parameter and are only applicable to complete data. The OLSE method limits the use of the Bernard estimator for estimating $F(t)$ for complete data, and the regression direction of $Y = \ln[-\ln(1-F)]$ on $X = \ln T$. If any of these two conditions is changed, a new bias correcting formula is needed. The same approach as the modified Ross' method or the modified Hirose's method can be used to derive the new bias correcting formulas. Section 5.4.1 presents the bias correcting formulas for the LS X on Y method applied to complete data.

It is also important to deal with the bias of the OLSE in the case of censored data, and a study is presented in Section 5.4.2. The bias as a function of the sample size and censoring level is shown by the 3-D surface plot, and it reflects the difficulty of proposing a single model for the bias due to the inconsistency. However, a bias correcting formula is proposed for multiply censored samples with $c \le 40\%$ and $n \le 100$.

5.4.1 Bias Correction for the Shape Parameter Estimator of LS *X* on *Y* for Complete Data

The LS *X* on *Y* method is presented in Section 4.4 and compared with the LS *Y* on *X* method. The simulation results in Section 4.4.3.1 show that for complete samples with $n \le 10$, $\hat{\beta}$ of the LS *X* on *Y* method always has larger bias than that of the LS *Y* on *X* method. There is certainly a need to correct the bias with the recent focus of Weibull researchers on small samples. In addition, from Figure 4-23 we can see that the curve of the LS *X* on *Y* looks smoother than that of the LS *Y* on *X*, which implies a higher efficiency of the potential bias correcting formula.

The modified Ross' method and the modified Hirose's method were used to propose two bias correcting formulas for the shape parameter estimator of the LS *X* on *Y* method. The procedures for developing these two formulas are similar to those described in Section 5.3.1 and Section 5.3.2, and hence are not repeated here.

The bias correcting formula of the modified Ross' method for the shape parameter estimator of LS *X* on *Y* is

$$\hat{\beta}_{U,MR(LS-XY)} = \hat{\beta}_{LS-XY} \cdot \left[\frac{1}{1+0.6(n-2)^{-1.1}} + 0.01 \right] \tag{5-25}$$

The bias correcting formula of the modified Hirose's method for the shape parameter estimator of LS *X* on *Y* is

$$\hat{\beta}_{U,MH(LS-XY)} \approx \hat{\beta}_{LS-XY} \cdot \left(1.0096 - \frac{0.2470}{n} - \frac{4.0751}{n^2} + \frac{12.0084}{n^3} - \frac{23.3542}{n^4}\right) \quad (5\text{-}26)$$

The values of $E(\hat{\beta}_{1,1})$ obtained from the Monte Carlo simulations, the unbiasing factor U from Equation (5-26) and Equation (5-27), and $E_U(\hat{\beta}_{1,1}) = E(\hat{\beta}_{1,1}) \cdot U$, at selected sample sizes, for the LS *X* on *Y* method, are shown in Table 5-6 and Table 5-7. As can be observed from both tables, the bias of $\hat{\beta}_{1,1}$ after correction is generally smaller, especially when $n \le 10$.

Table 5-6: Values of $E(\hat{\beta}_{1,1})$, U_{MR} and $E_{U,MR}(\hat{\beta}_{1,1})$ at selected sample sizes (the modified Ross' method for LS *X* on *Y*).

	n												
	3	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
$E(\hat{\beta}_{1,1})$	1.585	1.266	1.171	1.118	1.089	1.072	1.062	1.048	1.035	1.035	1.031	1.030	1.024
U_{MR}	0.635	0.791	0.858	0.894	0.917	0.933	0.944	0.953	0.959	0.965	0.969	0.972	0.976
$E_{U,MR}(\hat{\beta}_{1,1})$	1.007	1.002	1.005	1.000	0.999	1.000	1.003	0.999	0.993	0.998	0.999	1.002	0.999
	16	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**	**30**	**35**	**40**	**50**
$E(\hat{\beta}_{1,1})$	1.019	1.017	1.014	1.013	1.010	1.010	1.008	1.003	1.004	1.004	1.000	0.998	0.998
U_{MR}	0.978	0.980	0.982	0.984	0.986	0.988	0.990	0.992	0.994	0.995	0.997	0.999	1.002
$E_{U,MR}(\hat{\beta}_{1,1})$	0.996	0.997	0.997	0.997	0.995	0.998	0.998	0.995	0.998	0.999	0.997	0.998	1.000

Table 5-7: Values of $E(\hat{\beta}_{1,1})$, U_{MH} and $E_{U,MH}(\hat{\beta}_{1,1})$ at selected sample sizes (the modified Hirose's method for LS *X* on *Y*).

	n												
	3	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
$E(\hat{\beta}_{1,1})$	1.585	1.266	1.171	1.118	1.089	1.072	1.062	1.048	1.035	1.035	1.031	1.030	1.024
U_{MH}	0.631	0.790	0.856	0.893	0.916	0.933	0.945	0.954	0.961	0.967	0.971	0.975	0.978
$E_{U,MH}(\hat{\beta}_{1,1})$	1.000	0.999	1.002	0.998	0.998	1.000	1.003	1.000	0.995	1.000	1.001	1.004	1.002
	16	**17**	**18**	**19**	**20**	**22**	**24**	**26**	**28**	**30**	**35**	**40**	**50**
$E(\hat{\beta}_{1,1})$	1.019	1.017	1.014	1.013	1.010	1.010	1.008	1.003	1.004	1.004	1.000	0.998	0.998
U_{MH}	0.981	0.983	0.985	0.987	0.988	0.991	0.993	0.995	0.996	0.997	0.999	1.001	1.003
$E_{U,MH}(\hat{\beta}_{1,1})$	0.999	0.999	0.999	1.000	0.998	1.001	1.001	0.997	1.000	1.001	0.999	1.000	1.001

5.4.2 Bias Correction for the Shape Parameter Estimator of the OLSE for Censored Data

The bias of the OLSE of the shape parameter for censored data varies with sample size and censoring level. Based on the values of the bias under different combinations of sample sizes and censoring levels, obtained via Monte Carlo simulations, two 3-D surface plots were generated, as shown in Figure 5-4 (for OLSE) and Figure 5-5 (for MLE). The surface plots were generated by MATLAB 7 using the function *meshz*. There are three axes representing the bias of $\hat{\beta}_{1,1}$, sample size n and censoring level c, respectively. The color of the lines is proportional to the surface height and the color goes lighter as the bias goes larger.

Comparing the two surface plots, it can be observed that the surface plot of MLE shows a simpler relationship among the bias, sample size and censoring level. The bias of the MLE of the shape parameter is consistent with the sample size at any specific censoring level, and consistent with the censoring level at any specific sample size. However, the bias of the OLSE of the shape parameter is inconsistent in either way as the bias has a range of -10% – 15%. The surface plot in Figure 5-4 is further split in two parts, as shown in Figure 5-5: one shows the bias at low censoring levels (10% – 40%) and the other shows the bias at high censoring levels (50% – 80%). It can be observed from Figure 5-5 that, at low censoring levels (10% – 40%), the bias of the OLSE of the shape parameter presents good consistency and the bias is always negative when $10 \le n \le 100$, while the same is not true at high censoring levels (50% – 80%) because the bias reaches 0 at the combination of $c = 70\%$ and $n = 20 - 30$, or $c = 60\%$ and $n = 50 - 60$, or $c = 50\%$ and $n = 80 - 90$. The similar results have been presented in Section 3.3.3.2.

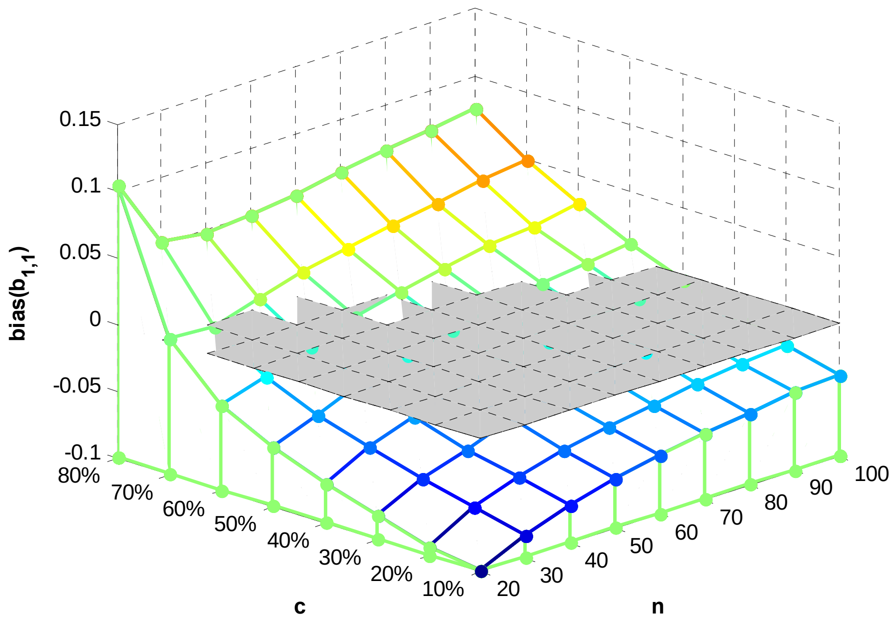


Figure 5-4: The surface plot of the bias of the shape parameter estimator of OLSE. The *Z* axis is the values of bias, the *Y* axis is censoring level (10% – 80%), and the *X* axis is sample size (20 – 100)§. The gray part in the second figure is the surface of bias = 0.

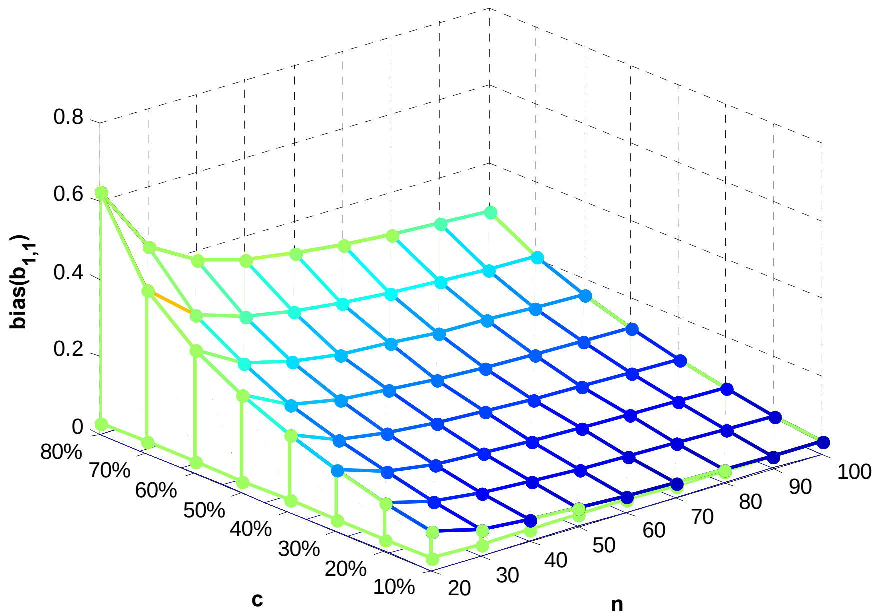


Figure 5-5: The surface plot of the bias of the shape parameter estimator of MLE. The *Z* axis is the values of bias, the *Y* axis is censoring level (10% – 80%), and the *X* axis is sample size (20 – 100)§.

§ The color of the line is proportional to the surface height (the value of bias).

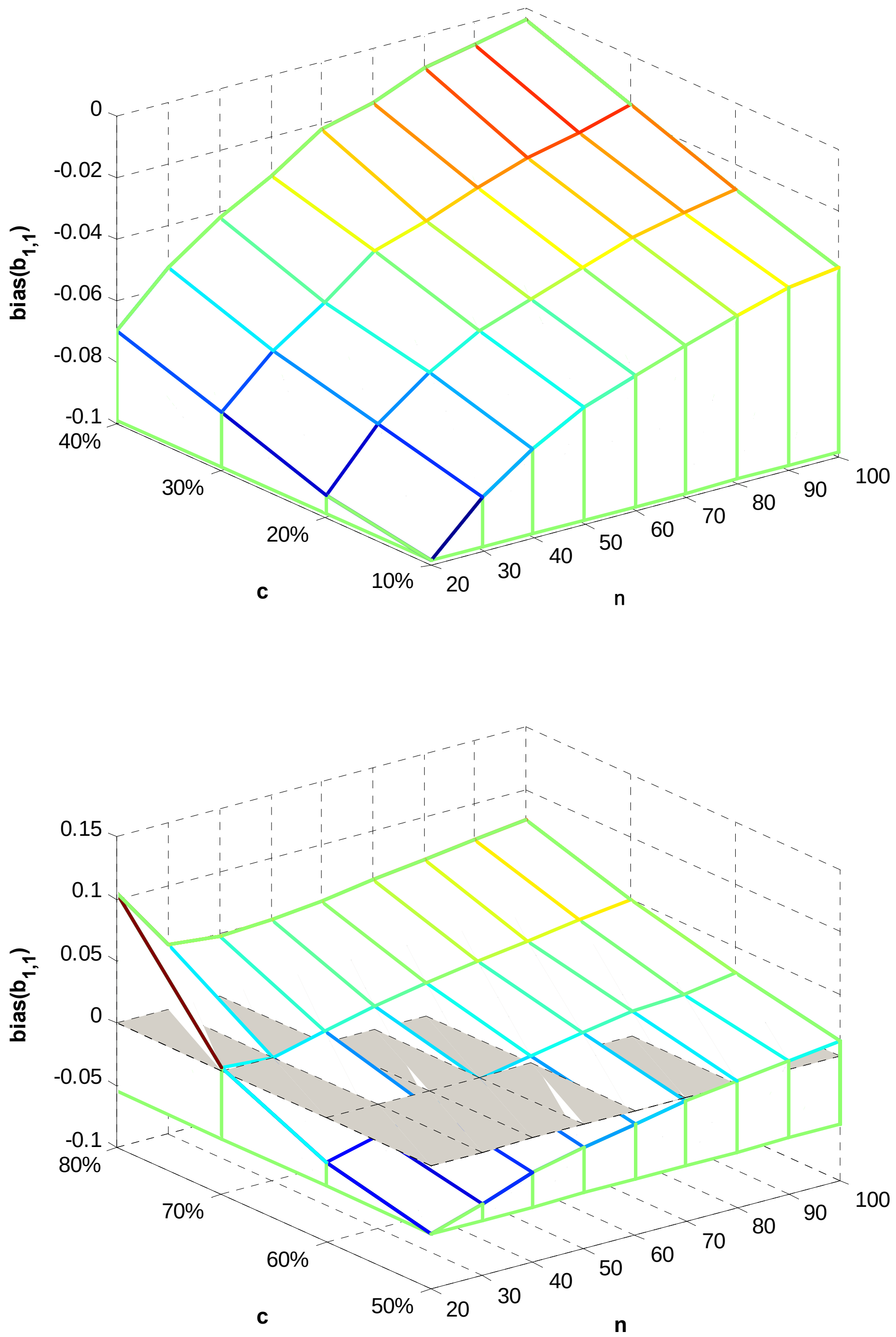


Figure 5-6: The surface plot of the bias of the shape parameter estimator of OLSE, split in two plots by censoring level. The *Z* axis is the values of bias, , the *Y* axis is censoring level (10% – 80%), and the *X* axis is sample size (20 – 100). The gray part in the second figure is the surface of bias = 0[§].

As the surface plots of the bias of the MLE at all censoring levels and the bias of the OLSE at low censoring levels show good consistency, it is possible to model the bias as a function of the sample size and censoring level. Ross (1996) proposed a bias correcting formula for the MLE of the shape parameter for singly right censored data, using the same model he used for the bias correction for complete data, i.e., Equation (5-9). The bias correcting formula for the MLE is given by

$$E\left[\hat{\beta}_{1,1}(n,r)\right] \approx 1 + \frac{1.37}{r-1.92}\sqrt{\frac{n}{r}} \tag{5-27}$$

In the following, a formula for correcting the bias of the OLSE of the shape parameter, applied to censored data with low censoring levels, is presented.

The proposed bias model is given by

$$B_{n,c}(\hat{\beta}) = -p_1 c^{p_2} n^{p_3} \tag{5-28}$$

where p_1, p_2, p_3 are the model parameters. $B_{n,c}(\hat{\beta})$ can only take negative values.

With the simulation generated values of the bias at different combinations of sample sizes and censoring levels, the values of p_1, p_2, p_3 were determined by the unconstrained nonlinear optimization technique (Nelder-Mead direct search method). The objective function is

$$\min \sum_{n_i, c_i} \left[\left(E(\hat{\beta}_{1,1})_{n_i,c_i} - 1 \right) - B_{n_i,c_i}(\hat{\beta}_{1,1}) \right]^2 = \sum_{n_i, c_i} \left[\left(E(\hat{\beta}_{1,1})_{n_i,c_i} - 1 \right) + p_1 c_i^{p_2} n_i^{p_3} \right]^2 \tag{5-29}$$

where n_i and c_i denote different sample sizes and censoring levels examined in the simulations. $E(\hat{\beta}_{1,1})_{n_i,c_i}$ denotes the value of $E(\hat{\beta}_{1,1})$ at a specific combination of n_i and c_i.

The parameter values were determined as

$$p_1 = 0.2211,\ p_2 = -0.3476,\ p_3 = -0.5430 \tag{5-30}$$

Substituting the values into Equation (5-28), the bias model is

$$B_{n,c}(\hat{\beta}) = -0.2211 c^{-0.3476} n^{-0.5430} \tag{5-31}$$

Thus the bias correcting formula is given by

$$\hat{\beta}_U(n,c) = \frac{\hat{\beta}}{E_{n,c}(\hat{\beta}_{1,1})} = \frac{\hat{\beta}}{1 + B_{n,c}(\hat{\beta}_{1,1})} = \frac{\hat{\beta}}{1 - 0.2211 c^{-0.3476} n^{-0.5430}} \tag{5-32}$$

Equation (5-32) can be added to the end of the conventional OLSE procedure for censored data in order to provide more accurate estimates. Table 5-8 tabulates the values of the simulation generated $E_{n,c}(\hat{\beta}_{1,1})$ and the corresponding unbiased estimates $E_U(\hat{\beta}_{1,1}) = E(\hat{\beta}_{1,1}) / (1 - 0.2211 c^{-0.3476} n^{-0.5430})$ at selected sample sizes and censoring levels. As can be seen from the table, $E_U(\hat{\beta}_{1,1})$ is more accurate than $E(\hat{\beta}_{1,1})$ at nearly all conditions. Great improvements can be observed when $c = 10\%$ and $n = 20, 30$.

Table 5-8: Values of $E(\hat{\beta}_{1,1})$ and $E_U(\hat{\beta}_{1,1})$ at selected sample sizes and censoring levels.

c \ n		20	30	40	50	60	70	80	90	100
10%	$E(\hat{\beta}_{1,1})$	0.902	0.917	0.928	0.938	0.943	0.949	0.955	0.957	0.962
	$E_U(\hat{\beta}_{1,1})$	0.998	0.994	0.994	0.997	0.996	0.998	1.000	1.000	1.002
20%	$E(\hat{\beta}_{1,1})$	0.908	0.927	0.938	0.948	0.954	0.961	0.963	0.968	0.973
	$E_U(\hat{\beta}_{1,1})$	0.983	0.988	0.990	0.994	0.996	0.999	0.998	1.002	1.005
30%	$E(\hat{\beta}_{1,1})$	0.917	0.935	0.946	0.957	0.965	0.970	0.972	0.979	0.982
	$E_U(\hat{\beta}_{1,1})$	0.982	0.988	0.991	0.998	1.002	1.003	1.004	1.009	1.010
40%	$E(\hat{\beta}_{1,1})$	0.930	0.945	0.961	0.970	0.975	0.983	0.988	0.989	0.994
	$E_U(\hat{\beta}_{1,1})$	0.989	0.992	1.002	1.007	1.008	1.014	1.016	1.016	1.019

It should be noted that the proposed bias correcting formula in Equation (5-32) is specially designed for censored samples satisfying $c \le 40\%$ and $n \le 100$. Since the OLSE of the shape parameter is inconsistent with either sample size or censoring level, it is difficult to have a single bias correcting formula that works for all conditions.

5.5 Summary

In this chapter, several bias correcting formulas for the OLSE of the Weibull shape parameter were proposed. These formulas can be added to the end of the conventional LSE procedure in order to provide more accurate estimates for the shape parameter.

The main work in this chapter is the bias correction applied to small and complete samples, where two methods, i.e., the modified Ross' bias correction method and the modified Hirose's bias correction method, were proposed for the OLSE of the shape parameter and examined in details. Although the bias correcting formulas were determined by numerical methods, they work very well as confirmed by the Monte Carlo simulation experiments. The bias is reduced to less than 1% and typically less than 0.5%. The application procedures were also provided for the proposed methods.

Two bias correcting formulas were also proposed for the shape parameter estimator of LS *X* on *Y* using the modified Ross' method and the modified Hirose's method.

Bias correction for the OLSE of the shape parameter in the case of multiply censored data is challenging. A simple bias correcting formula was proposed for multiply censored samples with $c \leq 40\%$ and $n \leq 100$. The bias is greatly reduced with the proposed formula.

One thing to note is that theoretically these bias correction methods can greatly reduce or eliminate the bias of the shape parameter estimator in the long run; however, they may not provide more accurate estimate for a single Weibull sample than OLSE.

Chapter 6

Weighted Least Squares Estimation Methods

This chapter presents the weighted least squares estimation methods. A simple approximation formula is proposed for calculating weights for small, complete samples. Through Monte Carlo simulations, the proposed WLSE method is compared with some existing WLSE methods and the OLSE method. The simulation results show that the proposed procedure is slightly better than the existing WLSE methods in terms of the properties of the estimators. The efficiency of the proposed WLSE method is 20% to 30% higher than that of the OLSE method. A bias correcting formula is also proposed to reduce the bias of the shape parameter estimator of the proposed WLSE method. WLSE for censored data is discussed and a tentative procedure is proposed for calculating weights.

6.1 Introduction

One problem with LSE is that it treats each data point equally under the assumption that the variance of the error term is constant. As shown in Section 3.2, this assumption cannot be satisfied. The variance of the errors can be calculated by Equations (3-8) – (3-11) under the assumption that the uncertainty of failure time can be neglected. By treating each data point equally, LSE has a low efficiency. WLSE, in theory, can maximize the efficiency of parameter estimation by giving each data point its proper amount of influence over the parameter estimates.

The biggest challenge of WLSE is to determine the appropriate weight for each data point. As a common practice, weights can be calculated by the reciprocal of the variances of the dependent variable values. The variances of the dependent variable values, in most cases, are estimated by repeated experiments. However, the values may also be obtained through analytical deduction, which is the current situation. In examining WLSE for the Weibull distribution, we still treat $X = \ln T$ as the independent variable and $Y = \ln[-\ln(1 - F(t))]$ as the dependent variable. We further assume that the uncertainty of failure time can be neglected so that the variance of the errors equals the variance of the dependent variable values. In particular, here the values of the dependent variable Y are not measured but estimated.

The weights in the WLSE method for the Weibull distribution can be calculated by

$$w_i = 1/Var(Y_{(i)}) \tag{6-1}$$

where $Var(Y_{(i)})$ can be obtained through analytical methods as shown later.

Several authors have examined the WLSE methods for the Weibull distribution and different methods for calculating weights have been proposed (White, 1969; Bergman, 1986; Faucher & Tyson, 1988; Hung, 2001; Lu et al., 2004). These methods are briefly described in Section 6.2. We noticed that, in most of the existing WLSE methods, $Var(Y_{(i)})$ is estimated via some kind of approximation method, e.g., the propagation of error. It is likely that errors are introduced by using such approximations. The exact values of $Var(Y_{(i)})$ are derived, which, in theory, are the best weights. To simplify the calculation for the best weights, a simple approximation formula is proposed through numerical method that can be used for small, complete samples, especially when $n \le 20$. With Monte Carlo simulations, the proposed methods are compared with the methods of Faucher & Tyson (1988), Lu et al. (2004) and OLSE for estimating the Weibull parameters. Censored data is also discussed and a method for calculating weights based on MFON is proposed. A numerical example clearly shows the proposed WLSE procedure for censored data. Some of the related work has been published by the authors (Zhang et al., 2006, 2008).

6.2 WLSE and Related Work

The idea of WLSE is to give each data point its proper amount of influence by assigning each data point a weight, denoted by w_i. The objective function of WLSE is

$$\min S = \sum_{i=1}^{r} w_i \left[y_i - (A + Bx_i) \right]^2 \tag{6-2}$$

where for complete data, $r = n$.

The conventional settings described in Section 2.3.1 for LSE are applicable, i.e., $X = \ln T$, $Y = \ln[-\ln(1-F)]$, $A = -\beta \ln \alpha$ and $B = \beta$. Given a Weibull sample, the values of x_i and y_i can be obtained in a similar approach as in the LSE method (see Section 2.3.1). From Equation (6-2), taking partial derivatives of S with regard to A and B, respectively, and setting the results to 0, we obtain

$$\begin{cases} \hat{B} = \dfrac{\sum_{i=1}^{r} w_i \cdot \sum_{i=1}^{r} w_i x_i y_i - \sum_{i=1}^{r} w_i x_i \cdot \sum_{i=1}^{r} w_i y_i}{\sum_{i=1}^{r} w_i \cdot \sum_{i=1}^{r} w_i x_i^2 - \left(\sum_{i=1}^{r} w_i x_i \right)^2} \\ \hat{A} = -\dfrac{\sum_{i=1}^{r} w_i y_i - \hat{\beta} \sum_{i=1}^{r} w_i x_i}{\sum_{i=1}^{r} w_i} \end{cases} \tag{6-3}$$

Thus the estimators of α and β for the WLSE method are given by

$$\begin{cases} \hat{\beta} = \dfrac{\sum_{i=1}^{r} w_i \cdot \sum_{i=1}^{r} w_i x_i y_i - \sum_{i=1}^{r} w_i x_i \cdot \sum_{i=1}^{r} w_i y_i}{\sum_{i=1}^{r} w_i \cdot \sum_{i=1}^{r} w_i x_i^2 - \left(\sum_{i=1}^{r} w_i x_i \right)^2} \\ \hat{\alpha} = \exp\left(- \dfrac{\sum_{i=1}^{r} w_i y_i - \hat{\beta} \sum_{i=1}^{r} w_i x_i}{\hat{\beta} \sum_{i=1}^{r} w_i} \right) \end{cases} \tag{6-4}$$

Equation (6-4) can be applied to both complete data and censored data. For complete data, $r = n$. As a special case, when $w_i = 1$ for all data points, the WLS estimators reduce to the OLS estimators.

The WLSE method can be easily carried out after the values of w_i are determined. As mentioned in Section 6.1, w_i can be calculated by Equation (6-1), i.e., the reciprocal of $Var(Y_{(i)})$. Following this rule, different methods for calculating $Var(Y_{(i)})$ have been proposed. White (1969) defined a log-Weibull order statistic and derived the formula for calculating its variance that equals the variance of $Var(Y_{(i)})$. The formula is complicated and the results are tabulated for selected sample sizes. White also gave a numerical example of using WLSE to estimate the Weibull parameters; however, without any discussion on the accuracy of the estimates. Another shortcoming of the White's method is that the regression of $X = \ln T$ on $Y = \ln[-\ln(1-F)]$ is used, which is not the conventional way nowadays. Therefore, there is no further discussion on this method. Besides, Bergman (1986), Faucher & Tyson (1988), Hung (2001) and Lu et al. (2004) each proposed a simple formula for calculating weights. These methods are briefly introduced below and some comments are given.

Bergman's WLSE (Bergman, 1986)

Bergman (1986) applied the formula of the propagation of error (the simple case that involves only a single variable) to the relationship $Y_{(i)} = \ln\left[-\ln(1-F_{(i)})\right]$ and obtained

$$S_{Y_{(i)}} = \frac{dY_{(i)}}{dF_{(i)}} \cdot S_{F_{(i)}} = -S_{F_{(i)}} \cdot \left[\left(1-\hat{F}_{(i)}\right)\ln\left(1-\hat{F}_{(i)}\right)\right]^{-1} \tag{6-5}$$

where $S_{Y_{(i)}}$ and $S_{F_{(i)}}$ denote the standard deviations of $Y_{(i)}$ and $F_{(i)}$, respectively.

By assuming $S_{F_{(i)}}$ is a constant, $S_{Y_{(i)}}$ is then proportional to $\left[\left(1-\hat{F}_{(i)}\right)\ln\left(1-\hat{F}_{(i)}\right)\right]^{-1}$ and $Var(Y_{(i)})$ is proportional to $\left[\left(1-\hat{F}_{(i)}\right)\ln\left(1-\hat{F}_{(i)}\right)\right]^{-2}$. Thus Bergman determined the formula for weights as

$$w_i = \left[\left(1-\hat{F}_{(i)}\right)\ln\left(1-\hat{F}_{(i)}\right)\right]^{2} \tag{6-6}$$

Bergman examined two non-parametric estimators for calculating $\hat{F}_{(i)}$, i.e., $i/(n+1)$ and $(i-0.5)/n$. He conducted a simulation experiment to compare his WLSE method with LSE on estimating the shape parameter with both plotting positions. The mean and standard deviation of $\hat{\beta}$, denoted by $E(\hat{\beta})$ and $S(\hat{\beta})$, were calculated. The comparison criteria were $E(\hat{\beta})/\beta$ and $S(\hat{\beta})/E(\hat{\beta})$ (coefficient of variation). The author concluded that 1) WLSE has little effect on the coefficient of variation; 2) in view of bias, the Hazen estimator $\hat{F}_{(i)} = (i-0.5)/n$ should be used for both LSE and WLSE; and 3) in view of bias, WLSE with $\hat{F}_{(i)} = i/(n+1)$ outperforms

LSE with $\hat{F}_{(i)} = i/(n+1)$, while WLSE with $\hat{F}_{(i)} = (i-0.5)/n$ performs similarly to LSE with $\hat{F}_{(i)} = (i-0.5)/n$. Obviously, the author did not focus on the efficiency improvement of WLSE over LSE, which should be measured directly by the standard deviation or variance of estimators.

<u>Hung's WLSE (Hung, 2001)</u>

Hung (2001) proposed a formula for calculating weights in a way similar to that of Bergman (1986). Hung's weights are given by

$$w_i = \frac{\left[\left(1-\hat{F}_{(i)}\right)\ln\left(1-\hat{F}_{(i)}\right)\right]^2}{\sum_{i=1}^{n}\left[\left(1-\hat{F}_{(i)}\right)\ln\left(1-\hat{F}_{(i)}\right)\right]^2} \tag{6-7}$$

Compared to Equation (6-6), this formula simply adds a denominator, i.e., $\sum_{i=1}^{n}\left[\left(1-\hat{F}_{(i)}\right)\ln\left(1-\hat{F}_{(i)}\right)\right]^2$. The author did not give the reason for adding this denominator, but clearly, it is independent of i and can be treated as constant. Since the weight for one observation is given relative to the weights for other observations, this denominator will not affect the estimation results. Therefore, Hung's weight formula is same as Bergman's weight formula.

Hung suggested $\hat{F}_{(i)}$ be calculated by the method of Drapella & Kosznik (1999), i.e., the expected plotting position described in Section 4.3.3. Via Monte Carlo simulations, Hung compared three estimation methods: WLSE with the Bernard estimator for calculating $\hat{F}_{(i)}$, LSE with the Bernard estimator for calculating $\hat{F}_{(i)}$ and LSE with the expected plotting position for calculating $\hat{F}_{(i)}$. The mean, variance and

MSE of the shape parameter estimator of each method were calculated as the comparison criteria. The simulation results showed that Hung's WLSE method provided the smallest variance and MSE in all cases examined.

Similar to Bergman's WLSE method, Hung's WLSE method involves an assumption that the uncertainty of $F_{(i)}$ is constant.

F&T's WLSE (Faucher and Tyson, 1988)

Faucher and Tyson (1988) pointed out that the order statistic $F_{(i)}$ has a beta distribution with parameters i and $n-i+1$; therefore, the uncertainty of $F_{(i)}$ cannot be constant. The authors proposed to estimate the uncertainty of $F_{(i)}$ through the difference of its two percentiles. The percentiles can be calculated by

$$\sum_{k=1}^{i} \binom{n}{k-1} F_{(i)}^{k-1} \left(1 - F_{(i)}\right)^{n+1-k} = 1 - p \tag{6-8}$$

The 20th percentile and the 80th percentile were selected to estimate the uncertainty of $F_{(i)}$. Then, with the relationship $Y_{(i)} = \ln\left[-\ln(1-F_{(i)})\right]$, the uncertainty of $Y_{(i)}$ is estimated by the difference of $\ln\left[-\ln(1-F'_{(i)})\right] - \ln\left[-\ln(1-F''_{(i)})\right]$, where $F'_{(i)}$ denotes the 80th percentile of $F_{(i)}$ that can be calculated by setting $p = 0.8$ in Equation (6-8) and $F''_{(i)}$ denotes the 20th percentile calculated by setting $p = 0.2$ in Equation (6-8). The weight formula is then expressed by

$$w_i = 1 \Big/ \left\{ \ln\left[-\ln(1-F'_{(i)})\right] - \ln\left[-\ln(1-F''_{(i)})\right] \right\}^2 \tag{6-9}$$

However, the selection of the two percentiles is somewhat subjective.

A simple approximation formula was also proposed via numerical methods as

$$w_i = 3.3\hat{F}_{(i)} - 27.5\left[1 - (1 - \hat{F}_{(i)})^{0.025}\right] \tag{6-10}$$

They suggested the Bernard estimator or exact median rank values to calculate $\hat{F}_{(i)}$.

The authors used Monte Carlo simulations to compare their WLSE with LSE in view of the bias and standard deviation of both scale and shape parameter estimators. The Hazen estimator and the Bernard estimator for calculating $\hat{F}_{(i)}$ were examined and compared. The results showed that their WLSE procedure with the Bernard estimator should be preferred because it generates smallest standard deviation of the estimators.

Lu et al's WLSE (Lu et al., 2004)

Lu et al. (2004) defined an intermediate variable C with $C = -\ln(1-F)$ and $Y = \ln C$. From the Weibull CDF, it can be easily obtained that C follows the standard exponential distribution. Therefore, the mean and variance of its order statistic $C_{(i)}$ can be determined as

$$E(C_{(i)}) = \sum_{j=1}^{i} \frac{1}{n-j+1} \text{ and } Var(C_{(i)}) = \sum_{j=1}^{i} \frac{1}{(n-j+1)^2} \tag{6-11}$$

$Var(Y_{(i)})$ can be approximated by applying the propagation of error formula on the relationship $Y = \ln C$, i.e.,

$$Var(Y_{(i)}) = Var\left(\ln C_{(i)}\right) = \frac{Var(C_{(i)})}{\left[E(C_{(i)})\right]^2} \tag{6-12}$$

Substituting Equation (6-11) in Equation (6-12) for $E(C_{(i)})$ and $Var(C_{(i)})$ yields

$$Var(Y_{(i)}) = \sum_{j=1}^{i} \frac{1}{(n-j+1)^2} \Bigg/ \left[\sum_{j=1}^{i} \frac{1}{(n-j+1)} \right]^2 \qquad (6\text{-}13)$$

The Lu's formula for weights is then given by

$$w_i = \left[\sum_{j=1}^{i} \frac{1}{(n-i+j)} \right]^2 \Bigg/ \sum_{j=1}^{i} \frac{1}{(n-i+j)^2} \qquad (6\text{-}14)$$

Besides its simplicity, this weight formula does not involve the selection of the estimator for calculating $\hat{F}_{(i)}$.

The authors compared several WLSE methods, including Bergman's WLSE, Hung's WLSE, F&T's WLSE, and their WLSE, via Monte Carlo simulations. For the Bergman's and F&T's methods, three estimators of $F_{(i)}$ including the Weibull estimator, the Hazen estimator, and the Bernard estimator, were examined. The mean, variance and MSE of the shape parameter estimators were calculated as the comparison criteria. It was concluded that Bergman's WLSE (as well as Hung's WLSE) in most cases generates a larger MSE than the others regardless of the plotting positions used. The authors' method and the F&T's method performed similarly.

Discussions

Equation (6-6), Equation (6-7), Equation (6-9) or Equation (6-10), and Equation (6-14) are the formulas for calculating weights proposed by different authors in the past. In summary, all these formulas are easy to use. Bergman's method, as well as Hung's, involves the assumption that the uncertainty of $F_{(i)}$ is constant. However, this is not a good assumption because the values of $F_{(i)}$ come from order statistics and the variance varies with the order number i. As $F_{(i)}$ can be treated as from a beta distribution, the variance of $F_{(i)}$ is given by

$$Var(F_{(i)}) = \frac{i(n+1-i)}{(n+1)^2(n+2)} \tag{6-15}$$

Lu et al. (2004) showed that Bergman's method (as well as Hung's) is inferior to the others from their simulation results. The F&T's and Lu et al.'s methods do not involve any assumptions; however, both methods' calculation for $Var(Y_{(i)})$ used approximation methods, and the values are only approximated values. It is likely that errors are introduced by using such approximations. In the next section, we present the method for calculating the exact values of $Var(Y_{(i)})$ that will generate the most appropriate weights, and compare it with the existing methods.

6.3 Method for Calculating Best Weights

The best values for weights can be obtained through the exact reciprocal values of $Var(Y_{(i)})$. As shown in Section 3.2.2, we have derived the formula for calculating the exact values of $Var(Y_{(i)})$, i.e.,

$$\begin{cases} E\left(Y_{(i)}\right) = i\binom{n}{i} \cdot \sum_{k=0}^{i-1}\left\{(-1)^k \binom{i-1}{k} \cdot \frac{-\gamma - \ln(n-i+k+1)}{n-i+k+1}\right\} \\ E\left(Y_{(i)}^2\right) = 1.978112 + i\binom{n}{i}\sum_{k=0}^{i-1}\left\{(-1)^k \binom{i-1}{k} \cdot \frac{2\gamma \ln(n-i+k+1) + \ln^2(n-i+k+1)}{n-i+k+1}\right\} \\ Var\left(Y_{(i)}\right) = E\left(Y_{(i)}^2\right) - E^2\left(Y_{(i)}\right) \end{cases}$$

(6-16)

The best weights can then be calculated by Equation (6-1), i.e., $w_i = 1/Var(Y_{(i)})$.

Since the weight for one observation is given relative to the weights for other observations, it can be normalized in some way. Here we normalize the weights by dividing the weight for each observation by the mean weight over the whole sample, i.e.,

$$w_{nor\_i} = \frac{w_i}{\overline{w}} \tag{6-17}$$

where w_{nor_i} denotes the normalized weight. $\overline{w} = \sum_{i=1}^{n} w_i \Big/ n$. In this way, the sum of the normalized weights equals the sample size, i.e., $\sum_{i=1}^{n} w_{nor\_i} = n$.

Table 6-1 lists the values of the normalized best weights for selected sample sizes. From the table it can be observed that, for $n = 5, 6$, the weights increase as the order number *i* increases, and the weights for the last two data points are much higher than those for the first two data points. From $n = 7$ onwards, however, the largest weights are not given to the last data point but a little bit earlier. The weights for the end part of the sample are still much larger than those for the beginning part of the sample.

Apparently, the weights differ greatly for the data points in a sample; therefore, the WLSE method and the OLSE method should perform differently.

Table 6-1: The normalized best weights for selected sample sizes (the largest value in each column is highlighted).

	n									
i	5	6	7	8	10	12	14	16	18	20
1	0.2675	0.2269	0.1970	0.1741	0.1414	0.1190	0.1028	0.0904	0.0807	0.0729
2	0.6779	0.5761	0.5009	0.4431	0.3600	0.3032	0.2619	0.2305	0.2058	0.1859
3	1.0838	0.9286	0.8108	0.7190	0.5857	0.4939	0.4269	0.3759	0.3357	0.3033
4	1.4263	1.2538	1.1071	0.9875	0.8091	0.6841	0.5921	0.5218	0.4663	0.4215
5	**1.5446**	1.5013	1.3673	1.2364	1.0250	0.8708	0.7556	0.6668	0.5964	0.5393
6		**1.5133**	**1.5416**	1.4440	1.2266	1.0509	0.9155	0.8097	0.7252	0.6564
7			1.4754	**1.5605**	1.4023	1.2202	1.0699	0.9494	0.8520	0.7721
8				1.4353	1.5306	1.3719	1.2158	1.0843	0.9758	0.8859
9					**1.5628**	1.4946	1.3489	1.2125	1.0957	0.9971
10					1.3564	**1.5667**	1.4626	1.3310	1.2100	1.1049
11						1.5410	1.5452	1.4355	1.3167	1.2080
12						1.2836	**1.5758**	1.5193	1.4129	1.3050
13							1.5089	**1.5709**	1.4942	1.3938
14							1.2181	1.5700	1.5539	1.4715
15								1.4726	**1.5810**	1.5337
16								1.1595	1.5554	1.5738
17									1.4351	**1.5811**
18									1.1071	1.5359
19										1.3979
20										1.0599

The weights calculated by other methods, e.g., the Bergman's, F&T's and Lu et al.'s methods, can be normalized in a similar way. After normalizing, the values of weights from different methods can be compared for the same sample size. Figure 6-1 and Figure 6-2 show the comparison of the best weights and the weights calculated by the Bergman's method in Equation (6-6), the F&T's method in Equation (6-10), and the Lu et al.'s method in Equation (6-14), for two sample sizes, $n = 5$ and $n = 15$, respectively. The following can be observed from the two figures: 1) the weights of the F&T method are close to the best weights; 2) compared to the best weights, the weights of the Bergman's and Lu et al.'s methods present reversed trends. Bergman's method underestimates the last few points and overestimates the remaining points, while Lu et al.'s method overestimates the last few points and underestimates the remaining points; and 3) at $n = 15$, the weight for the 10th point, calculated by different methods, are almost same.

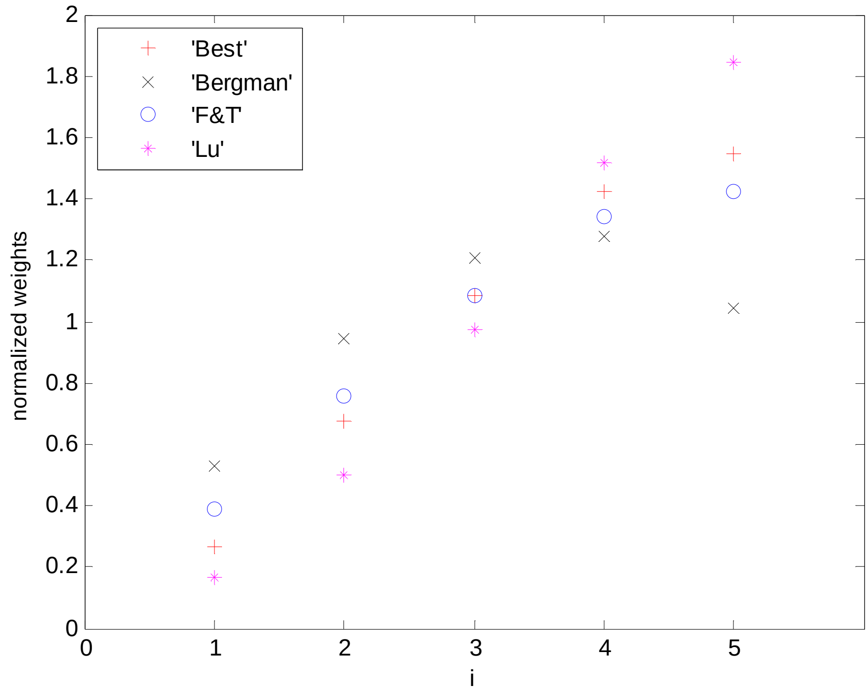


Figure 6-1: Comparison of normalized weights calculated by different methods at *n* = 5.

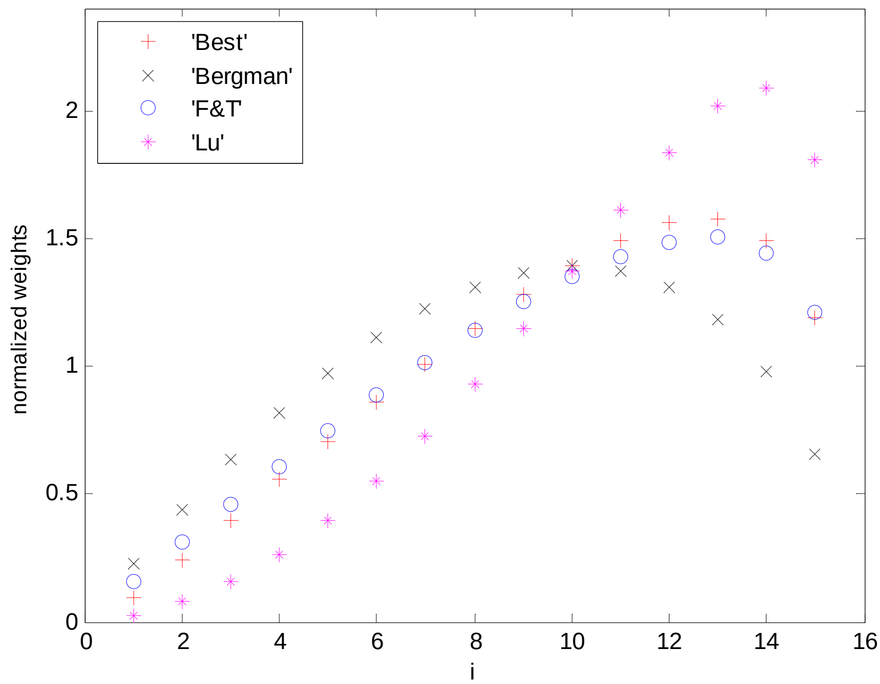


Figure 6-2: Comparison of normalized weights calculated by different methods at *n* = 15.

6.4 An Approximation Formula for Calculating Weights for Small, Complete Samples

Using Equation (6-16) and then Equation (6-1) to calculate weights is not convenient without the aid of a computer program. Also note that when the sample size becomes large, say $n \ge 30$, the binomial coefficients in Equation (6-16) will become extremely large, and it is hard to generate accurate results. Considering the fact that OLSE does not perform very well mainly for small samples, say $n \le 20$, the examination of WLSE also focuses on small samples.

To simplify the calculation for weights, it is possible to use numerical methods to derive an approximation formula for calculating the best weights. w_i can be modelled as the function of order number *i* and sample size *n*, as can be seen in Table 6-1. However, intuitively, it is easier to model it as the function of $\hat{F}_{(i)}$, like in the Bergman's, Hung's and F&T's methods. To study the relationship between the best weights and $\hat{F}_{(i)}$, Figure 6-3 plots the best weights calculated by Equation (6-16) and then Equation (6-1) at selected sample sizes, and Figure 6-4 plots the values of $\hat{F}_{(i)}$, calculated by the Bernard estimator, i.e., $\hat{F}_{(i)} = (i - 0.3)/(n + 0.4)$, at the same sample sizes. As can be seen, the two figures show similar patterns. Therefore, the best weights can be modelled as a function of $\hat{F}_{(i)}$.

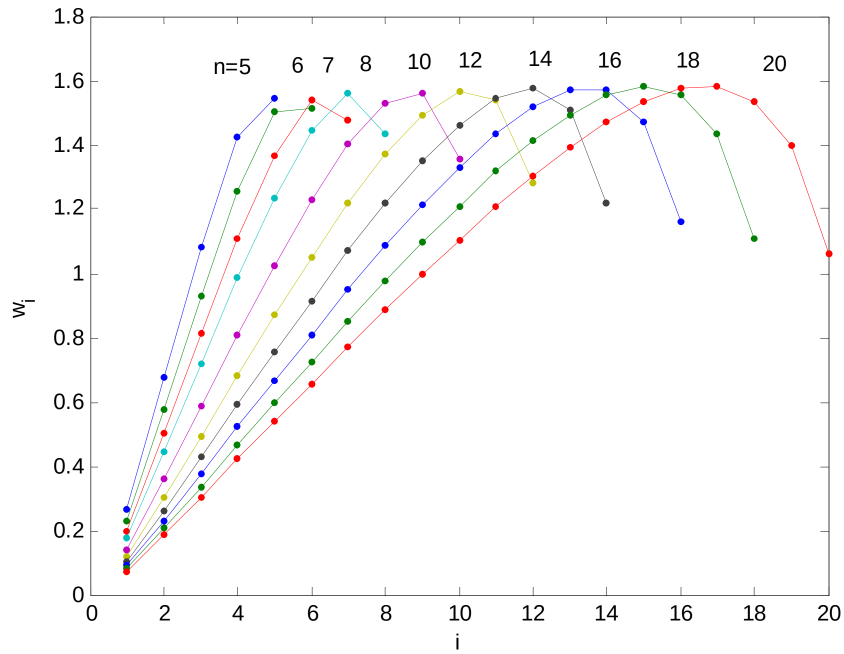


Figure 6-3: Plot of best weights as a function of *i* and *n*.

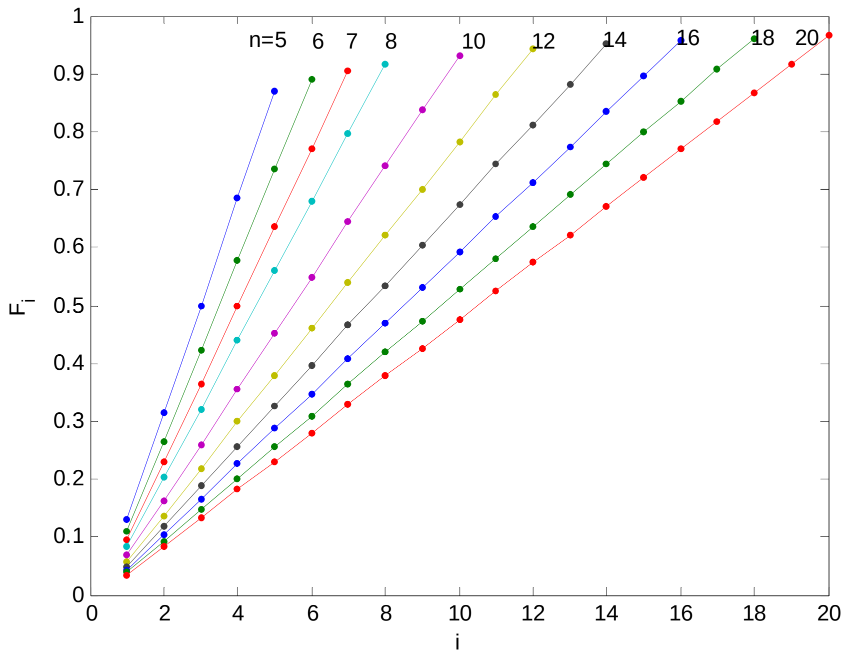


Figure 6-4: Plot of $\hat{F}_{(i)}$ (calculated by the Bernard estimator) as a function of *i* and *n*.

6.4.1 The Approximation Formula

The relationship between w_i and $\hat{F}_{(i)}$ can be modelled by a polynomial function, i.e.,

$$w\_app(i) = p_0 + p_1\hat{F}_{(i)} + p_2\hat{F}_{(i)}^2 + p_3\hat{F}_{(i)}^3 + p_4\hat{F}_{(i)}^4 \quad (6\text{-}18)$$

where $w\_app(i)$ denotes the approximated value of w_i, and p_0, p_1, p_2, p_3, p_4 are the model parameters to be determined.

The model parameters can be determined by the nonlinear curve fitting technique. The objective function is

$$\min \sum_{i=1}^{n} \left[w_i - (p_0 + p_1\hat{F}_{(i)} + p_2\hat{F}_{(i)}^2 + p_3\hat{F}_{(i)}^3 + p_4\hat{F}_{(i)}^4) \right]^2 \quad (6\text{-}19)$$

To solve this function, the multidimensional unconstrained nonlinear minimization method, i.e., the Nelder-Mead method (Nelder & Mead, 1965), was used. The calculation was executed in MATLAB 7, and the built-in function *fminsearch* was used.

The best values of w_i, calculated by Equation (6-16) and then Equation (6-1), and $\hat{F}_{(i)}$, calculated by $\hat{F}_{(i)} = (i - 0.3)/(n + 0.4)$, for samples of sizes 2 to 20 were used in Equation (6-19) to determine the five model parameters. Finally, the values of p_0, p_1, p_2, p_3, p_4 were determined as

$$p_0 = -0.076, p_1 = 3.610, p_2 = -6.867, p_3 = 13.54, p_4 = -9.231 \quad (6\text{-}20)$$

Thus the approximation formula for calculating the best weights is

$$w\_app(i) = -0.076 + 3.610\hat{F}_{(i)} - 6.867\hat{F}_{(i)}^2 + 13.54\hat{F}_{(i)}^3 - 9.231\hat{F}_{(i)}^4 \quad (6\text{-}21)$$

6.4.2 Application Procedure

The application procedure of the proposed WLSE method for estimating the Weibull parameters in the case of small, complete samples is summarized as follows:

Step 1: Rank the failure times from smallest to largest and calculate the *Y*-axis plotting positions by $\hat{F}_{(i)} = (i-0.3)/(n+0.4)$.

Step 2: Plot the failure times t_i against $\hat{F}_{(i)}$ on WPP. If the Weibull distribution fits, the data points should appear to be on a straight line.

Step 3: Calculate the values of the weights for each data point using Equation (6-21).

***Step 4*:** Calculate the estimates for α and β using Equation (6-4).

Nowadays, many statistical software packages and electrical spreadsheets provide the WLS programs, and users just need to provide x_i, y_i and the weights. Therefore, WLSE can be easily applied.

6.4.3 A Numerical Example

Below is a randomly generated Weibull sample with $\alpha = 1$ and $\beta = 2$:
0.2153, 0.6394, 0.7607, 0.8112, 1.0024, 1.2612, 1.3418, 1.4468, 1.5011, 1.8998.

Five methods, including the proposed one, Bergman's WLSE, F&T's WLSE, Lu et al.'s WLSE, and OLSE, were used to estimate the two Weibull parameters for this sample. The results are shown in Table 6-2. Figure 6-5 is the WPP with straight lines generated by each method.

It can be observed from Figure 6-5 that the OLSE line is greatly affected by the first point. In Table 6-2 we can see that OLSE highly underestimates β. The proposed method and the F&T's method provide the best estimates for β and the bias is very small. Bergman's WLSE method underestimates β, and Lu et al.'s WLSE method overestimates β. The differences in $\hat{\alpha}$ among these methods are smaller compared to $\hat{\beta}$; however, overall the bias of $\hat{\alpha}$ is larger than that of $\hat{\beta}$.

Table 6-2: Estimates of α and β generated by different WLSE methods and OLSE.

	Proposed	Bergman	F&T	Lu	OLSE
$\hat{\alpha}$	1.2526	1.2774	1.2547	1.2465	1.2863
$\hat{\beta}$	2.0639	1.9350	2.0318	2.2034	1.7221

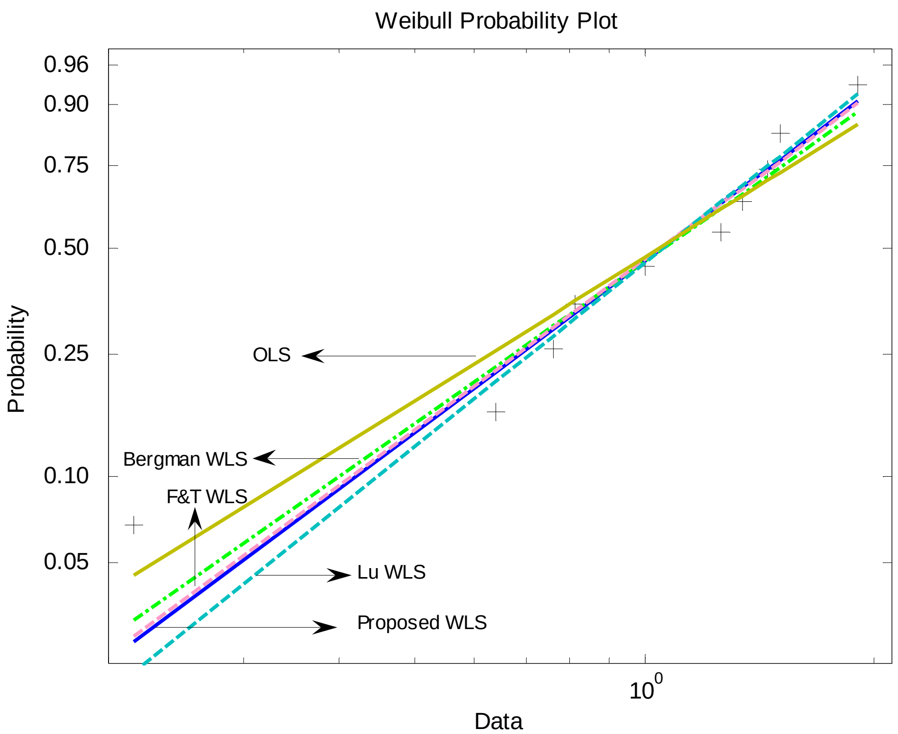


Figure 6-5: WPP with straight lines generated by different WLSE methods and OLSE.

6.4.4 Monte Carlo Study: A Comparison of Different WLSE Methods and OLSE

Monte Carlo experiments were conducted to examine the proposed WLSE method for estimating the Weibull parameters for small, complete data sets.

Five methods were compared in this experiment, including the following:

1. **Best W**: A WLSE method with best weights calculated by Equation (6-16) and then Equation (6-1).
2. **App. W**: The proposed WLSE method, where the approximated best weights calculated by Equation (6-21) are used.
3. **F&T**: A WLSE method where weights are calculated by Equation (6-10).
4. **Lu**: A WLSE method where weights are calculated by Equation (6-14).
5. **OLSE**.

The Bergman's method was not considered in this experiment because it has been shown inferior to the other existing WLSE methods (Lu et al., 2004) and it involves an inappropriate assumption that $F_{(i)}$ has no uncertainty.

Weibull samples of different sizes were randomly generated with selected values of α and β. For each sample generated, the above techniques were used to obtain the estimates of α and β simultaneously. By repeating this process for 10000 times, the mean, standard deviation, and MSE of the parameter estimates were calculated as the comparison criteria. The setting of the experiment factors is given in Table 6-3. For all the methods, $\hat{F}_{(i)}$ is calculated by $\hat{F}_{(i)} = (i-0.3)/(n+0.4)$, i.e., the Bernard estimator.

Table 6-3: Setting of experiment factors. The experiment is to compare four WLSE methods and OLSE.

Factors	Values
α_T	1
β_T	0.5, 1, 2
n	5, 6, …, 19, 20
M	10000
Methods	Best W, App. W, F&T, Lu, OLSE

It should be noted that the weights of all WLSE techniques examined here are independent of the values of α and β. Therefore, the two pivotal functions, $\hat{\beta}/\beta$ and $\hat{\beta}\ln(\hat{\alpha}/\alpha)$, for the LS estimated α and β, are also true for the WLS estimated Weibull parameters. The advantage of using the pivotal functions is that their distributions can be derived from the normalized Weibull distribution ($\alpha = \beta = 1$), so that the simulation work can be greatly reduced. In this experiment, the true value of β was fixed at 1 to assess the estimators of β, and to assess the estimators of α, three true values of β were used, i.e., $\beta_T = 0.5, 1, 2$. Since α is a scale parameter, we fixed its true value to 1 in the whole experiment.

The simulation results are shown in Table 6-4 and Table 6-5. The results for selected sample sizes are omitted; however, it will not affect the following conclusions which can be observed from the tabulated values.

Simulation Results for Estimators of β (Table 6-4)

1) In view of both the standard deviation and MSE of $\hat{\beta}$, the WLSE methods are significantly better than OLSE. The ratio of the MSE of $\hat{\beta}$ between App. W and OLSE is about 70% at $n = 20$. Among the WLSE methods examined, Best W and App. W always generate the smallest standard deviation and

MSE, followed by the F&T's method. Lu's method generates slightly larger standard deviation and MSE than the other three WLSE methods.

2) In view of bias, all WLSE methods perform similarly, and they only outperform OLSE at n = 5. In most cases, the bias of $\hat{\beta}$ of OLSE is much smaller, say, about 2 – 3% less than that of the other methods.

3) App. W performs very close to Best W.

Simulation Results for Estimators of α (Table 6-5)

1) In view of both the standard deviation and MSE of $\hat{\alpha}$, the WLSE methods always outperform OLSE. Among the WLSE methods examined, Lu's method always generates the smallest standard deviation and MSE, followed by Best W, App. W and F&T. At $\beta = 0.5$, the ratio of the MSE between the Lu's method and OLSE is about 70%.

2) In view of the bias of $\hat{\alpha}$, the WLSE methods outperform OLSE in nearly all cases. Lu's method always generates the smallest bias among the WLSE methods, followed by Best W and App. W. The bias of $\hat{\alpha}$ of the Lu's method is 5 – 10% less than that of the OLSE, and the bias of $\hat{\alpha}$ of Best W and App. W is 3 – 5% less than that of the OLSE.

3) The standard deviation and MSE of $\hat{\alpha}$ of all methods decrease as β_T increases.

4) App. W performs very close to Best W.

Table 6-4: Simulation results of $\hat{\beta}_{1,1}$, generated by different WLSE methods and OLSE at different *n*: the values of $E(\hat{\beta}_{1,1}) \pm S(\hat{\beta}_{1,1})$ and $MSE(\hat{\beta}_{1,1})$ (in parentheses).

Method	n = 5	6	8	10	12	15	18	20
Best W	1.006 ± 0.563 (0.317)	0.972 ± 0.447 (0.201)	0.948 ± 0.335 (0.115)	0.946 ± 0.286 (0.085)	0.945 ± 0.254 (0.068)	0.950 ± 0.222 (0.051)	0.953 ± 0.198 (0.041)	0.958 ± 0.188 (0.037)
App. W	1.006 ± 0.562 (0.316)	0.971 ± 0.447 (0.201)	0.947 ± 0.335 (0.115)	0.946 ± 0.286 (0.085)	0.946 ± 0.254 (0.068)	0.950 ± 0.222 (0.051)	0.955 ± 0.198 (0.041)	0.960 ± 0.189 (0.037)
F&T	1.007 ± 0.562 (0.316)	0.972 ± 0.448 (0.201)	0.948 ± 0.336 (0.116)	0.946 ± 0.287 (0.085)	0.945 ± 0.255 (0.068)	0.950 ± 0.223 (0.052)	0.953 ± 0.199 (0.042)	0.957 ± 0.189 (0.037)
Lu	1.005 ± 0.567 (0.321)	0.971 ± 0.451 (0.204)	0.948 ± 0.339 (0.118)	0.948 ± 0.292 (0.089)	0.949 ± 0.261 (0.070)	0.953 ± 0.228 (0.054)	0.956 ± 0.204 (0.044)	0.961 ± 0.195 (0.040)
OLSE	1.046 ± 0.592 (0.353)	1.009 ± 0.481 (0.231)	0.978 ± 0.370 (0.137)	0.968 ± 0.319 (0.103)	0.962 ± 0.287 (0.084)	0.961 ± 0.255 (0.067)	0.959 ± 0.230 (0.057)	0.961 ± 0.220 (0.051)

Table 6-5: Simulation results of $\hat{\alpha}_{1,\beta_T}$, generated by different WLSE methods and OLSE at different *n* and β_T: the values of $E(\hat{\alpha}_{1,\beta_T}) \pm S(\hat{\alpha}_{1,\beta_T})$ and $MSE(\hat{\alpha}_{1,\beta_T})$ (in parentheses).

	Method	n = 5	6	8	10	12	15	18	20
$\beta_T=0.5$	Best W	1.383 ± 1.344 (1.954)	1.316 ± 1.154 (1.431)	1.219 ± 0.921 (0.896)	1.191 ± 0.807 (0.687)	1.154 ± 0.721 (0.544)	1.129 ± 0.624 (0.406)	1.103 ± 0.557 (0.320)	1.093 ± 0.523 (0.282)
	App. W	1.395 ± 1.356 (1.996)	1.325 ± 1.162 (1.455)	1.224 ± 0.925 (0.906)	1.194 ± 0.808 (0.691)	1.155 ± 0.722 (0.545)	1.127 ± 0.623 (0.405)	1.101 ± 0.556 (0.319)	1.091 ± 0.522 (0.281)
	F&T	1.405 ± 1.365 (2.026)	1.334 ± 1.169 (1.478)	1.233 ± 0.931 (0.921)	1.202 ± 0.813 (0.703)	1.163 ± 0.727 (0.555)	1.135 ± 0.627 (0.412)	1.108 ± 0.559 (0.324)	1.098 ± 0.525 (0.286)
	Lu	1.333 ± 1.302 (1.806)	1.264 ± 1.115 (1.313)	1.167 ± 0.888 (0.817)	1.142 ± 0.780 (0.629)	1.109 ± 0.700 (0.501)	1.088 ± 0.608 (0.377)	1.066 ± 0.544 (0.300)	1.059 ± 0.513 (0.266)
	OLSE	1.528 ± 1.495 (2.512)	1.454 ± 1.286 (1.861)	1.342 ± 1.032 (1.183)	1.304 ± 0.902 (0.906)	1.256 ± 0.812 (0.725)	1.216 ± 0.689 (0.521)	1.181 ± 0.612 (0.408)	1.167 ± 0.573 (0.356)
$\beta_T=1$	Best W	1.067 ± 0.497 (0.251)	1.055 ± 0.447 (0.203)	1.040 ± 0.384 (0.149)	1.035 ± 0.344 (0.120)	1.027 ± 0.311 (0.097)	1.022 ± 0.278 (0.078)	1.020 ± 0.256 (0.066)	1.016 ± 0.240 (0.058)
	App. W	1.072 ± 0.499 (0.254)	1.059 ± 0.449 (0.205)	1.042 ± 0.385 (0.150)	1.036 ± 0.344 (0.120)	1.027 ± 0.311 (0.097)	1.021 ± 0.278 (0.078)	1.019 ± 0.255 (0.066)	1.015 ± 0.240 (0.058)
	F&T	1.076 ± 0.500 (0.256)	1.063 ± 0.450 (0.207)	1.046 ± 0.386 (0.151)	1.039 ± 0.346 (0.121)	1.030 ± 0.312 (0.098)	1.025 ± 0.279 (0.079)	1.023 ± 0.256 (0.066)	1.018 ± 0.241 (0.058)
	Lu	1.047 ± 0.489 (0.242)	1.033 ± 0.440 (0.195)	1.017 ± 0.378 (0.144)	1.012 ± 0.339 (0.115)	1.006 ± 0.307 (0.094)	1.003 ± 0.276 (0.076)	1.003 ± 0.254 (0.064)	1.000 ± 0.239 (0.057)
	OLSE	1.122 ± 0.523 (0.288)	1.109 ± 0.472 (0.235)	1.090 ± 0.407 (0.173)	1.081 ± 0.363 (0.138)	1.069 ± 0.328 (0.112)	1.060 ± 0.293 (0.090)	1.055 ± 0.269 (0.075)	1.049 ± 0.252 (0.066)
$\beta_T=2$	Best W	1.006 ± 0.234 (0.055)	1.004 ± 0.214 (0.046)	1.003 ± 0.186 (0.035)	1.003 ± 0.168 (0.028)	1.003 ± 0.153 (0.024)	1.002 ± 0.136 (0.019)	1.002 ± 0.125 (0.016)	1.002 ± 0.119 (0.016)
	App. W	1.008 ± 0.235 (0.055)	1.006 ± 0.215 (0.046)	1.004 ± 0.186 (0.035)	1.003 ± 0.168 (0.028)	1.004 ± 0.153 (0.024)	1.001 ± 0.136 (0.019)	1.002 ± 0.125 (0.016)	1.002 ± 0.119 (0.016)
	F&T	1.010 ± 0.235 (0.055)	1.008 ± 0.215 (0.046)	1.006 ± 0.187 (0.035)	1.005 ± 0.168 (0.028)	1.005 ± 0.154 (0.024)	1.003 ± 0.136 (0.019)	1.004 ± 0.126 (0.016)	1.004 ± 0.120 (0.016)
	Lu	0.996 ± 0.233 (0.054)	0.993 ± 0.213 (0.045)	0.991 ± 0.185 (0.034)	0.992 ± 0.167 (0.028)	0.993 ± 0.153 (0.023)	0.992 ± 0.136 (0.019)	0.994 ± 0.126 (0.016)	0.996 ± 0.120 (0.016)
	OLSE	1.032 ± 0.240 (0.059)	1.030 ± 0.220 (0.049)	1.027 ± 0.191 (0.037)	1.025 ± 0.172 (0.030)	1.024 ± 0.158 (0.026)	1.020 ± 0.140 (0.020)	1.019 ± 0.129 (0.017)	1.018 ± 0.123 (0.017)

6.4.5 A Bias Correcting Formula for the Proposed Method

The proposed simple formula for calculating weights in Equation (6-21) is limited to small, complete samples. The proposed WLSE method helps to improve the efficiency of parameter estimation, which has been justified by the Monte Carlo experiment. However, the experiment results also show that the shape parameter estimators of the WLSE methods in most cases have larger bias than that of the OLSE. The bias is most significant for very small samples. This can be dangerous. Therefore, a bias correcting formula is proposed for the proposed WLSE method.

The modified Hirose's method presented in Section 5.3.2 for unbiasing the OLSE of the shape parameter can also be used for the WLSE of the shape parameter. As shown in Figure 6-6, the plot of the bias of the WLS estimated β vs. n presents a hyperbolic appearance; therefore, Hirose's bias model in Equation (5-17) can be applied. The process for deriving the five model parameters are same as that presented in Section 5.3.2 and is not repeated here.

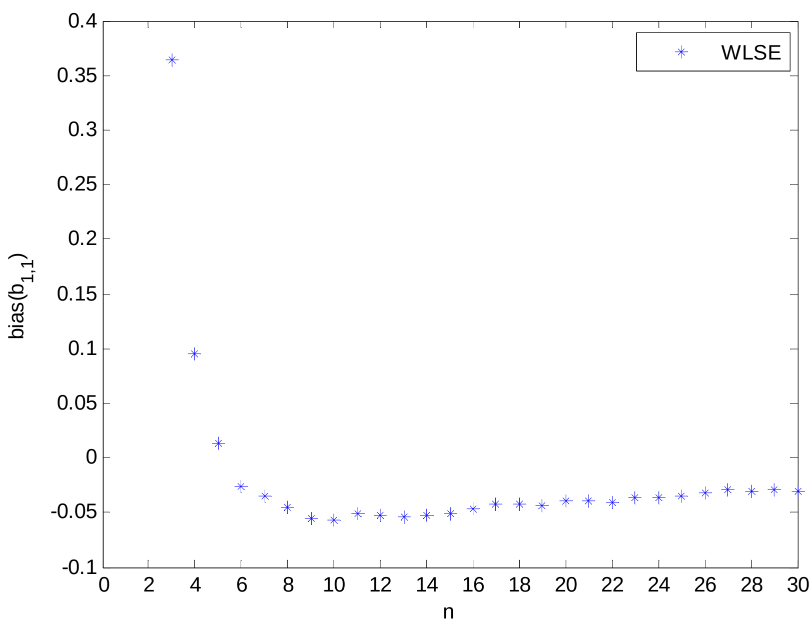


Figure 6-6: Plot of the bias of the proposed WLS estimated $\hat{\beta}_{1,1}$ vs. *n*.

The bias correcting formula for the proposed WLSE of the shape parameter is given by

$$\hat{\beta}_U = \hat{\beta} \cdot \left(0.986 + \frac{1.521}{n} - \frac{8.339}{n^2} + \frac{3.527}{n^3} + \frac{6.345}{n^4} \right) \qquad (6\text{-}22)$$

This equation can be added in the end of the WLSE procedure described in Section 6.4.2.

6.5 Discussions on Large Samples and Censored Samples

As the proposed WLSE method presented in last section is limited to small, complete samples, this section discusses WLSE for large samples and censored samples, respectively.

6.5.1 WLSE for Large Samples

As previously mentioned, it is difficult to calculate weights by Equation (6-16) and then Equation (6-1) when the sample size is large, say $n \geq 30$. For example, MATLAB 7 generates negative values for the weights at $n = 30$, which is obviously wrong. A possible solution for large samples is to use one of the intermediate results in the process of deriving $Var(Y_{(i)})$, as shown in Appendix A, i.e.,

$$\begin{cases} E\left(Y_{(i)}\right) = E\left(Z_{(i)}\right) = i\binom{n}{i}\int_0^{+\infty} \ln v \cdot (e^v - 1)^{i-1} e^{-nv} dv \\ E\left(Y_{(i)}^2\right) = E\left(Z_{(i)}^2\right) = i\binom{n}{i}\int_0^{+\infty} \ln^2 v \cdot (e^v - 1)^{i-1} \cdot e^{-nv} dv \\ Var\left(Y_{(i)}\right) = E\left(Y_{(i)}^2\right) - E^2\left(Y_{(i)}\right) \end{cases} \qquad (6\text{-}23)$$

where $v = e^z$. The Simpson rule (Thisted & Thisted, 1988) may be applied to calculate the integrals in this equation and finally solve the weights.

The F&T's and Lu et al.'s methods, i.e., Equation (6-10) and Equation (6-14), can be used for large samples. However, the accuracy needs to be checked. As shown in Faucher & Tyson's (1988) simulation experiment, Equation (6-10) works well at $n = 100$. Also, Lu et al.'s (2004) simulation experiment showed that Equation (6-14) works well at $n = 50$.

6.5.2 WLSE for Censored Samples

Censored data are commonly encountered in reliability data analysis and it adds the difficulty for parameter estimation. For a censored sample, LSE uses only failure data points to conduct regression, and the influence of the censored items is reflected through the estimation of $\hat{F}_{f,(j)}$, or through the MFON of each failure data point. Several methods have been proposed for calculating the MFON for multiply censored data, as shown in Section 4.3.2. The JM method, i.e., Equation (4-9), is widely used.

Like LSE, the WLSE methods can also be applied to multiply censored data; however, this is less discussed in the literature. Lu et al. (2004), via Monte Carlo simulations, examined several WLSE methods for censored samples of size 20 with 18 different predetermined censoring patterns. The weights for a complete sample of size 20 are selected for the failures in the censored samples based on their event numbers and used directly. The authors concluded that the simulation results for censored samples are in accordance with those for complete data. Obviously, their determination for weights is questionable. The effect of censoring on the failure items is not taken into consideration. To apply WLSE to multiply censored data, instead of the event number, the MFON of the failure data points should be used. The weights for complete samples cannot be directly used for censored samples of the same size.

If the MFON of a failure data point is non-integer, its weight might be calculated by linear interpolation, i.e.,

$$w_{f,j} = w_{Int_j} + (m_{f,(j)} - Int_j)\lfloor w_{Int_j+1} - w_{Int_j} \rfloor \qquad (6\text{-}24)$$

where $m_{f,(j)}$ denotes the MFON of the j^{th} failure and $Int_j = \text{int}[m_{f,(j)}]$ denotes the integral part of $m_{f,(j)}$. For small samples, w_{Int_j} can be calculated by Equation (6-21), and for large samples, it can be calculated by Equation (6-11) or Equation (6-14).

Thus, the step-by-step procedures of WLSE applied to multiply censored samples are given as follows:

***Step 1*:** Calculate $m_{f,(j)}$ and $\hat{F}_{f,(j)}$ for each failure data point using the JM method, i.e., Equation (4-9).

***Step 2*:** Plot the failure times $t_{f,(j)}$ against $\hat{F}_{f,(j)}$ on WPP. If the Weibull distribution fits, the data points should appear to be on a straight line.

***Step 3*:** Calculate the weight for each failure data point based on its MFON. If the MFON is non-integer, the weight is calculated through linear interpolation, i.e., Equation (6-24).

***Step 4*:** Calculate the estimates for α and β using Equation (6-4).

6.5.2.1 A Numerical Example

The following example illustrates the proposed WLSE procedure for a censored sample. This data set, as shown in Table 6-6, has been used for several times, see, e.g., Campean (2000) and Hastings & Bartlett (1997). Table 6-7 shows the spreadsheet used for the calculation of $\hat{\alpha}$ and $\hat{\beta}$.

Table 6-6: A multiply censored data set.

Unit	Failure/Censor	Age (hr)
1	F	112
2	C	213
3	F	250
4	C	484
5	C	500
6	F	572

Table 6-7: The calculation spreadsheet (WLSE for a multiply censored sample).

i	Failure/ Censor	t_i	$\hat{F}_{(i)}$	w_i	j	$m_{f,(j)}$	$\hat{F}_{f,(j)}$	$w'_i = w_{f,j}$	x_i	y_i	$w'_i x_i$	$w'_i y_i$	$w'_i x_i y_i$	$w'_i x_i^2$
1	F	112	0.1094	0.2269	1	1	0.1094	0.2269	4.7185	-2.1556	1.0706	-0.4891	-2.3079	5.0518
2	C		0.2656	0.5761										
3	F	250	0.4219	0.9286	2	2.2	0.2969	0.6466	5.5215	-1.0435	3.5702	-0.6747	-3.7255	19.7126
4	C		0.5781	1.2538										
5	C		0.7344	1.5013										
6	F	572	0.8906	1.5133	3	4.6	0.6719	1.4023	6.3491	0.1083	8.9034	0.1518	0.9641	56.5289
sum								2.2758			13.5442	-1.0120	-5.0693	81.2933

In Table 6-7, $\hat{F}_{(i)}$ and w_i are calculated for a complete sample of size 6. The Bernard estimator is used for $\hat{F}_{(i)}$, i.e., $\hat{F}_{(i)} = (i-0.3)/(n+0.4)$. The values of w_i are extracted from Table 6-1 but can also be calculated by Equation (6-21). The values of $m_{f,(j)}$, $\hat{F}_{f,(j)}$ and w'_i are calculated only for failure data points. The calculations are shown below.

Calculation of $m_{f,(j)}$ (use the JM method) and $\hat{F}_{f,(j)}$:

$$m_{f,(1)} = 1,\ \hat{F}_{f,(1)} = (m_{f,(1)} - 0.3)/(n+0.4) = (1-0.3)/6+0.4) = 0.1094$$

$$m_{f,(2)} = m_{f,(1)} + \frac{n+1-m_{f,(1)}}{I_2^*+1} = 1 + \frac{6+1-1}{4+1} = 2.2,\ \hat{F}_{f,(2)} = (2.2-0.3)/6.4 = 0.2969$$

$$m_{f,(3)} = m_{f,(2)} + \frac{n+1-m_{f,(2)}}{I_3^*+1} = 2.2 + \frac{6+1-2.2}{1+1} = 4.6,\ \hat{F}_{f,(3)} = (4.6-0.3)/6.4 = 0.6719$$

Calculation of w'_i by linear interpolation:

$$w'_1 = w_1 = 0.2269$$

$$w'_2 = w_2 + (m_{f,(2)} - 2) \times (w_3 - w_2) = 0.5761 + (2.2 - 2) \times (0.9286 - 0.5761) = 0.6466$$

$$w'_3 = w_4 + (m_{f,(3)} - 4) \times (w_5 - w_4) = 1.2538 + (4.6 - 4) \times (1.5013 - 1.2538) = 1.4023$$

Calculation of parameter estimates:

x_i and y_i are calculated by $x_i = \ln(t_{f,(j)})$ and $y_i = \ln\left[-\ln(1 - \hat{F}_{f,(j)})\right]$. From Equation (6-4), the parameter estimates are calculated by

$$\begin{cases} \hat{\beta} = \dfrac{\sum_{i=1}^{r} w'_i \cdot \sum_{i=1}^{r} w'_i x_i y_i - \sum_{i=1}^{r} w'_i x_i \cdot \sum_{i=1}^{r} w'_i y_i}{\sum_{i=1}^{r} w'_i \cdot \sum_{i=1}^{r} w'_i x_i^2 - \left(\sum_{i=1}^{r} w'_i x_i \right)^2} \\ \quad = \dfrac{2.2758 \times (-5.0693) - 13.5442 \times (-1.0120)}{2.2758 \times 81.2933 - (13.5442)^2} = 1.3894 \\ \hat{\alpha} = \exp\left(-\dfrac{\sum_{i=1}^{r} w'_i y_i - \hat{\beta} \sum_{i=1}^{r} w'_i x_i}{\hat{\beta} \sum_{i=1}^{r} w'_i} \right) \\ \quad = \exp\left(-\dfrac{-1.0120 - 1.3894 \times 13.5544}{1.3894 \times 2.2758} \right) = 529.2410 \end{cases}$$

6.6 Summary

In this chapter, a simple formula for calculating the weights to be used in WLSE for estimating the two Weibull parameters in the case of small, complete samples of size $n \le 20$ were proposed. Compared to the existing WLSE methods for the Weibull distribution, the proposed method has a better statistical foundation because it is based on the theoretical deduction of the variance of $Y_{(i)}$. The Monte Carlo experiment showed that the proposed method performs closely to the best W method and is slightly better than the other WLSE methods and significantly better than OLSE in view of the standard deviation and MSE for estimating β. For estimating α, the Lu et al.'s (2004) method performs better than the other WLSE methods and OLSE; however, it performs inferior to the other WLSE methods for estimating β. The bias of $\hat{\beta}$ of the proposed WLSE method is larger than that of the OLSE; therefore, a bias correcting formula is proposed using the modified Hirose's method.

The WLSE method for multiply censored data was also proposed, where the weights can be calculated by the modified failure order number. When the MFON is non-integer, the weight can be calculated by linear interpolation. A numerical example clearly illustrated the proposed WLSE procedure for censored data.

Chapter 7

Robust Regression Estimation Methods

This chapter presents a study of using robust regression methods to estimate the Weibull parameters. The robust M-estimation method is proposed and compared with OLSE and MLE via Monte Carlo simulations. Both the case of small data sets with outliers and the case of data sets with multiply censoring are considered. Simulation results show that the proposed method is an effective method in reducing bias and it performs well in most cases with or without outliers.

7.1 Introduction

The quality of data is very important in parameter estimation. Complete data with large sample size are always preferred to achieve a high accuracy on parameter estimation. Unfortunately, reliability engineers often face the problem of small data sets or data sets with censors. In addition, it is also common to have extremely early or late failures in life testing experiments. These harsh data conditions may lead to the estimators of the Weibull parameters, obtained by the traditional methods such as MLE and LSE, to be significantly biased.

In the previous chapter, we have examined the efficiency of the WLSE methods over the OLSE method on Weibull parameter estimation. The proposed WLSE method assumes there is no uncertainty on the failure time so that the weights used are theoretically optimal. Obviously, this is seldom true for field data. Field data may have some outliers, e.g., extremely early or late failures, caused by readout error or irrelevant failure modes, etc. As is well known, the robust regression techniques are good alternatives to the least squares technique when outliers present in a data set. By replacing the LS regression with the robust regression, we call the estimation method the robust regression estimation method.

This chapter is organized as follows. Section 7.1.1 and Section 7.1.2 present the general knowledge including the definition of outliers, types of outliers and common robust regression techniques. Six robust regression techniques and the OLS technique are summarized in a table and compared. Section 7.1.3 overviews the related work of applying robust regression techniques for Weibull parameter estimation. There is very limited work on this topic. Section 7.2 studies the possible outlier configurations of the Weibull samples and presents an important finding which narrows the selection of the robust regression techniques for Weibull parameter estimation. Then, as a preliminary study, the robust M-estimation method is proposed and examined in details, as shown in Section 7.3 and Section 7.4. The simulation results may provide useful information on the use of the robust M-estimation method. Some of the work has been published in Zhang et al. (2006).

7.1.1 Concepts of Outliers

It is not easy to give a mathematically precise definition of an outlier, but there is a commonly used rule, i.e., a point that is at least three or four standard deviations from the center of the data set is considered an outlier (Ryan, 1997). For example, if x_i is suspected as an outlier in a sample, we can exclude it first and calculate the sample mean $\bar{x}$ and sample standard deviation s of the remaining data points, then calculate a standardized value $|x_i - \bar{x}|/s$ for x_i. If this value is large (e.g., > 4), then x_i can be considered as an outlier. Outliers can have many causes, for example, data-entry or recording error. It can also occur because it is truly from another population, or it may present an atypical observation. In general, outliers caused by errors should be discarded from analysis.

Outliers can be classified based on the direction of outlying. Outlying can occur in the *X*-axis direction only, *Y*-axis direction only, or both axes directions simultaneously. Such a point is called an *X*-outlier, a *Y*-outlier or an *X*&*Y*-outlier, respectively. Figure 7-1 illustrates the three types of outliers.

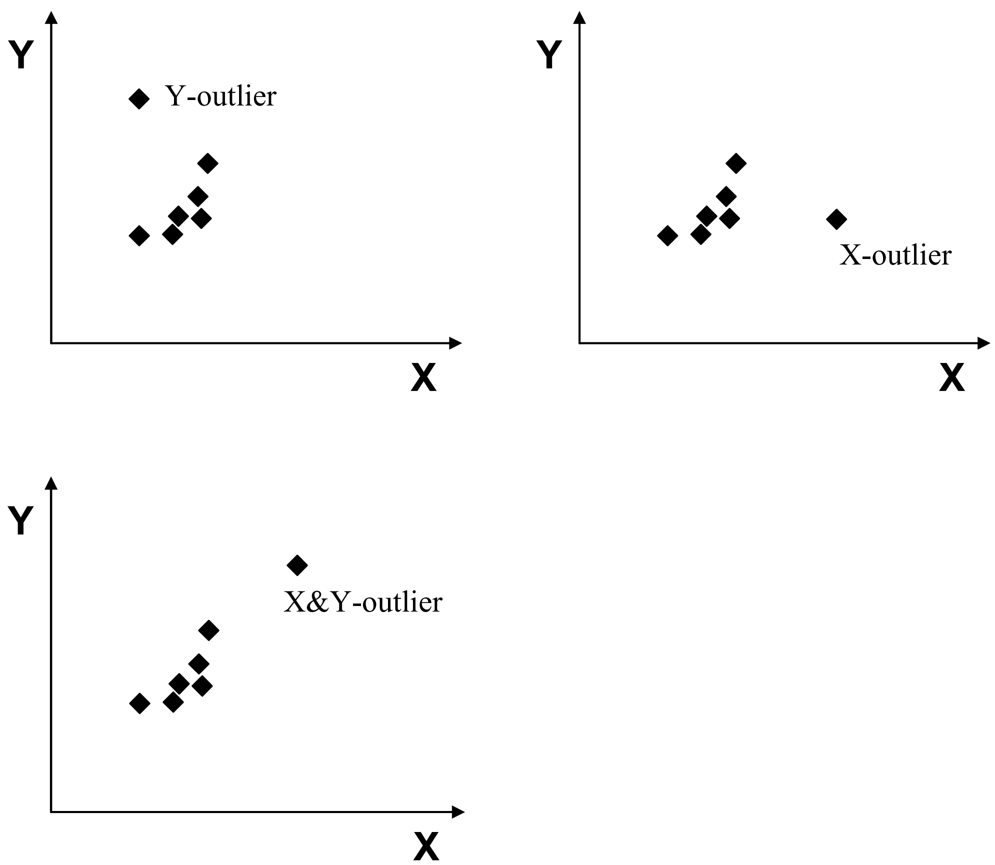


Figure 7-1: Three types of outliers.

7.1.2 Common Robust Regression Techniques

Robust regression techniques are good alternatives to LS that can be appropriately used when there is evidence that the distribution of the error term is (considerably) nonnormal, and/or there are outliers (Ryan, 1997). These techniques aim to reject or limit the influence of the outliers in a sample in order to provide a better fit to the majority of the data points.

Robust regression techniques have a large family. Typical ones include least absolute value (LAV) (Schwarz, 1987), least median of squares (LMS) (Rousseeuw, 1984), least trimmed squares (LTS) (Ruppert & Carroll, 1980), Huber's M-estimation (Huber, 1973), generalized M-estimation (GM-estimation) (Hampel et al., 1986) and MM-estimation (Yohai, 1987). These methods are distinguished by their objective functions and can be assessed by several properties, e.g., efficiency, breakdown point, etc. A brief description and comparison of these methods are given in Anderson & Schumacker (2003). Table 7-1 presents a summary of six commonly used robust regression techniques including LAV, LMS, LTS, M-estimation with unbounded influence function, M-estimation with bounded influence function or GM-estimation, and MM-estimation, together with the OLS technique on several aspects. The table shows the objective functions of each method, their breakdown points (i.e., the smallest fraction of contamination that can cause an estimator to take on values arbitrarily far from its true value), the outlier configurations that they can be applied to, their drawbacks, and their availability in the common statistical software packages.

Table 7-1: Summary of six typical robust regression techniques and OLS.

Regression Method	*Objective Function*	*Breakdown Point*	*Outlier Configurations Apply to*	*Drawbacks*	*Statistical Software Packages Available*
OLS	$\min \sum_{i=1}^{n} e_i^2$	1/n	no influential points or outliers	any outlier can corrupt the estimator	all
Least Absolute Value (LAV)	$\min \sum_{i=1}^{n} \lvert e_i \rvert$	1/n	only Y-axis outliers	can not deal with X-axis outlier (leverage point)	SAS, STATA
Least Median of Squares (LMS)	$\min \ median \ e_i^2$	50%	outlier density is within 50% and sample size is not large	ignores the fit of points other than median, and has a asymptotic relative efficiency of 0	S-PLUS, SAS, STATA
Least Trimmed Squares (LTS)	$\min \sum_{i=1}^{h} e_{(i)}^2$	50% when $h=1/2$	outlier density is within 50%	some good data may be trimmed from computation resulting low efficiency	S-PLUS, SAS, STATA
M-estimation with unbounded influence function	$\min \sum_{i=1}^{n} \rho(e_i)$	1/n	only Y-axis outliers	can not deal with X-axis outlier (leverage point)	MATLAB, S-PLUS, SAS, STATA
M-estimation with bounded influence function (or GM-estimation)	$\min \sum_{i=1}^{n} \rho(e_i)$	50% at best	any type of outliers	"good" leverage points may be down weigted resulting low efficiency	MATLAB, S-PLUS, SAS, STATA
MM-estimation	$\min \sum_{i=1}^{n} \rho(e_i / \hat{s})$	50%	any type of outliers	computation intensive	S-PLUS, SAS, STATA

Most RR methods have both advantages and disadvantages. Some methods, e.g., LAV, are not good at dealing with *X*-outliers. Although robust regression methods usually are computation intensive, many of them are available in common statistical software packages such as S-PLUS, SAS, MATLAB and STATA. For example, S-PLUS 7 has a robust regression library including methods of LMS, LTS and MM-estimation (*S-PLUS 6 Robust Library User's Guide*, 2002). MATLAB 7 has the functions for calculating M-estimators with different weight functions available (*Statistics Toolbox for Use with MATLAB, User's Guide Version 5*, 2004).

The RR methods are still emerging nowadays and it is impossible to examine all of them for estimating the Weibull parameters. On the other hand, since the performance of a RR method is closely related to the outlier configuration, a blind examination of all RR methods should be avoided. In Section 7.2, the special outlier configuration of the Weibull samples is presented. With this finding, some of the RR methods can be excluded from examination.

7.1.3 Related Work

Few papers can be found on the use of RR methods to estimate the Weibull parameters. Lawson et al. (1997) examined the M-estimators (the authors use the term "ML-estimators") for Weibull samples under four outlier conditions: with no outlier or influential data point, with outliers in the right tail area, with outliers in the left tail area, and with two or more near neighbors along the *X*-axis. Different weight functions for the M-estimators were examined including Huber, Andrews, Hampel and Ramsey (Huber, 1973; Andrews et al., 1972; Hampel et al., 1986; Ramsay, 1977) via Monte Carlo simulations. OLSE was also included in the simulation experiment. The comparisons were made on two aspects: model statistics and parameter estimation. The authors concluded that the robust M methods always perform better or at least equally well than OLS in terms of fitting on the probability plot, judged by three model statistics: R^2, MS_{Error} and F-statistic. The Andrews' and the Ramsay's weights are recommended. For parameter estimation, however, the authors found not much difference between M-estimators and OLSE, especially for samples with tail area outliers. This result is disappointing as we expect robust regression methods to perform better. Considering the authors only used three sample sizes and 1000 iteration in their simulation experiments, it is possible that the results are incomplete. In this chapter, we focus on the comparison of the robust M-estimators (with bounded influence functions) with the OLS estimators on Weibull parameter estimation via intensive simulation experiments.

7.2 Special Outlier Configuration of Weibull Samples

As previously mentioned, there are three types of outliers based on the direction of outlying: *X*-outlier, *Y*-outlier and *X&Y*-outlier. Sometimes all three types of outliers can happen in a sample; however, for the Weibull sample, this is not the case. As is well-known, the *X*-axis of the WPP represents the measured values or observations t (i.e., failure time) from a life testing experiment or field. The *Y*-axis of the WPP represents the cumulative probability of failure $F(t)$ at each failure data point. With the use of some non-parametric estimator for $F(t)$, the plotting positions along the *Y*-axis are independent of the values of t along the *X*-axis and can be treated as known constants. Therefore, there is no outlying in the *Y*-axis direction. In other words, there should be no *Y*-outliers and *X&Y*-outliers. Such condition, according to Ryan, (1997), is called fixed regressor case.

This special condition for Weibull samples violates the use of some RR methods that are only robust to the *Y*-outliers, such as LAV and M-estimation with unbounded influence function. However, LMS, LTS, M-estimation with bounded influence functions (or GM-estimation) and MM-estimation are robust to *X*-outliers so that they are the potential candidates for examination. As a preliminary study, this chapter presents the study of M-estimation methods with different bounded ρ functions. The theoretical background of this type of method is presented in the next section.

In the following, for simplicity, the M-estimators refer to the M-estimators with bounded influence functions.

7.3 Robust M-estimators of the Weibull Parameters

7.3.1 Estimating Equation

The M-estimation of Weibull parameters belongs to the simple linear regression context. Let's consider a simple linear regression model $y_i = A + Bx_i + e_i$; for simplicity, the matrix form is used here, i.e., $y_i = \mathbf{x}_i'\theta + e_i$, where $\mathbf{x}_i' = \begin{pmatrix} 1 & x_i \end{pmatrix}$ and $\theta = \begin{pmatrix} A \\ B \end{pmatrix}$. As is known, the objective function of the LS estimators is given by

$$\min \sum_{i=1}^{n} (y_i - \mathbf{x}_i'\theta)^2 = \min \sum_{i=1}^{n} e_i^2 \tag{7-1}$$

The idea of M-estimation is simply to replace the squared residuals e_i^2 by another function of the residuals, thus the objective function of an M-estimator is

$$\min \sum_{i=1}^{n} \rho(y_i - \mathbf{x}_i'\theta) = \min \sum_{i=1}^{n} \rho(e_i) \tag{7-2}$$

where ρ is typically a symmetric, positive-definite function with a unique minimum at zero. The maximum likelihood estimator is a special case when $\rho(e) = -\ln f(e)$ (or $\rho(e) = -\log f(e)$) and hence the name "M-estimation" is used.

To solve Equation (7-2), the normal way is to differentiate the sum of ρ with respect to the two regression coefficients and set the results to zero. This gives

$$\sum_{i=1}^{n} \psi(y_i - \mathbf{x}_i'\theta)\mathbf{x}_i = 0 \text{ or } \sum_{i=1}^{n} \psi(e_i)\mathbf{x}_i = 0 \tag{7-3}$$

where ψ is the first derivative of ρ, i.e., $\psi = d\rho/de$. ψ is called the influence function and it measures the influence of a data point on the value of the parameter estimate. Besides being a bounded function, ψ should satisfy that the robust estimator is unique. To meet this, the residuals need to be standardized by a robust estimate of their scale, denoted by $\hat{\sigma}_e$. Thus the estimating equation becomes

$$\sum_{i=1}^{n} \psi(e_i/\hat{\sigma}_e)\mathbf{x}_i = 0 \tag{7-4}$$

where the median absolute deviation (MAD) is often used for calculating $\hat{\sigma}_e$ and the formula is $\hat{\sigma}_e = 1.4826 \times MAD = 1.4826 \times median\left(\left|\hat{e}_i - median(\hat{e}_i)\right|\right)$.

Define a weight function as $w(e) = \psi(e)/e$, thus Equation (7-4) becomes

$$\sum_{i=1}^{n} w(e_i/\hat{\sigma}_e) \cdot e_i \cdot \mathbf{x}_i = 0 \tag{7-5}$$

Different functions for $\rho(e)$ and $w(e)$ (or $\psi(e)$) have been proposed. Table 7-2 lists some of them. The Huber's (Huber, 1973) and Tukey's biweight (also known as bisquare) (Beaton & Tukey, 1974) functions are two common choices.

Table 7-2: Typical ρ functions and weight functions used in the M-estimation method.

	$\rho(e)$	$w(e)$	k
Huber	$\rho = \begin{cases} \frac{e^2}{2} & \text{for } \lvert e\rvert \le k \\ k\left(\lvert e\rvert - \frac{k}{2}\right) & \text{for } \lvert e\rvert > k \end{cases}$	$w = \begin{cases} 1 & \text{for } \lvert e\rvert \le k \\ \frac{k}{\lvert e\rvert} & \text{for } \lvert e\rvert > k \end{cases}$	2.0
Bisquare	$\rho = \begin{cases} \frac{k^2}{6}\left\{1-\left[1-\left(\frac{e}{k}\right)^2\right]^3\right\} & \text{for } \lvert e\rvert \le k \\ \frac{k^2}{6} & \text{for } \lvert e\rvert > k \end{cases}$	$w = \begin{cases} \left[1-\left(\frac{e}{k}\right)^2\right]^2 & \text{for } \lvert e\rvert \le k \\ 0 & \text{for } \lvert e\rvert > k \end{cases}$	4.685
Andrews	$\rho = \begin{cases} k\left[1-\cos\left(\frac{e}{k}\right)\right] & \text{for } \lvert e\rvert \le \pi \cdot k \\ 2k & \text{for } \lvert e\rvert > \pi \cdot k \end{cases}$	$w = \begin{cases} \frac{\sin(e/k)}{e/k} & \text{for } \lvert e\rvert \le \pi \cdot k \\ 0 & \text{for } \lvert e\rvert > \pi \cdot k \end{cases}$	1.339
Cauchy	$\rho = \frac{k^2}{2} \cdot \log\left[1+\left(\frac{e}{k}\right)^2\right]$	$w = \frac{1}{1+(e/k)^2}$	2.385
Hampel	$\rho = \begin{cases} \frac{e^2}{2} & \text{for } \lvert e\rvert \le k \\ k\left(\lvert e\rvert - \frac{k}{2}\right) & \text{for } k \le \lvert e\rvert \le k' \\ \left[k \cdot \frac{\left(k'' \cdot \lvert e\rvert - \frac{e^2}{2}\right)}{k'' - k'}\right] - \frac{7}{6}k^2 & \text{for } k' \le \lvert e\rvert \le k'' \\ k \cdot (k' + k'' - k) & \text{for } \lvert e\rvert > k'' \end{cases}$	$w = \begin{cases} 1 & \text{for } \lvert e\rvert \le k \\ \frac{k}{\lvert e\rvert} & \text{for } k \le \lvert e\rvert \le k' \\ k \cdot \frac{(k'' - \lvert e\rvert)}{\lvert e\rvert \cdot (k'' - k')} & \text{for } k' \le \lvert e\rvert \le k'' \\ 0 & \text{for } \lvert e\rvert > k'' \end{cases}$	$k = 1.7$ $k' = 3.4$ $k'' = 8.5$
Welsch	$\rho = \frac{k^2}{2} \cdot \left[1-\exp\left(-\left(\frac{e}{k}\right)^2\right)\right]$	$w = \exp\left(-\left(\frac{e}{k}\right)^2\right)$	2.985

The estimating equation, Equation (7-4) or Equation (7-5), can be solved by the iteratively reweighted least squares method (see, e.g., Holland & Welsch, 1977). The procedure of the method is summarized as follows.

Step 1: Select the initial estimates, for example, using the least squares estimates.

Step 2: Compute the residuals.

Step 3: Calculate weights and solve the weighted least squares estimates.

Step 4: Recalculate the residuals.

Step 5: Repeat Steps 3 and 4 until the estimates convergence.

Same as OLSE, the robust M-estimation methods enjoy graphical presentation, i.e., the WPP. The Weibull shape parameter is the slope of the regression line generated by the robust regression method.

7.3.2 Practical Application with Statistical Software

Since the estimating equation of the M-estimation method has to be solved iteratively until the convergence is reached, the computation can be highly complicated; however, this is not a big problem nowadays as several statistical software packages have functions or dialogs of various robust M-estimation methods. MATLAB 7 is used in this study and it has a function, *robustfit*, to generate the M-estimates directly. The syntax (*Statistics Toolbox for Use with MATLAB, User's Guide Version 5*, 2004) is given by

$$\left[\mathrm{b, stats}\right] = \mathrm{robustfit}(x, y, wfun, tune) \tag{7-6}$$

The left side of the equation is the output, where b returns the M-estimates of the regression coefficients, *stats* is optional and it includes several statistical measures such as the standard errors of the coefficient estimates. The right side of the equation is the input, where for the estimation of Weibull parameters, $x = \ln t$ and

$y = \ln\left[-\ln(1-\hat{F})\right]$ should be provided, same as in the OLSE method. *wfun* is the weight function; by default the bisquare weight is used but we are free to change it to 'andrews', 'cauchy', 'fair', 'huber', 'logistic', 'talwar' and 'welsch'. *Tune* is the tuning constant related to the weight function and it has a default value for each of them.

Besides MATLAB, the *robustreg* procedure in SAS 9 and the *rreg* command in STATA 11 can also be used to generate M-estimates for a data set with no difficulty. SAS 9 provides ten weight functions and the bisquare weight is still the default one. STATA 11, however, does not offer the selection of the weight functions and use the bisquare only.

7.3.3 Numerical Examples

Example 1 (A Complete Data Set with An Extremely Early Failure)

In this example, ten fatigue specimens were put on test and all tested to failure. The failure times in hours are as follows: 150, 50, 250, 240, 135, 200, 240, 150, 200, and 190. This data set is used in Abernethy (2000) but we modified the second observation to 50 to generate an extremely early failure. Early failures are very common in life testing and it can be caused by many reasons, for example, the experiment conditions are unstable at the beginning, or the failure is caused by other failure modes that are not of concern.

The robust M-estimation method (with the bisquare weight) and the OLSE method were used to estimate the shape parameter for this data set, and the results are 3.781 and 2.123, respectively. Figure 7-2 is the WPP for the data set, where the regression lines are generated by the two methods. It can be seen that the first data point (which can be considered as an outlier) moves the OLS regression line toward it while the M-estimation regression line is nearly unaffected by it and fits the other data points well. The OLSE method results in the highly over-estimated shape parameter estimate.

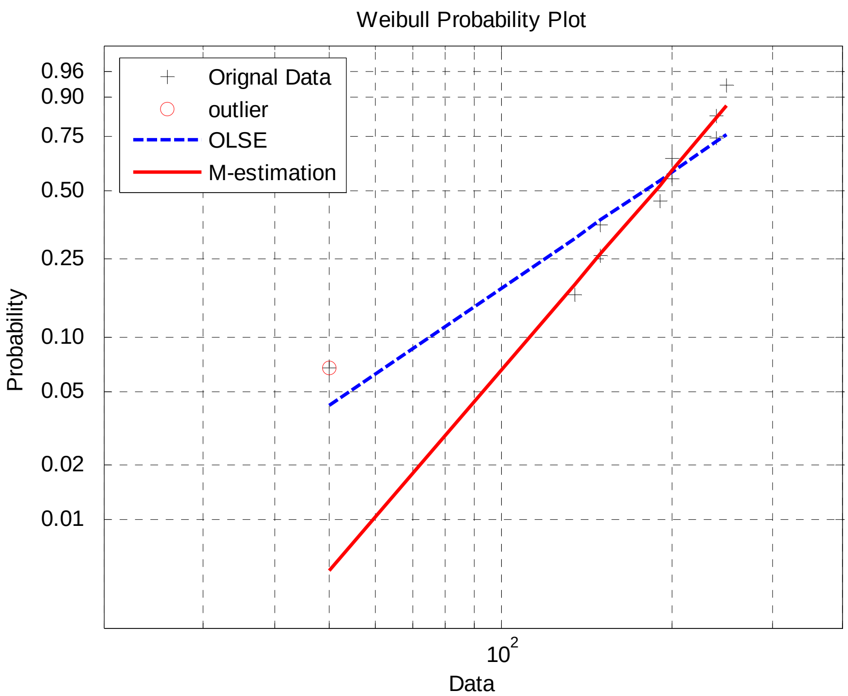


Figure 7-2: A numerical example to compare OLSE and robust M-estimation with WPP in the case of complete data.

Example 2 (A Multiply Censored Data Set)

Censored data often add difficulty to parameter estimation, even if there is no outlier. This sample, as shown in Table 7-3, was randomly generated from the Weibull distribution with $\alpha = 1000$ and $\beta = 1.5$.

Table 7-3: A computer-generated multiply censored example ("F" denotes failure and "C" denotes censor).

54.6	1077.6	831.4	134.4	172.8	1749.5	189.7	1385.5	820.6	13.2
C	C	F	C	C	F	F	C	F	C
685.7	578.8	596.1	1182.4	1081	497.7	375.4	2008.5	951.5	135.1
F	C	C	C	F	C	C	F	F	C

The robust M-estimation method (with the bisquare weight) and the OLSE method were used to estimate the shape parameter for this data set, and the results are 1.307 and 0.927, respectively. Figure 7-3 shows the WPP. As can be seen from the plot, the first failure data point in this sample is far from the others, and it moves the OLS regression line toward it. The M-estimation regression line is less affected by this point and fits the majority of data points well. The OLSE method results in a under-estimated β for this sample.

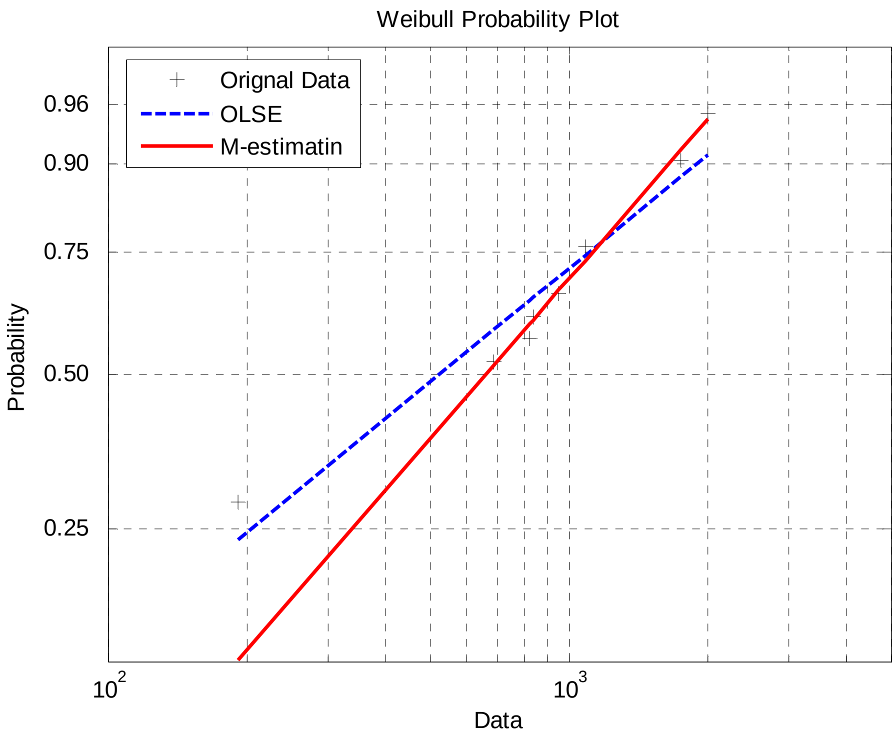


Figure 7-3: A numerical example to compare OLSE and robust M-estimation with WPP in the case of censored data.

7.4 Monte Carlo Study of the Robust M-estimators of the Shape Parameter

Monte Carlo simulation experiments have been carried out to compare the performance of the OLSE and the robust M-estimation methods on parameter estimation when dealing with small, complete data sets with outliers, and multiply censored data sets. Different weight functions including bisquare, Andrews, Cauchy and Welsch, were examined for the robust M-estimation methods. The selection of the four weight functions comes from their popularity and availability in MATLAB 7. The MLE method is also included in the comparison due to its wide application.

As shown in Section 7.2, it only makes sense to have *X*-outliers for Weibull samples. Given this, four outlier configurations were generated in this experiment including one left tail *X*-outlier, one right tail *X*-outlier, two left tail *X*-outliers and two right tail *X*-outliers. The method to generate samples with these types of outliers is as follows: Firstly, generate a random, complete Weibull sample following the first two steps in the procedure described in Section 3.3.1; Secondly, calculate the standard deviation of this sample; Finally, to generate one left/right tail *X*-outlier, shift the first/last failure data point in the original sample four standard deviations (of the original sample) to the left/right in the *X*-axis direction, or, to generate two left/right tail *X*-outliers, simultaneously shift the first/last two failure data points in the original sample in such way.

Because multiply censored data often have large scatter and involve influential points, it is less important to further add *X*-outliers to the randomly generated samples. It is expected that, if the robust M-estimation methods perform well when there is no real outliers, it will surely perform well when there are.

The setting of experiment factors is given in Table 7-4. For each combination of the simulation factors, for example, $\alpha_T = 1$, $\beta_T = 0.5$, $n = 5$ and complete sample with one left tail *X*-outlier, 10000 random samples were generated and parameter estimates were obtained from OLSE, four M-estimation methods, and MLE. The results of the shape parameter estimators are the focus of this experiment. The mean, standard deviation and MSE of the parameter estimates were calculated and analyzed. The experiment was executed in MATLAB 7. Simulation results are presented in the following sections for complete data and censored data, respectively.

Table 7-4: Setting of experiment factors. The experiment is to examine robust M-estimators and compare them with OLSE and MLE.

Factors	Values
α_T	1
β_T	0.5, 1, 2, 4, 10
n	5, 6, …, 10, 15, 18, 20, 25, 30 (for complete data) 10, 20, 30, 50, 80, 100 (for censored data)
c	20%, 40%, 60%, 80%
Outlier type	left tail X-outliers and right tail X-outliers
M	10000
Methods	M-estimation methods (bisquare, Andrews, Cauchy and Welsch), OLSE, MLE

7.4.1 Simulation Results for Complete Samples with Outliers

General Observations

1) The four M-estimators associated with different weight functions including bisquare, Andrews, Cauchy and Welsch perform similar.
2) The comparison result for the outlier configuration type one left X-outlier is similar to that of one right X-outlier, and the result for two left X-outliers is similar to that of two right X-outliers.

Based on the above two observations, the simulation results are only partially tabulated, as shown in Table 7-5 and Table 7-6. M-estimator with the bisquare weight function is selected to represent the performance of the robust M-estimator, and the results for the outlier configurations of one left X-outlier and two left X-outliers are omitted. The omitted results will not affect the following conclusions which can be observed from the tabulated values.

Simulation Results for Data Sets with One X-outlier (Table 7-5)

1) M-estimator performs best in view of bias among the three estimators in most times except when $\beta_T = 4, 10$ and $n = 5, 6$. MLE performs best when $\beta_T = 4, 10$ and $n = 5, 6$.

2) Compare to OLSE, M-estimator has smaller bias in almost all combinations of n and β_T. The differences in bias between the two estimators are small at $n = 5, 6$, but become significant as n and β_T increase. OLSE is highly biased at all sample sizes when $\beta_T = 10$, while the bias of the M-estimator is within 5% when $\beta_T = 10$ and $n \geq 8$. At all β_T, the bias of the M-estimator is within 10% when $n \geq 10$. The differences in MSE between the two estimators are small when $\beta_T = 0.5, 1, 2$, but the MSE of the M-estimator is much smaller when $\beta_T = 4, 10$ and $n \geq 8$.

3) Compare to MLE, M-estimator is significantly better when $\beta_T = 0.5, 1$ and $n \leq 10$ in view of both bias and MSE.

4) The bias of all the estimators is decreasing with the increase of sample size. However, the bias is inconsistent with β_T.

Table 7-5: Simulation results of $\hat{\beta}$ for complete samples with one right tail *X*-outlier: the values of $E(\hat{\beta}_{1,1}) \pm S(\hat{\beta}_{1,1})$ and $MSE(\hat{\beta}_{1,1})$ (in parentheses).

		n						
	Method	5	6	8	10	15	20	30
β_T=0.5	OLSE	0.291 ± 0.314 (0.142)	0.403 ± 0.140 (0.029)	0.421 ± 0.133 (0.024)	0.438 ± 0.120 (0.018)	0.453 ± 0.105 (0.013)	0.462 ± 0.093 (0.010)	0.470 ± 0.080 (0.007)
	M-estimator (bisquare)	0.289 ± 0.314 (0.143)	0.404 ± 0.150 (0.032)	0.441 ± 0.182 (0.036)	0.456 ± 0.155 (0.026)	0.465 ± 0.127 (0.017)	0.470 ± 0.106 (0.012)	0.475 ± 0.086 (0.008)
	MLE	0.944 ± 0.786 (0.815)	1.023 ± 0.220 -(0.049)	1.020 ± 0.191 (0.037)	1.018 ± 0.172 (0.030)	1.012 ± 0.140 (0.020)	1.012 ± 0.123 (0.017)	1.012 ± 0.123 (0.017)
β_T=1	OLSE	0.627 ± 0.218 (0.186)	0.652 ± 0.203 (0.162)	0.719 ± 0.199 (0.119)	0.775 ± 0.190 (0.087)	0.847 ± 0.179 (0.056)	0.874 ± 0.166 (0.043)	0.914 ± 0.148 (0.029)
	M-estimator (bisquare)	0.626 ± 0.220 (0.188)	0.658 ± 0.227 (0.168)	0.859 ± 0.421 (0.197)	0.910 ± 0.370 (0.145)	0.941 ± 0.283 (0.083)	0.932 ± 0.227 (0.056)	0.948 ± 0.179 (0.035)
	MLE	1.445 ± 0.789 (0.820)	1.343 ± 0.620 (0.502)	1.222 ± 0.433 (0.237)	1.172 ± 0.362 (0.160)	1.112 ± 0.257 (0.079)	1.073 ± 0.202 (0.046)	1.050 ± 0.156 (0.027)
β_T=2	OLSE	1.436 ± 0.155 (0.343)	0.934 ± 0.293 (1.222)	1.080 ± 0.283 (0.926)	1.217 ± 0.283 (0.694)	1.436 ± 0.284 (0.399)	1.557 ± 0.277 (0.273)	1.704 ± 0.264 (0.158)
	M-estimator (bisquare)	1.436 ± 0.156 (0.343)	0.959 ± 0.342 (1.202)	1.674 ± 0.875 (0.873)	1.927 ± 0.735 (0.545)	1.914 ± 0.541 (0.299)	1.905 ± 0.449 (0.211)	1.913 ± 0.356 (0.135)
	MLE	2.430 ± 0.771 (0.778)	2.678 ± 1.184 (1.861)	2.436 ± 0.831 (0.880)	2.352 ± 0.717 (0.638)	2.223 ± 0.513 (0.313)	2.155 ± 0.411 (0.193)	2.099 ± 0.315 (0.109)
β_T=4	OLSE	3.270 ± 0.112 (0.546)	1.145 ± 0.415 (8.326)	1.373 ± 0.408 (7.068)	1.593 ± 0.424 (5.973)	2.073 ± 0.456 (3.921)	2.403 ± 0.465 (2.767)	2.875 ± 0.460 (1.476)
	M-estimator (bisquare)	3.270 ± 0.113 (0.546)	1.162 ± 0.443 (8.253)	3.434 ± 2.033 (4.451)	3.861 ± 1.367 (1.887)	3.813 ± 1.026 (1.087)	3.813 ± 0.882 (0.812)	3.836 ± 0.700 (0.518)
	MLE	4.449 ± 0.769 (0.794)	5.323 ± 2.373 (7.383)	4.896 ± 1.767 (3.923)	4.647 ± 1.399 (2.376)	4.388 ± 0.992 (1.134)	4.302 ± 0.823 (0.769)	4.207 ± 0.643 (0.457)
β_T=10	OLSE	1.562 ± 0.653 (71.620)	1.521 ± 0.530 (72.171)	1.647 ± 0.546 (70.074)	1.855 ± 0.604 (66.707)	2.460 ± 0.754 (57.421)	3.064 ± 0.879 (48.875)	4.224 ± 1.023 (34.414)
	M-estimator (bisquare)	1.565 ± 0.656 (71.586)	1.527 ± 0.538 (72.079)	9.873 ± 4.339 (18.844)	9.820 ± 3.469 (12.065)	9.554 ± 2.615 (7.038)	9.538 ± 2.232 (5.196)	9.543 ± 1.726 (3.187)
	MLE	14.337 ± 7.327 (72.496)	13.286 ± 6.058 (47.501)	12.225 ± 4.453 (24.780)	11.704 ± 3.487 (15.060)	11.008 ± 2.536 (7.448)	10.775 ± 2.094 (4.987)	10.477 ± 1.551 (2.634)

Simulation Results for Data Sets with Two X-outliers (Table 7-6)

1) M-estimator performs best among the three estimators when $\beta_T = 0.5$ and $n \geq 15$ in view of both bias and MSE.

2) M-estimator outperforms the OLSE in view of both bias and MSE in almost all combinations of n and β_T, even when $n = 5, 6$. The differences between the two estimators increase as β_T increases.

3) Both OLSE and M-estimator perform badly when $\beta_T = 10$ and $n \leq 20$, and their bias and MSE are much larger than those of the MLE. As the sample size increases, say at $n = 30$, however, the bias and MSE of the M-estimator

is comparable to those of the ML estimator, but those of the OLSE is still unacceptable.

4) MLE is better than the two linear regression methods in most cases, especially when sample size is small, say $n < 20$.

5) The bias of all three estimators decreases as the sample size increase. However, the bias is inconsistent with β_T.

Table 7-6: Simulation results of $\hat{\beta}$ for complete samples with two right tail *X*-outliers: the values of $E(\hat{\beta}_{1,1}) \pm S(\hat{\beta}_{1,1})$ and $MSE(\hat{\beta}_{1,1})$ (in parentheses).

		n					
	Method	10	15	18	20	25	30
β_T=0.5	OLSE	0.374 ± 0.091 (0.024)	0.408 ± 0.084 (0.016)	0.421 ± 0.081 (0.013)	0.428 ± 0.081 (0.012)	0.441 ± 0.075 (0.009)	0.449 ± 0.071 (0.008)
	M-estimator (bisquare)	0.377 ± 0.106 (0.026)	0.425 ± 0.115 (0.019)	0.445 ± 0.117 (0.017)	0.452 ± 0.114 (0.015)	0.460 ± 0.100 (0.012)	0.465 ± 0.090 (0.009)
	MLE	0.584 ± 0.179 (0.039)	1.023 ± 0.220 (0.049)	1.020 ± 0.191 (0.037)	1.018 ± 0.172 (0.030)	1.012 ± 0.140 (0.020)	1.012 ± 0.123 (0.017)
β_T=1	OLSE	0.583 ± 0.132 (0.191)	0.686 ± 0.135 (0.117)	0.729 ± 0.131 (0.091)	0.749 ± 0.131 (0.080)	0.791 ± 0.129 (0.060)	0.826 ± 0.123 (0.045)
	M-estimator (bisquare)	0.604 ± 0.171 (0.187)	0.840 ± 0.279 (0.103)	0.908 ± 0.266 (0.079)	0.923 ± 0.248 (0.067)	0.932 ± 0.217 (0.052)	0.940 ± 0.185 (0.038)
	MLE	1.167 ± 0.350 (0.150)	1.106 ± 0.257 (0.077)	1.093 ± 0.227 (0.060)	1.077 ± 0.207 (0.049)	1.059 ± 0.181 (0.036)	1.047 ± 0.158 (0.027)
β_T=2	OLSE	0.782 ± 0.201 (1.525)	0.987 ± 0.212 (1.072)	1.085 ± 0.221 (0.886)	1.142 ± 0.229 (0.788)	1.267 ± 0.227 (0.588)	1.362 ± 0.216 (0.454)
	M-estimator (bisquare)	0.816 ± 0.265 (1.471)	1.640 ± 0.591 (0.479)	1.851 ± 0.512 (0.285)	1.888 ± 0.484 (0.247)	1.911 ± 0.414 (0.180)	1.913 ± 0.377 (0.150)
	MLE	2.329 ± 0.694 (0.590)	2.215 ± 0.511 (0.308)	2.177 ± 0.447 (0.231)	2.161 ± 0.421 (0.203)	2.126 ± 0.359 (0.145)	2.100 ± 0.324 (0.115)
β_T=4	OLSE	0.910 ± 0.282 (9.627)	1.179 ± 0.322 (8.060)	1.329 ± 0.351 (7.256)	1.423 ± 0.363 (6.771)	1.673 ± 0.388 (5.565)	1.868 ± 0.402 (4.708)
	M-estimator (bisquare)	0.905 ± 0.307 (9.673)	2.431 ± 1.340 (4.259)	3.354 ± 1.218 (1.902)	3.643 ± 1.000 (1.128)	3.820 ± 0.815 (0.697)	3.824 ± 0.731 (0.565)
	MLE	4.663 ± 1.418 (2.450)	4.427 ± 1.014 (1.211)	4.356 ± 0.891 (0.922)	4.302 ± 0.803 (0.736)	4.250 ± 0.726 (0.590)	4.192 ± 0.641 (0.448)
β_T=10	OLSE	1.295 ± 0.275 (75.856)	1.408 ± 0.337 (73.934)	1.525 ± 0.397 (71.978)	1.600 ± 0.430 (70.752)	1.842 ± 0.523 (66.823)	2.069 ± 0.606 (63.267)
	M-estimator (bisquare)	1.286 ± 0.273 (76.010)	1.651 ± 1.275 (71.338)	3.504 ± 3.215 (52.540)	5.251 ± 3.756 (36.658)	8.500 ± 2.893 (10.619)	9.381 ± 1.972 (4.271)
	MLE	11.675 ± 3.506 (15.096)	11.063 ± 2.521 (7.487)	10.821 ± 2.203 (5.528)	10.755 ± 2.048 (4.767)	10.601 ± 1.780 (3.530)	10.490 ± 1.583 (2.745)

7.4.2 Simulation Results for Censored Data

The simulation results for censored data are presented in Table 7-7. The following conclusions can be observed.

1) In general, M-estimator performs better than OLSE in view of both bias and MSE.

2) In view of both bias and MSE, M-estimator and OLSE perform better than MLE in most conditions.

3) MLE performs slightly better than the M-estimator when $\beta_T = 2$ and $n \geq 30$, and significantly better when $\beta_T = 10$. M-estimator and OLSE perform badly when $\beta_T = 10$. The increase of sample size does not improve their performance.

4) The estimator of MLE deteriorates as the censoring level increases, but the estimators of OLSE and M-estimator are inconsistent with the censoring level.

Table 7-7: Simulation results of $\hat{\beta}$ for multiply censored samples, generated by robust M-estimation, OLSE and MSE: the values of $E(\hat{\beta}_{1,1}) \pm S(\hat{\beta}_{1,1})$ and $MSE(\hat{\beta}_{1,1})$ (in parentheses).

c	Method	*n* = 20	30	50	80	100
20%	OLSE	0.908 ± 0.227 (0.060)	0.927 ± 0.192 (0.042)	0.948 ± 0.154 (0.026)	0.963 ± 0.125 (0.017)	0.973 ± 0.113 (0.014)
	M-estimator (bisquare)	0.943 ± 0.238 (0.049)	0.966 ± 0.200 (0.038)	0.991 ± 0.158 (0.010)	1.008 ± 0.130 (0.009)	1.021 ± 0.118 (0.004)
	MLE	1.123 ± 0.234 (0.070)	1.094 ± 0.183 (0.042)	1.069 ± 0.134 (0.023)	1.055 ± 0.102 (0.013)	1.052 ± 0.091 (0.011)
40%	OLS	0.930 ± 0.269 (0.078)	0.945 ± 0.216 (0.050)	0.970 ± 0.176 (0.032)	0.988 ± 0.143 (0.021)	0.994 ± 0.132 (0.017)
	M-estimator (bisquare)	0.958 ± 0.288 (0.066)	0.971 ± 0.229 (0.037)	1.001 ± 0.187 (0.016)	1.022 ± 0.153 (0.003)	1.029 ± 0.141 (0.002)
	MLE	1.195 ± 0.300 (0.128)	1.148 ± 0.217 (0.069)	1.121 ± 0.159 (0.040)	1.103 ± 0.120 (0.025)	1.096 ± 0.108 (0.021)
60%	OLS	0.971 ± 0.339 (0.116)	0.976 ± 0.268 (0.072)	0.999 ± 0.216 (0.046)	1.014 ± 0.177 (0.031)	1.023 ± 0.161 (0.026)
	M-estimator (bisquare)	0.988 ± 0.361 (0.111)	0.995 ± 0.283 (0.062)	1.018 ± 0.226 (0.035)	1.034 ± 0.187 (0.018)	1.044 ± 0.169 (0.013)
	MLE	1.298 ± 0.408 (0.255)	1.233 ± 0.291 (0.139)	1.191 ± 0.205 (0.079)	1.166 ± 0.156 (0.052)	1.158 ± 0.138 (0.044)
80%	OLS	1.111 ± 0.621 (0.398)	1.057 ± 0.410 (0.171)	1.053 ± 0.312 (0.100)	1.053 ± 0.247 (0.064)	1.059 ± 0.226 (0.054)
	M-estimator (bisquare)	1.111 ± 0.621 (0.398)	1.057 ± 0.415 (0.172)	1.066 ± 0.326 (0.092)	1.065 ± 0.257 (0.056)	1.071 ± 0.234 (0.046)
	MLE	1.524 ± 0.779 (0.882)	1.395 ± 0.497 (0.403)	1.313 ± 0.341 (0.215)	1.262 ± 0.241 (0.127)	1.251 ± 0.211 (0.107)

7.5 Summary

Robust regression methods provide another alternative to OLS to fit the regression line on WPP. Common robust regression methods can be easily applied with many statistical software packages.

The results of this study indicated that the robust M-estimator of the Weibull shape parameter almost always outperforms the OLS estimator for small, complete samples with one X-outlier or two X-outliers in the right or left tail. The differences in bias between the M-estimator and OLSE become significant as n and β_T increase. For samples with one X-outlier, M-estimator performs best in most cases except at the combinations of very large β_T and very small n. For samples with two X-outliers, the M-estimator performs best when $\beta_T \leq 1$, while the ML estimator is the best in most cases, especially when the sample size is very small and β_T is very large. Finally, for multiply censored samples, M-estimator also performs better than OLSE in view of both bias and MSE, and they perform better than MLE in most cases. In general, M-estimator outperforms OLSE and thus should be recommended for use.

The robust regression methods are highly dependent on the outlier configurations, and may not provide better estimates than OLSE even when outliers exist in the samples. We recommend that the OLSE and RRE methods should be used always with WPP to judge their performance.

Chapter 8

A Procedure for Implementation of Linear Regression Estimation Methods and Case Studies

This chapter presents a procedure which serves the purpose to guide the practitioners on the selection of linear regression estimation methods, among those discussed in this thesis, for different types of data. Case studies are provided to further illustrate the application process.

8.1 Introduction

As mentioned in the beginning of this thesis, the analysis of life data is complex because different types of data require different approaches of processing. This is particularly true for parameter estimation. Accurate parameter estimates contribute to an appropriate model for life data, and the parameter estimation results can directly affect other aspects of life data analysis and hence have great impacts on reliability-related activities and even business decisions. Therefore, the selection of parameter estimation methods is very important in life data analysis.

In the previous chapters, various linear regression estimation methods for the Weibull distribution have been presented. The step-by-step procedures were provided for these methods so that there is no difficulty to apply them if the practitioners are told which method to use. In this chapter, some suggestions on the selection of the estimation methods, among those discussed in this thesis, under different data conditions, are presented. Three case studies are also presented for illustration.

Different from the numerical examples in the previous chapters, which are mostly computer generated, here the cases selected are more like from the real conditions.

8.2 Implementation Procedure on the Use of Linear Regression Estimation Methods

A flowchart is proposed to illustrate the process for selecting an appropriate linear regression estimation method, as shown in Figure 8-1. The foundation of this chart is the results, both analytical and experimental, presented in the previous chapters. It mainly serves the purpose to provide accurate shape parameter estimates because the shape parameter is usually more important than the scale parameter. The process is described as follows.

The process begins when one have a data set consisting of several observations, i.e., failure times and censoring times. First, draw a WPP for this data set to check whether the data are Weibull distributed. Note that WPP is a simple model validation tool and may not be accurate. If the majority of data points do not nicely form a straight line, before reject the Weibull distribution assumption, it is necessary to use specially designed goodness-of-fit tests, e.g., Chi-Square goodness-of-fit, to check again.

If there is no doubt on the Weibull distribution assumption, then use the WPP to check whether there are outliers or influential points in the sample. The judgment, however, is subjective. If we suspect there are one or more outliers, the RRE methods should be used for parameter estimation. It is not recommended to remove the outliers or influential points from analysis because data are precious and every data point conveys information.

If there are no outliers or influential points in the sample, check whether this is a complete sample or a multiply censored sample. Note that the sample can also be a singly Type I or Type II censored sample, and it is suggested that the same procedure for multiply censored data be applied to singly censored data, because singly censored data can be treated as a special case of multiply censored data. For multiply censored samples, the selection of estimation methods mainly depends on the censoring level of the sample. It is suggested that LS *Y* on *X* with the HJ estimator, i.e., the OLSE method, be used for a highly censored sample ($c \geq 50\%$), and LS *X* on *Y* with the JM estimator for a lowly censored sample ($c < 50\%$). This is based on the simulation results presented in Section 4.4.3.

On the other hand, if the sample is a complete sample, the selection of estimation methods is based on the sample size. If it is a small sample with $n \leq 10$, the OLSE method is recommended; if $n > 10$, LS *Y* on *X* with the Ross estimator or LS *X* on *Y* with the Bernard estimator is recommended. For small samples with $n \leq 20$, the WLSE methods can also be used as a supplementary.

It is important to point out that the flowchart is mainly based on the examination results on the bias, standard deviation and MSE of the linear regression estimators. Therefore, it is correct in the long run but may not be correct for a single Weibull sample. In fact, no estimation method can always provide accurate point estimates for any sample. Facing this problem, it is important to improve data collection methods including data recording, instrumentation calibrations, etc., and try to reduce the scatter of data, eliminate outliers or identify the causes for them.

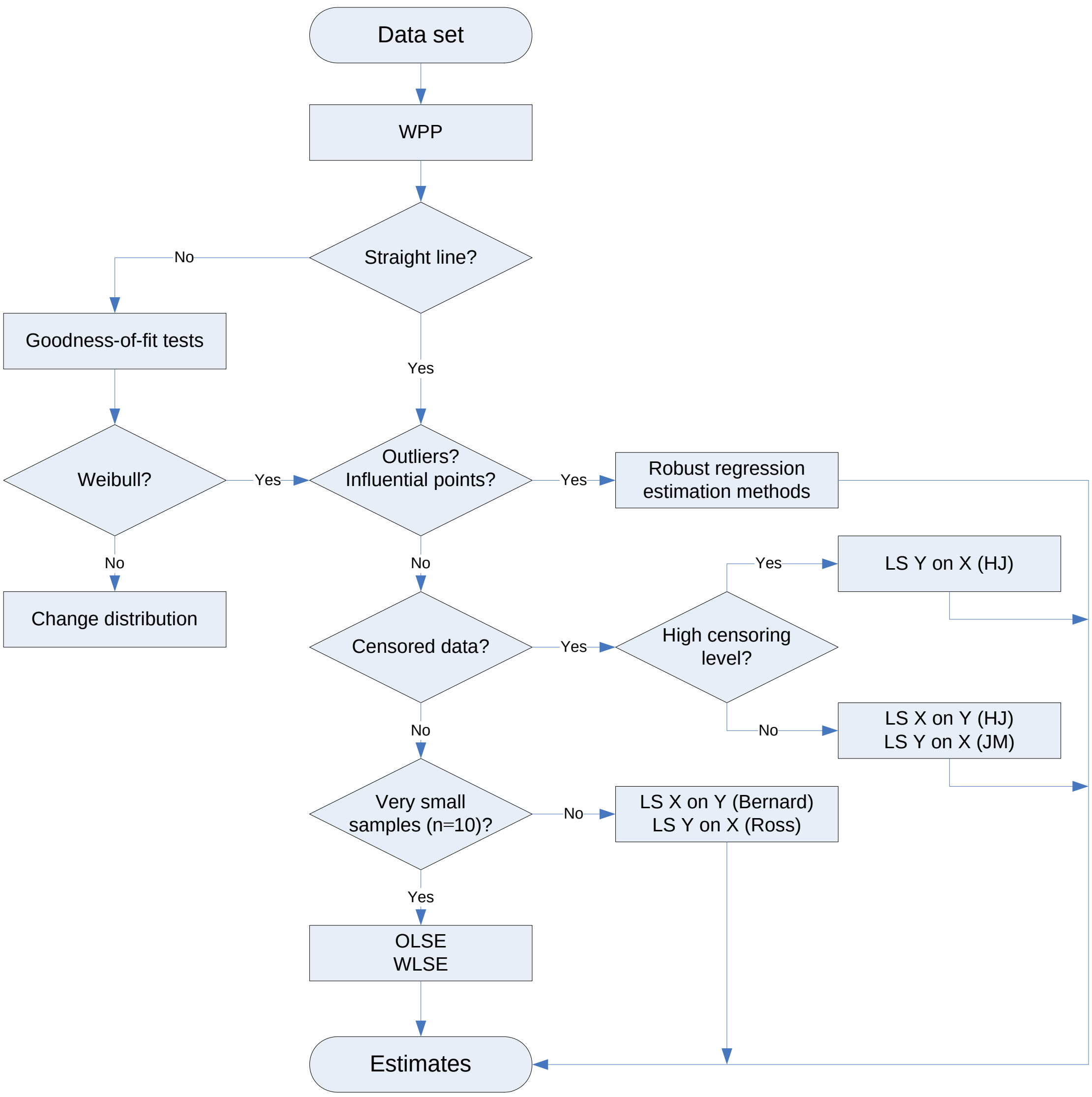


Figure 8-1: Flowchart on the selection of linear regression estimation methods.

8.3 Case Studies

8.3.1 Case Study 1: Life of Compressor (Complete Data)

This case study examines the life of compressors. The source of data is in the work of Moss (2005).

Scenario: Four large, identical, horizontal reciprocating compressors were monitored over a period for piston/liner failures. Since both piston and liner were replaced after failure, the lifetimes observed were treated as a complete sample. For each compressor, the failure times were recorded for five times as shown in Table 8-1.

Table 8-1: Original data of case study 1.

Compressor	*1st failure*	*2nd failure*	*3rd failure*	*4th failure*	*5th failure*
A	3600	3803	630	4001	7010
B	4200	4710	4600	1902	3808
C	2408	3018	1650	4926	2415
D	3003	5405	3609	5909	2806

Analysis: All the failure records are merged to form a complete sample of size 20. As this is a complete sample, the selection of the estimation methods is based on the sample size. According to the flowchart in Figure 8-1, LS X on Y with the Bernard estimator or LS Y on X with the Ross estimator can be used to estimate the parameters. Table 8-2 tabulates the calculation spreadsheet and below the table the calculations of the estimates are presented. OLSE and MLE were also used for this sample and the comparison of estimation results are shown in Table 8-3. Figure 8-2 shows the WPP with the straight line fit by LS X on Y (Bernard).

For this sample, LS Y on X (Ross) and LS X on Y (Bernard) provide similar parameter estimates for the shape parameter, and the WPP shows a good fit.

Table 8-2: Parameter estimation of case 1: the calculation spreadsheet.

i	$t_{(i)}$	$F_{(i),Bernard}$	$F_{(i),Ross}$	x_i	$y_{i,Bernard}$	$y_{i,Ross}$	x_i^2	$y_{i,Bernard}^2$	$x_i \cdot y_{i,Bernard}$	$x_i \cdot y_{i,Ross}$
1	630	0.03	0.03	6.45	-3.35	-3.57	41.55	11.25	-21.62	-23.04
2	1650	0.08	0.08	7.41	-2.44	-2.52	54.89	5.96	-18.09	-18.70
3	1902	0.13	0.13	7.55	-1.95	-2.00	57.01	3.81	-14.74	-15.11
4	2408	0.18	0.18	7.79	-1.61	-1.64	60.63	2.59	-12.53	-12.80
5	2415	0.23	0.23	7.79	-1.34	-1.37	60.68	1.80	-10.44	-10.64
6	2806	0.28	0.27	7.94	-1.12	-1.14	63.04	1.24	-8.86	-9.02
7	3003	0.33	0.32	8.01	-0.92	-0.94	64.12	0.85	-7.37	-7.51
8	3018	0.38	0.37	8.01	-0.75	-0.76	64.20	0.56	-5.98	-6.09
9	3600	0.43	0.42	8.19	-0.59	-0.60	67.05	0.34	-4.81	-4.90
10	3609	0.48	0.47	8.19	-0.44	-0.45	67.10	0.19	-3.59	-3.67
11	3803	0.52	0.52	8.24	-0.30	-0.31	67.96	0.09	-2.44	-2.51
12	3808	0.57	0.57	8.24	-0.16	-0.17	67.98	0.03	-1.32	-1.38
13	4001	0.62	0.62	8.29	-0.03	-0.03	68.80	0.00	-0.22	-0.27
14	4200	0.67	0.67	8.34	0.11	0.10	69.60	0.01	0.90	0.85
15	4600	0.72	0.72	8.43	0.24	0.24	71.13	0.06	2.05	2.01
16	4710	0.77	0.77	8.46	0.38	0.38	71.53	0.15	3.25	3.22
17	4926	0.82	0.82	8.50	0.53	0.53	72.29	0.29	4.55	4.52
18	5405	0.87	0.87	8.60	0.70	0.70	73.88	0.50	6.05	6.04
19	5909	0.92	0.92	8.68	0.91	0.91	75.42	0.83	7.90	7.90
20	7010	0.97	0.97	8.86	1.22	1.22	78.41	1.48	10.76	10.78
sum				161.97	-10.89	-11.41	1317.24	32.02	-76.55	-80.32

Calculation of Estimates by LS Y on X (Ross)

$$\hat{\beta} = \frac{n\sum_{i=1}^{n} x_i y_{i,Ross} - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_{i,Ross}}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{20\times(-76.55) - 161.97\times(-11.41)}{20\times 1317.24 - (161.97)^2} = 2.21$$

$$\hat{\alpha} = \exp\left(-\frac{\sum_{i=1}^{n} y_{i,Ross} - \hat{\beta}\sum_{i=1}^{n} x_i}{\hat{\beta} n} \right) = \exp\left(-\frac{-11.41 - 2.21\times 161.97}{20\times 2.21} \right) = 4257.19$$

Calculation of Estimates by LS X on Y (Bernard)

$$\hat{\beta} = \frac{n\sum_{i=1}^{n} y_{i,Bernard}^2 - \left(\sum_{i=1}^{n} y_{i,Bernard}\right)^2}{n\sum_{i-1}^{n} x_i y_{i,Bernard} - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_{i,Bernard}} = \frac{20\times 32.02 - (-10.89)^2}{20\times(-76.55) - 161.97\times(-10.89)} = 2.24$$

$$\hat{\alpha} = \exp\left(\frac{\hat{\beta}\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_{i,Bernard}}{\hat{\beta} n} \right) = \exp\left(\frac{20 \times 161.97 + 10.89}{20 \times 2.24} \right) = 4194.85$$

Table 8-3: Comparison results of different estimation methods (case study 1).

	OLSE	LS X on Y (Bernard)	LS Y on X (Ross)	MLE
$\hat{\alpha}$	4248.33	4257.19	4194.85	4121.75
$\hat{\beta}$	2.13	2.21	2.24	2.64

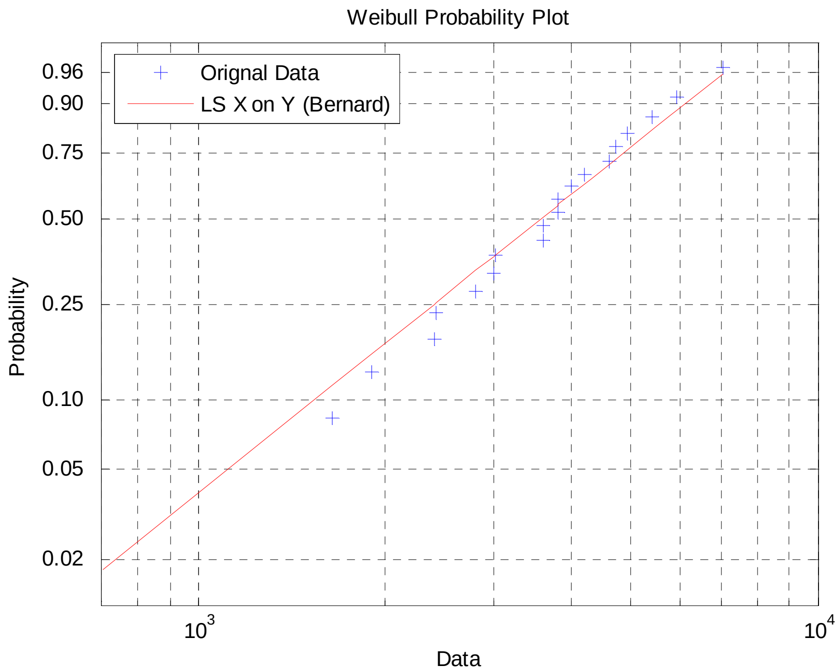


Figure 8-2: WPP of case 1. The straight line is fit by the LS X on Y (Bernard) method.

8.3.2 Case Study 2: Life of Capacitor (Multiply Censored Data with a Low Censoring Level)

The scenario of this example is described in Tobias & Trindade (1995) without providing observations. The experiment was slightly modified and a data set which is Weibull distributed with $\alpha_T = 1000$ and $\beta_T = 1$ was randomly generated.

Scenario: An experiment was carried out to test capacitors on fixtures mounted in ovens. Assume the test started with 20 capacitors in four ovens, each containing 5 units. The units were subject to a fixed high voltage and high temperature. All units are expected to be tested to failure, however, at 250hr, the experimenter found one of the ovens malfunctions, causing all further data in this oven invalid. The other ovens and units continued till all of them failed.

Analysis: The experiment output is a multiply censored data set with a censoring level $c = 25\%$ (5 censors in 20 observations). Since the censoring level is low, according to Figure 8-1, LS *Y* on *X* with the JM estimator or LS *X* on *Y* with the HJ estimator is preferred to estimate the parameters. Table 8-4 tabulates the the calculation spreadsheet and below the table the calculations of the estimates are presented. Table 8-5 shows the estimation results from five methods including LS *Y* on *X* with the HJ estimator and the JM estimator, LS *X* on *Y* with the HJ estimator and the JM estimator, and MLE. As can be seen, LS *X* on *Y* with the JM estimator provides very accurate estimate for the shape parameter for this sample. Figure 8-3 shows the WPP with the straight line fit by this method. MLE tends to overestimate β while HJ tends to underestimate β. This can be dangerous because $\beta > 1$, $\beta = 1$ and $\beta < 1$ represent different failure modes.

Table 8-4: Parameter estimation of case 2: the calculation spreadsheet.

i	Failure/ Censor Index	$t_{(i)}$	$m_{f,(j),JM}$	$F_{(j),JM}$	$F_{(j),HJ}$	x_i	$y_{i,JM}$	$y_{i,HJ}$	x_i^2	$y_{i,HJ}^2$	$x_i y_{i,JM}$	$x_i y_{i,HJ}$
1	F	62.29	1.00	0.03	0.05	4.13	-3.35	-3.02	17.07	9.12	-13.86	-12.48
2	F	75.07	2.00	0.08	0.10	4.32	-2.44	-2.30	18.65	5.30	-10.54	-9.94
3	F	104.99	3.00	0.13	0.14	4.65	-1.95	-1.87	21.66	3.50	-9.08	-8.70
4	F	184.73	4.00	0.18	0.19	5.22	-1.61	-1.55	27.24	2.42	-8.40	-8.11
5	F	185.49	5.00	0.23	0.24	5.22	-1.34	-1.30	27.28	1.70	-7.00	-6.80
6	F	209.76	6.00	0.28	0.29	5.35	-1.12	-1.09	28.58	1.19	-5.96	-5.82
7	F	219.22	7.00	0.33	0.33	5.39	-0.92	-0.90	29.05	0.81	-4.96	-4.87
8	F	225.13	8.00	0.38	0.38	5.42	-0.75	-0.73	29.34	0.54	-4.04	-3.98
9	C	250.00										
10	C	250.00										
11	C	250.00										
12	C	250.00										
13	C	250.00										
14	F	999.95	9.63	0.46	0.46	6.91	-0.49	-0.49	47.72	0.24	-3.40	-3.38
15	F	1126.22	11.25	0.54	0.54	7.03	-0.26	-0.26	49.37	0.07	-1.84	-1.86
16	F	1398.03	12.88	0.62	0.61	7.24	-0.04	-0.05	52.46	0.00	-0.31	-0.37
17	F	1528.17	14.50	0.70	0.69	7.33	0.17	0.16	53.76	0.03	1.28	1.17
18	F	1708.08	16.13	0.78	0.77	7.44	0.40	0.38	55.40	0.14	2.99	2.82
19	F	1741.19	17.75	0.86	0.85	7.46	0.66	0.62	55.69	0.39	4.92	4.65
20	F	1897.15	19.38	0.94	0.92	7.55	1.01	0.94	56.97	0.88	7.59	7.09
sum						90.66	-12.04	-11.48	570.23	26.32	-52.63	-50.59

Calculation of Estimates by LS Y on X (JM)

$$\begin{cases} \hat{\beta} = \dfrac{r\sum_{i=1}^{r} x_i y_{i,JM} - \sum_{i=1}^{r} x_i \sum_{i=1}^{r} y_{i,JM}}{r\sum_{i-1}^{r} x_i^2 - \left(\sum_{i=1}^{r} x_i\right)^2} = \dfrac{15\times(-52.63) - 90.66\times(-12.04)}{15\times 570.23 - 90.66^2} = 0.90 \\ \hat{\alpha} = \exp\left(-\dfrac{\sum_{i=1}^{r} y_{i,JM} - \hat{\beta}\sum_{i=1}^{r} x_i}{\hat{\beta} r} \right) = \exp\left(-\dfrac{-12.04 - 0.90\times 90.66}{15\times 0.90} \right) = 1028.50 \end{cases}$$

Calculation of Estimates by LS X on Y (HJ)

$$\begin{cases} \hat{\beta} = \dfrac{r\sum_{i=1}^{r} y_{i,HJ}^2 - \left(\sum_{i=1}^{r} y_{i,HJ}\right)^2}{r\sum_{i-1}^{r} x_i y_{i,HJ} - \sum_{i=1}^{r} x_i \sum_{i=1}^{r} y_{i,HJ}} = \dfrac{15\times 26.32 - (-11.48)^2}{15\times(-50.59) - 90.66\times(-11.48)} = 0.93 \\ \hat{\alpha} = \exp\left(\dfrac{\hat{\beta}\sum_{i=1}^{r} x_i - \sum_{i=1}^{r} y_{i,HJ}}{\hat{\beta} r} \right) = \exp\left(\dfrac{0.93\times 90.66 - (-11.48)}{15\times 0.93} \right) = 960.00 \end{cases}$$

Table 8-5: Comparison results of different estimation methods (case study 2).

	LS Y on X (HJ)	LS X on Y (HJ)	LS Y on X (JM)	LS X on Y (JM)	MLE
$\hat{\alpha}$	1044.56	960.07	1028.29	932.97	915.78
$\hat{\beta}$	0.84	0.93	0.90	1.01	1.07

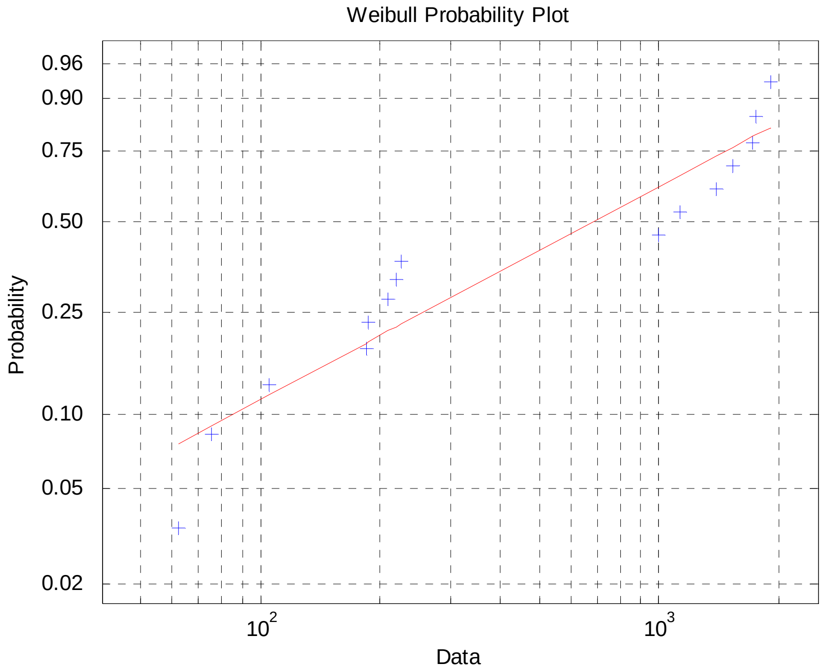


Figure 8-3: WPP of case 2. The straight line is fit by LS X on Y with the JM estimator.

8.3.3 Case Study 3: Life of Radio (Type II Censored Data with a High Censoring Level)

The source of this case study comes from the work of Lawson et al. (1997).

Scenario: 20 radios were placed in an environment test chamber and tested until 8 radios failed. Cycles-to-failure data were collected. Based on similar product history, the distribution is assumed to be Weibull.

Analysis: This data set is a singly Type II censored data set. The censoring level is 60%, which is a high censoring level. For censored data of high censoring levels, the flowchart in Figure 8-1 suggests the method of LS Y on X (HJ), i.e., OLSE, be used. Table 8-6 tabulates the calculation spreadsheet and below the table the calculations of the estimates are presented. The calculation for the estimates of LS Y on X (JM) is also presented. Figure 8-4 shows the WPP for both methods.

The estimates of the two methods are close. Since we do not know the true parameter values, it is hard to judge which one is better. On the other hand, from the WPP, it can be observed that the first data point is suspicious to be an outlier, indicating the robust regression estimation methods should be used.

Applying robust M-estimation (bisquare) to this data set, the estimation results are: $\hat{\alpha} = 1284.21, \hat{\beta} = 1.33$. Figure 8-5 shows the WPP with straight lines fit by LS Y on X (HJ) or OLSE and the robust M-estimation (bisquare). It can be seen from the figure that the robust regression line is less affected by the first data point.

Table 8-6: Parameter estimation of case 3: the calculation spreadsheet.

i	$t_{(i)}$	$F_{(j),JM}$	$F_{(j),HJ}$	x_i	$y_{i,JM}$	$y_{i,HJ}$	x_i^2	$y_{i,JM}^2$	$y_{i,HJ}^2$	$x_i y_{i,JM}$	$x_i y_{i,HJ}$
1	260	0.03	0.05	5.56	-3.35	-3.02	30.92	11.25	9.12	-18.65	-16.79
2	265	0.08	0.10	5.58	-2.44	-2.30	31.13	5.96	5.30	-13.62	-12.84
3	300	0.13	0.14	5.70	-1.95	-1.87	32.53	3.81	3.50	-11.13	-10.67
4	305	0.18	0.19	5.72	-1.61	-1.55	32.72	2.59	2.42	-9.20	-8.89
5	425	0.23	0.24	6.05	-1.34	-1.30	36.63	1.80	1.70	-8.11	-7.88
6	545	0.28	0.29	6.30	-1.12	-1.09	39.70	1.24	1.19	-7.03	-6.86
7	620	0.33	0.33	6.43	-0.92	-0.90	41.34	0.85	0.81	-5.92	-5.80
8	870	0.38	0.38	6.77	-0.75	-0.73	45.81	0.56	0.54	-5.05	-4.97
sum				48.12	-13.48	-12.78	290.79	28.06	24.57	-78.73	-74.72

Calculation of Estimates by LS Y on X (HJ)

$$\hat{\beta} = \frac{r\sum_{i-1}^{r} x_i y_{i,HJ} - \sum_{i=1}^{r} x_i \sum_{i=1}^{r} y_{i,HJ}}{r\sum_{i=1}^{r} x_i^2 - \left(\sum_{i=1}^{r} x_i\right)^2} = \frac{8\times(-74.72) - 48.12\times(-12.78)}{8\times 290.79 - (48.12)^2} = 1.60$$

$$\hat{\alpha} = \exp\left(-\frac{\sum_{i=1}^{r} y_{i,HJ} - \hat{\beta}\sum_{i=1}^{r} x_i}{\hat{\beta} r} \right) = \exp\left(-\frac{-12.78 - 1.60 \times 48.12}{8 \times 1.60} \right) = 1111.47$$

Calculation of Estimates by LS Y on X (JM)

$$\hat{\beta} = \frac{r\sum_{i-1}^{r} x_i y_{i,JM} - \sum_{i=1}^{r} x_i \sum_{i=1}^{r} y_{i,JM}}{r\sum_{i=1}^{r} x_i^2 - \left(\sum_{i=1}^{r} x_i\right)^2} = \frac{8 \times (-78.73) - 48.12 \times (-13.48)}{8 \times 290.79 - (48.12)^2} = 1.74$$

$$\hat{\alpha} = \exp\left(-\frac{\sum_{i=1}^{r} y_{i,JM} - \hat{\beta}\sum_{i=1}^{r} x_i}{\hat{\beta} r} \right) = \exp\left(-\frac{-13.48 - 1.74 \times 48.12}{8 \times 1.74} \right) = 1078.57$$

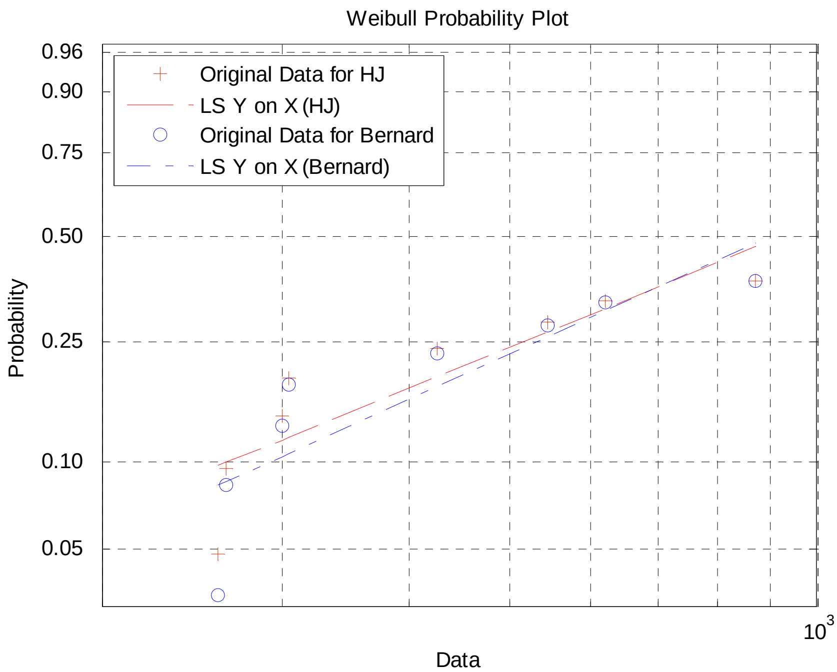


Figure 8-4: WPP of case 3. The straight lines are fit by LS *Y* on *X* (JM) and LS *Y* on *X* (HJ).

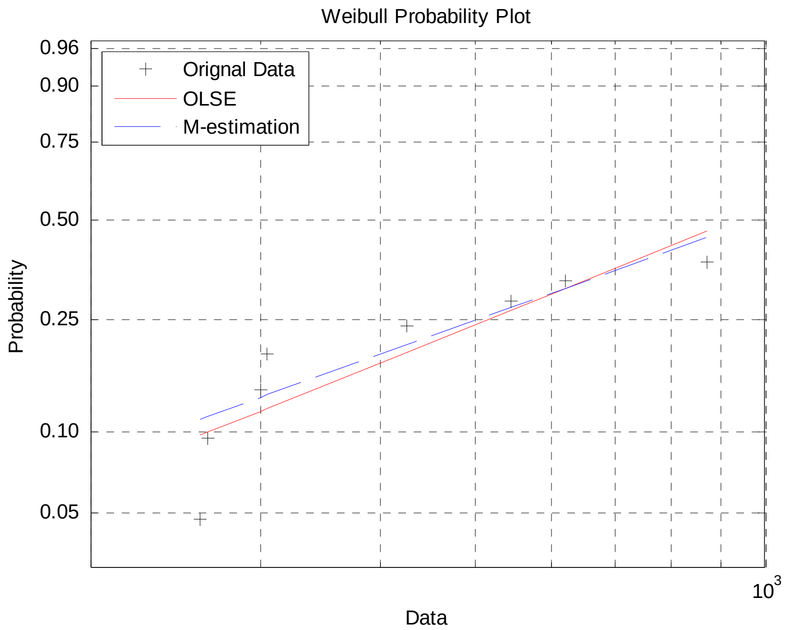


Figure 8-5: WPP of case 3. The straight lines are fit by OLSE and M-estimation (bisquare).

Chapter 9

Conclusions and Future Work

9.1 Conclusions

This thesis explored a group of linear regression estimation methods for the Weibull distribution. LSE is the basic method in this group which is traditionally considered simple but inaccurate. The LSE method in the general sense has the flexibility on the selection of failure probability estimators and the regression direction. We defined the OLSE method which uses the most widely used failure probability estimators (i.e., the Bernard estimator for complete data and the HJ estimator for censored data), and the regression direction of *Y* on *X*. Due to the simplicity, the OLSE method is widely used by Weibull practitioners. On contrary, it has been less discussed by researchers compared to other analytical estimation methods such as MLE.

The statistical properties of the OLS estimators of the Weibull scale and shape parameters were carefully studied via both theoretical analyses and Monte Carlo simulation experiments. In the theoretical analyses, firstly, we showed that the parameter estimators of OLSE are not BLUE given that the variance of errors cannot be constant and the covariance of errors is correlated. Secondly, assuming the *Y*-axis plotting positions are pre-determined and can be treated as fixed values, we deduced the analytical expressions of the bias of the OLS estimators as a function of the *Y*-axis plotting positions. Thirdly, we proved that $\hat{\beta}/\beta$ and $\hat{\beta}\ln(\hat{\alpha}/\alpha)$, whose distributions are independent of α and β, of the LS estimators are two pivotal functions. This applies to both complete data and censored data. The first pivotal function $\hat{\beta}/\beta$ is the theoretical foundation of the proposed bias correction methods (Chapter 5). In addition, we pointed out that the two pivotal functions have great impact on the Monte Carlo experiments described throughout this thesis. First of all, the functions can be used to check the reliability of the simulation results. Second, the functions provide theoretical support for simplifying the setting of the true parameter values of α_T and β_T in the simulation experiment and hence save much effort in the simulation. Since it is difficult to identify the distributions of the estimators of OLSE or other linear regression estimation methods via analytical approaches, the Monte Carlo method was used frequently to study the properties of the estimators. The simulation results for the OLSE of the shape and scale parameters for complete data and multiply censored data at different sample sizes and censoring levels were tabulated. We found that for complete data, the OLSE of the shape parameter is inconsistent with sample size n and the bias reaches smallest at $n = 6-7$. During $10 < n < 30$, the bias keeps around 4%. For multiply censored data, the bias of the OLS shape parameter estimator is inconsistent with the censoring level c and reaches smallest at different combinations of n and c, e.g., $c = 30\%$ and $n = 150-200$, $c = 40\%$ and $n = 100-150$, $c = 50\%$ and $n = 80-100$, $c = 60\%$ and $n = 50-60$, and $c = 70\%$ and $n = 20-30$. For estimating α, the results are unsatisfactory when $\beta_T \leq 1$, but the bias generally decreases as β_T increases. We also found that the magnitude of the standard deviation of both estimators of OLSE is much larger than the magnitude of the bias in most cases, indicating that improving the efficiency of OLSE is important.

Some arguments were made on the procedure of the OLSE method. A frequently discussed issue toward LSE among Weibull researchers is the estimation of failure probability, also known as the determination of *Y*-axis plotting positions. We summarized the existing estimators of F for complete data and censored data, respectively, and divided them into different categories. Two tables were provided for easy references. These estimators were compared in terms of several aspects including the theoretical foundation and the application simplicity. Then, the properties of the LS estimators with different estimators of F used in the regression, were examined via the Monte Carlo simulation experiment. We focused on those relatively new estimators of F proposed in the last decade including the Ross estimator (Ross, 1994), the Drap-Kos estimator (Drapella & Kosznik, 1999), the age sensitive estimator (Hastings and Bartlett, 1997) and the RRR estimator (Wang, 2001, 2004). The simulation results showed that for complete data, the Bernard estimator outperforms the Ross estimator or the Drap-Kos estimator for estimating the shape parameter only when $n < 10$ in view of the bias. The Ross estimator or the Drap-Kos estimator can generate nearly unbiased $\hat{\beta}$ when $n \geq 10$. However, we also found that the Ross estimator or the Drap-Kos estimator cannot improve the estimation efficiency. The simulation results for censored data showed that JM, ASM and RRR are good for samples with low censoring levels, say $c < 50\%$, and the three methods perform similarly. For application simplicity, JM should be used. The simplest method HJ was found to perform best for samples with high censoring levels, say $c \geq 50\%$.

Another argument of the OLSE procedure is the determination of the independent and dependent variables when conducting least squares regression. OLSE treats $X = \ln T$ as independent variable and $Y = \ln[-\ln(1-F)]$ as dependent variable which is consistent with WPP where the *X*-axis is t and the *Y*-axis is F. However, we noticed that in the early literature (see, e.g. Weibull, 1967; White, 1969 and Mann et al., 1974) such a setting is reversed. The two methods are named LS *Y* on *X* and LS *X* on *Y*. We compared them in terms of model statistics and parameter estimation. As is known, a model comparison and a parameter estimation comparison are two different things for Weibull parameter estimation methods. We proved that the two regression models of LS *Y* on *X* and LS *X* on *Y* have same R^2 and the ratio of their MS_{Error} equals to β^2. Thus LS *X* on *Y* has a smaller MS_{error} when $\beta > 1$ and LS *Y* on *X* has a smaller MS_{error} when $\beta < 1$. This provides a rule for model selection between the two when we have information about the value of β. For parameter estimation, our simulation results showed that for complete samples, LS *Y* on *X* is recommended for estimating β for very small samples, say $n \le 10$, and LS *X* on *Y* is recommended for estimating β for medium to large samples, say $n \ge 30$. For censored samples, LS *Y* on *X* is recommended for estimating β for samples with high censoring levels ($c \ge 50\%$), and LS *X* on *Y* is recommended for estimating β for samples with low censoring levels ($c < 50\%$). For estimating α, LS *X* on *Y* is recommended for both complete and censored samples.

In view of the bias of the OLSE of the shape parameter, we proposed several simple bias correcting formulas which can be used in the end of the OLSE procedure. The bias correcting formulas were determined based on the modeling of the unbiasing factors. In the case of complete data, the modified Ross' method and the modified Hirose's method were proposed. The simulation results showed that the proposed methods reduce bias to less than 1% and typically less than 0.5%. The bias correction for the OLSE of the shape parameter was also examined for multiply censored data.

We found that due to the inconsistency of the OLS shape parameter estimator, it is difficult to propose a general model of the bias as a function of the sample size and censoring level. However, when the censoring level is low ($c < 50\%$) and the sample size is within 100, the bias as a function of the sample size and censoring level shows good consistency. Therefore, a simple bias correcting formula was proposed that can be applied to multiply censored samples with $c \le 40\%$ and $n \le 100$. The bias is greatly reduced with the proposed formula.

Besides LSE, the family of linear regression estimation methods also includes WLSE and RRE methods. WLSE methods have been studied by some researchers and a few weight formulas have been proposed. We proposed a novel formula for calculating weights applied to small, complete samples. This formula gives the approximated values of the best weights. Theoretically, the proposed formula is more accurate than the existing ones because it is based on the analytical deduction of the exact values of the variances of predictor variable values. The proposed WLSE method was compared with selected WLSE methods in the literature and OLSE for estimating the Weibull parameters via Monte Carlo simulations. The results showed that it is slightly better than the others and significantly better than OLSE in terms of the standard deviation and MSE of the estimators. Given that the shape parameter estimator of the proposed WLSE method still has a large bias, a simple bias correcting formula was proposed which can be used as an add-on. We also discussed WLSE for large samples and censored samples. The proposed formula for weights cannot be used for large samples and approximation methods have to be used. For censored samples, we suggested to calculate weights by the MFON of each failure data point. The step-by-step procedures of the proposed WLSE method applied to censored data were provided and we also presented a numerical example to illustrate the calculation process.

Robust regression techniques are known to be good at dealing with outliers. As a preliminary study, we mainly examined robust M-estimation methods (with bounded influence functions). We pointed out the special outlier data configuration of the Weibull samples, that is, there should be no *Y*-outliers and *X&Y*-outliers because the plotting positions along the *Y*-axis in WPP are independent of failure times and can be treated as known constants. This makes it unnecessary to examine some of the robust regression techniques that are robust only to the *Y*-outliers. With Monte Carlo simulations, we examined robust M-estimators with different weight functions (bisquare, Andrews, Cauchy and Welsch) on parameter estimation for complete data with one left tail *X*-outlier, one right tail *X*-outlier, two left tail *X*-outliers and two right tail *X*-outliers. We also examined robust M-estimators for multiply censored data. The results of our study indicated that the robust M-estimator of the Weibull shape parameter is more efficient than the OLS estimator for small, complete samples with one *X*-outlier in the left or right tail, and especially when $\beta_T > 1$ and $n \geq 8$. For small complete samples with two *X*-outliers in the tail, the M-estimator still outperforms the OLS estimator. For multiply censored samples, M-estimator performs better than OLSE in most cases in view of both bias and MSE and thus should also be recommended for use.

In the beginning of the thesis we have pointed out that reliability data analysis requires different estimation methods for different types of data. We provided a flowchart to instruct the use of the linear regression estimation methods discussed in this thesis for different types of data. And we used some cases studies to illustrate the process. For all the methods discussed in this thesis, step-by-step procedures were provided so that these methods can be easily applied by engineers and practitioners conducting Weibull analysis. The proposed methods are of great practical value, but there are some assumptions which need to be checked and some problems may be encountered in the future.

9.2 Suggestions for Future Work

An underlying assumption in this study is that the data is known to be from a two-parameter Weibull distribution, or it can at least be best modelled by such a distribution. This assumption can be roughly checked by WPP. If the data points form the approximation of a straight line on WPP, we can say that the assumption is satisfied. However, elaborate statistical tests may be necessary to confirm this assumption.

A large portion of the results in this thesis was obtained via Monte Carlo simulations. We selected only limited values for the experiment factors including α, β, n and c. Moreover, due to the focus on small samples, n was mainly set to within 30. Large sample properties of the proposed methods were not carefully examined, though we have noticed that the OLSE of the shape parameters is inconsistent with the sample size.

During the presentation of the WLSE methods, a tentative method for calculating weights applied to multiply censored data was proposed. Future work could be conducted to further investigate this procedure both theoretically and via Monte Carlo simulations. In the proposed procedure, the JM estimator was recommended for calculating the MFON; however, it would be nice to check other estimators as well.

The study of the RRE methods is just a beginning. Besides the robust M-estimation methods, other robust regression methods could be examined in future work. In addition, we considered only tail area outliers in our experiment due to its popularity; however, outliers can occur in other places in a sample. Future work could be conducted to examine such conditions.

The shape parameter estimators are the focus of this study and we assumed the shape parameter is more important than the scale parameter. There are circumstances that people have knowledge about the shape parameter and the scale parameter is of more concern. We have found that OLSE can perform badly for estimating the scale parameter when the shape parameter is small (within 1). Therefore, future work could focus on the scale parameter, e.g., to propose bias correction methods for the scale parameter.

Finally, the linear regression estimation methods could be extended to other distributions in the Weibull family such as the three-parameter Weibull distribution, the extended Weibull distributions and modified Weibull distributions. The WLSE methods, RRE methods and bias correction methods could be proposed for these distributions to generate more accurate parameter estimates or improve the estimation efficiency.

Bibliography

Abbasi, B. et al., 2006. Estimating the Parameters of Weibull Distribution Using Simulated Annealing Algorithm. *Applied Mathematics and Computation*, 183(1), 85-93.

Abernethy, R.B., 2000. *The New Weibull Handbook (Fourth Edition)*, FL: Robert B. Abernethy.

Allen, M.P., 1997. *Understanding Regression Analysis*, NY: Plenum Press.

Ambrozic, M. & Vidovic, K., 2007. Reliability of the Weibull Analysis of the Strength of Construction Materials. *Journal of Materials Science*, 42(23), 9645-9653.

Anderson, C. & Schumacker, R.E., 2003. A Comparison of Five Robust Regression Methods with Ordinary Least Squares Regression: Relative Efficiency, Bias, and Test of Null Hypothesis. *Understanding Statistics*, 2(2), 79-103.

Andrews, D.F. et al., 1972. *Robust Estimates of Location: Survey and Advances*, NJ: Princeton University Press.

Arora, M.S., 2000. *Theory and Applications of the 4-Parameter Generalized Weibull Distribution (Ph.D. Thesis)*, Ann Arbor, Mich.: University Microfilms International.

Bain, L.J., 1972. Inferences Based on Censored Sampling from the Weibull or Extreme-Value Distribution. *Technometrics*, 14(3), 693-702.

Bain, L.J. & Antle, C.E., 1967. Estimation of Parameters in the Weibull Distribution. *Technometrics*, 9(4), 621-627.

Beaton, A.E. & Tukey, J.W., 1974. The Fitting of Power Series, Meaning Polynomials, Illustrated on Band-Spectroscopic Data. *Technometrics*, 16(2), 147-185.

Bebbington, M., Lai, C.D. & Zitikis, R., 2007. A Flexible Weibull Distribution. *Reliability Engineering and System Safety*, 92(6), 719-726.

Bebbington, M., Lai, C.D. & Zitikis, R., 2008. Estimating the Turning Point of a Bathtub-Shaped Failure Distribution. *Journal of Statistical Planning and Inference*, 138(4), 1157-1166.

Berger, R.W. & Lawrence, K., 1974. Estimating Weibull Parameters by Linear and

Non Linear Regression. *Technometrics*, 16(4), 617-619.

Bergman, B., 1986. Estimation of Weibull Parameters Using a Weight Function. *Journal of Materials Science Letters*, 5, 611–614.

Bernard, A. & Bosi-Levenbach, E.C., 1953. The Plotting of Observations on Probability Paper. *Statistica Neerlandica*, 7, 163-173.

Billmann, B.R., Antle, C.E. & Bain, L.J., 1972. Statistical Inference from Censored Weibull Samples. *Technometrics*, 14(4), 831-840.

Blom, G., 1958. *Statistical Estimates and Transformed Beta Variables*, NY: Wiley.

Breyfogle, F.W., 1992. *Statistical Methods for Testing, Development, and Manufacturing*, NY: Wiley.

Cacciari, M., Mazzanti, G. & Montanari, G., 1996. Comparison of Maximum Likelihood Unbiasing Methods for the Estimation of the Weibull Parameters. *IEEE Transactions on Dielectrics and Electrical Insulation*, 3(1), 18-27.

Cacciari, M. & Montanari, G.C., 1991. Discussion: Estimating the Cumulative Probability of Failure Data Points to be Plotted on Weibull and Other Probability Paper. *IEEE Transactions on Electrical Insulation*, 26(6), 1224-1229.

Campean, I.F., 2000. Exponential Age Sensitive Method for Failure Rank Estimation. *Quality and Reliability Engineering International*, 16(4), 291-300.

Cohen, A.C., 1965. Maximum Likelihood Estimation in the Weibull Distribution Based on Complete and on Censored Samples. *Technometrics*, 7(4), 579-588.

Cran, G.W., 1976. Graphical Estimation Methods for Weibull Distributions. *Microelectronics Reliability*, 15(1), 47-52.

D’Agostino, R.B., 1971. Linear Estimation of the Weibull Parameters. *Technometrics*, 13(1), 171-182.

Dimitrakopoulou, T., Adamidis, K. & Loukas, S., 2007. A Lifetime Distribution with an Upside-Down Bathtub-Shaped Hazard Function. *IEEE Transactions on Reliability*, 56(2), 308-311.

Dodson, B., 2006. *The Weibull Analysis Handbook*, Milwaukee: American Society for Quality.

Drapella, A. & Kosznik, S., 1999. An Alternative Rule for Placement of Empirical Points on Weibull Probability Paper. *Quality and Reliability Engineering*

International, 15(1), 57-59.

Dubey, S.D., 1966. *Asymptotic Efficiencies of the Moment Estimators for the Parameters of the Weibull Laws*, IS: Naval Research Logistics Quarterly.

Dubey, S.D., 1965. Asymptotic Properties of Several Estimators of Weibull Parameters. *Technometrics*, 7(3), 423-434.

Engelhardt, M., 1975. On Simple Estimation of the Parameters of the Weibull or Extreme-Value Distribution. *Technometrics*, 17(3), 369-374.

Engelhardt, M. & Bain, L.J., 1977. Simplified Statistical Procedures for the Weibull or Extreme-value Distribution. *Technometrics*, 19(3), 323-331.

Engelhardt, M. & Bain, L.J., 1973. Some Complete and Censored Sampling Results for the Weibull or Extreme-Value Distribution. *Technometrics*, 15(3), 541-549.

Engelhardt, M. & Bain, L.J., 1974. Some Results on Point Estimation for the Two-Parameter Weibull or Extreme-Value Distribution. *Technometrics*, 16(1), 49-56.

Farnum, N. & Booth, P., 1997. Uniqueness of Maximum Likelihood Estimators of the 2-parameter Weibull Distribution. *IEEE Transactions on Reliability*, 46(4), 523–525.

Faucher, B. & Tyson, W.R., 1988. On the Determination of Weibull Parameters. *Journal of Materials Science Letters*, 7, 1199–1203.

Ferrari, S.L.P., Silva, M.F. & Cribari-Neto, F., 2007. Adjusted Profile Likelihoods for the Weibull Shape Parameter. *Journal of Statistical Computation and Simulation*, 77(7), 531-548.

Filliben, J., 1975. The Probability Plot Correlation Coefficient Test for Normality. *Technometrics*, 17(1), 111-117.

Fothergill, J.C., 1990. Estimating the Cumulative Probability of Failure Data Points to be Plotted on Weibull and Other Probability Paper. *IEEE Transactions on Electrical Insulation*, 25(3), 489-492.

Garthwaite, P.H., Jolliffe, I. & Jones, B., 2002. *Statistical Inference*, NY: Oxford University Press.

Gibbons, J.D., Olkin, I. & Sobel, M., 1999. *Selecting and Ordering Populations: A New Statistical Methodology*, Society for Industrial and Applied Mathematics.

Hampel, F.R. et al., 1986. *Robust Statistics: The Approach Based on Influence Functions*, NY: Wiley.

Harter, H.L. & Moore, A.H., 1967. Asymptotic Variances and Covariances of Maximum-Likelihood Estimators, from Censored Samples, of the Parameters of Weibull and Gamma Populations. *Annals of Mathematical Statistics*, 38(2), 557-570.

Harter, H.L. & Moore, A.H., 1965. Maximum-likelihood Estimation of the Parameters of Gamma and Weibull Populations from Complete and from Censored Samples. *Technometrics*, 7(4), 639-643.

Hastings, N.A.J. & Bartlett, H.J.G., 1997. Estimating the Failure Order-Number from Reliability Data with Suspended Items. *IEEE Transactions on Reliability*, 46(2), 266-268.

Hazen, A., 1930. *Flood Flows*, NY: Wiley.

Herd, G.R., 1960. Estimation of Reliability from Incomplete Data. *Proceedings of the 6th International Symposium on Reliability and Quality Control*.

Hirose, H., 1999. Bias Correction for the Maximum Likelihood Estimates in the Two Parameter Weibull Distribution. *IEEE Transactions on Dielectrics and Electrical Insulation*, 6(1), 66-68.

Hirose, H., 1996. Maximum Likelihood Estimation in the 3-parameter Weibull Distribution. *IEEE Transactions on Dielectrics and Electrical Insulation*, 3(1), 43-55.

Holland, P.W. & Welsch, R.E., 1977. Robust Regression Using Iteratively Reweighted Least-Sqaures. *Communications in Statistics: Theory and Methods*, 6, 813-827.

Hossain, A.M. & Zimmer, W.J., 2003. Comparisons of Estimation Methods for Weibull Parameters: Complete and Censored Samples. *Journal of Statistical Computation and Simulation*, 73(2), 145-153.

Huber, P.J., 1973. Robust Regression: Asymptotics, Conjectures, and Monte Carlo. *The Annals of Statistics*, 1(5), 799-821.

Hung, W.L., 2004. Estimation Of Weibull Parameters Using a Fuzzy Least-Squares Method. *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems*, 12(5), 701-711.

Hung, W.L., 2001. Weighted Least Squares Estimation of the Shape Parameter of the Weibull Distribution. *Quality and Reliability Engineering International*, 17(6), 467-469.

IEEE Guide for the Statistical Analysis of Electrical Insulation Voltage Endurance Data, 1987. *IEEE Guide for the Statistical Analysis of Electrical Insulation Voltage Endurance Data, ANSI/IEEE Std 930-1987*, IEEE.

Jacquelin, J., 1993. Generalization of the Method of Maximum Likelihood. *IEEE Transactions on Dielectrics and Electrical Insulation*, 28(1), 65-72.

Johnson, L.G., 1964. *The Statistical Treatment of Fatigue Experiments*, NY: Elsevier.

Kaminskiy, M.P. & Krivtsov, W., 2005. A Simple Procedure for Bayesian Estimation of the Weibull Distribution. *IEEE Transactions on Reliability*, 54(4), 612-616.

Kao, J.H.K., 1959. A Graphical Estimation of Mixed Weibull Parameters in Life-testing of Electron Tubes. *Technometrics*, 1(4), 389-407.

Kaplan, E.L. & Meier, P., 1958. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, 53(282), 457-481.

Keats, J.B., Nahar, P.C. & Korbel, K.M., 2000. A Study of the Effects of Mis-specification of the Weibull Shape Parameter on Confidence Bounds Based on the Weibull-to-Exponential Transformation. *Quality and Reliability Engineering International*, 16(1), 27-31.

Kimball, B.F., 1960. On the Choice of Plotting Positions of Probability Paper. *Journal of the American Statistical Association*, 55(291), 546- 560.

Lai, C.D. & Xie, M., 2006. *Stochastic Ageing and Dependence for Reliability*, NY: Springer.

Lawson, C., Keats, J.B. & Montgomery, D.C., 1997. Comparison of Robust and Least-Squares Regression in Computer-Generated Probability Plots. *IEEE Transactions on Reliability*, 46(1), 108-115.

Liu, C.C., 1997. *A Comparison Between the Weibull and Lognormal Models Used to Analyze Reliability Data*, University of Nottingham Doctoral Thesis. Available at: http://www.barringer1.com/wa.htm [Accessed December 8, 2009].

Looney, S.W. & Gulledge, T.R., 1985. Probability Plotting Positions and Goodness of Fit for the Normal Distribution. *The Statistician*, 34(3), 297-303.

Lu, H.L., Chen, C.H. & Wu, J.W., 2004. A Note on Weighted Least-squares Estimation of the Shape Parameter of the Weibull Distribution. *Quality and Reliability Engineering International*, 20(6), 579-586.

Mann, N.R., 1968. Point and Interval Estimation Procedures for the Two-Parameter Weibull and Extreme-Value Distribution. *Technometrics*, 10(2), 231-256.

Mann, N.R., 1967. Tables for Obtaining the Best Linear Invariant Estimates of Parameters of the Weibull Distribution. *Technometrics*, 9(4), 629-645.

Mann, N.R., Schafer, R.E. & Singpurwalla, N.D., 1974. *Methods for Statistical Analysis of Reliability and Life Data*, NY: Wiley.

Mark, N.B., 2005. Estimation of Weibull Parameters from Common Percentiles. *Journal of Applied Statistics*, 32(1), 17-24.

MATLAB, 2004. *Statistics Toolbox for Use with MATLAB, User's Guide Version 5*, Natick, MA: The MathWorks.

McCool, J.I., 1970. Inferences on Weibull Percentiles and Shape Parameter from Maximum Likelihood Estimates. *IEEE Transactions on Reliability*, 19(1), 2-9.

McCool, J.I., 1965. The Construction of Good Linear Unbiased Estimates from the Best Linear Estimates for a Smaller Sample Size. *Technometrics*, 7(4), 543-552.

Mischke, C.R., 1979. A Distribution-Independent Plotting Rule for Ordered Failures. *An ASME publication 79-DET-112, American Society of Mechanical Engineers*. Available at: http://www.barringer1.com/wa.htm [Accessed December 8, 2009].

Montanari, G.C. et al., 1997a. In Search of Convenient Techniques for Reducing Bias in the Estimation of Weibull Parameter for Uncensored Tests. *IEEE Transactions on Dielectrics and Electrical Insulation*, 4(3), 306-313.

Montanari, G.C. et al., 1998. Optimum Estimators for the Weibull Distribution from Censored Test Data: Progressively-censored Tests. *IEEE Transactions on Dielectrics and Electrical Insulation*, 5(2), 157-164.

Montanari, G.C. et al., 1997b. Optimum Estimators for the Weibull Distribution of Censored Data: Singly-censored Tests. *IEEE Transactions on Dielectrics and Electrical Insulation*, 4(4), 462-469.

Montgomery, D.C. & Runger, G.C., 2003. *Applied Statistics and Probability for Engineers*, NY: Wiley.

Moss, T.R., 2005. *The Reliability Data Handbook*, London: Professional Engineering.

Munkholm, L. & Perfect, E., 2005. Brittle Fracture of Soil Aggregates: Weibull

Models and Methods of Parameter Estimation. *Soil Science Society of America Journal*, 69(5), 1565-1571.

Murthy, D.N.P., Xie, M. & Jiang, R., 2004. *Weibull Models*, NY: Wiley.

Nelder, J.A. & Mead, R., 1965. A Simplex Method for Function Minimization. *Computer Journal*, 7, 308-313.

Nelson, W., 1972. Theory and Applications of Hazard Plotting for Censored Failure Data. *Technometrics*, 14(4), 945-966.

Nelson, W. & Thompson, V.C., 1971. Weibull Probability Papers. *Journal of Quality Technology*, 3(2), 45-50.

Nelson, W.B., W., 2004. *Applied Life Data Analysis*, NY: Wiley.

Newby, M.J., 1980. The Properties of Moment Estimators for the Weibull Distribution Based on the Sample Coefficient of Variation. *Technometrics*, 22(2), 187-194.

Pham, H. & Lai, C.D., 2007. On Recent Generalizations of the Weibull Distribution. *IEEE Transactions on Reliability*, 56(3), 454-458.

Ramsay, J.O., 1977. A Comparative Study of Several Robust Estimates of Slope, Intercept, and Scale in Linear Regression. *Journal of the American Statistical Association*, 72(359), 608-615.

Rockette, H., Antle, C. & Klimko, L.A., 1974. Maximum Likelihood Estimation with the Weibull Model. *Journal of the American Statistical Association*, 69(345), 246-249.

Ross, R., 1996. Bias and Standard Deviation Due to Weibull Parameter Estimation for Small Data Sets. *IEEE Transactions on Dielectrics and Electrical Insulation*, 3(1), 28-42.

Ross, R., 1999. Comparing Linear Regression and Maximum Likelihood Methods to Estimate Weibull Distributions on Limited Data Sets: Systematic and Random Errors. In Conference on Electrical Insulation and Dielectric Phenomena. Austin, TX.

Ross, R., 1994a. Formulas to Describe the Bias and Standard Deviation of the ML-estimated Weibull Shape Parameter. *IEEE Transactions on Dielectrics and Electrical Insulation*, 1(2), 247-253.

Ross, R., 1994b. Graphical Methods for Plotting and Evaluating Weibull Distributed Data. In *Proceedings of the 4th International Conference on Properties and*

Applications of Dielectric Materials. Brisbane, Australia.

Rousseeuw, P.J., 1984. Least Median of Squares Regression. *Journal of the American Statistical Association*, 79(388), 871-880.

Ruppert, D. & Carroll, R.J., 1980. Trimmed Least Squares Estimation in the Linear Model. *Journal of the American Statistical Association*, 75(372), 828-838.

Ryan, T.P., 1997. *Modern Regression Methods*, NY: Wiley.

Schwarz, G., 1987. Least-Absolute-Value Prediction Lines in Closed Form. *Journal of the American Statistical Association*, 82(400), 1150-1152.

Seki, T. & Yokoyama, S., 1993. Simple and Robust Estimation of the Weibull Parameters. *Microelectronics Reliability*, 33(1), 45-52.

Skinner, K.R., Keats, J.N. & Zimmer, W.J., 2001. A Comparison of Three Estimators of the Weibull Parameters. *Quality and Reliability Engineering International*, 17(4), 249-256.

Soliman, A.A., Abd Ellah, A.H. & Sultan, K.S., 2006. Comparison of Estimates Using Record Statistics from Weibull Model: Bayesian and Non-Bayesian Approaches. *Computational Statistics and Data Analysis*, 51(3), 2065-2077.

Somboonsavatdee, A., Nair, V.N. & Sen, A., 2007. Graphical Estimators from Probability Plots with Right-Censored Data. *Technometrics*, 49(4), 420-429.

S-PLUS, 2002. *S-PLUS 6 Robust Library User's Guide*, Seattle, WA: Insightful Corporation.

Tang, L.C., 2003. Failure-free Life in Reliability Modeling. *International Journal of Industrial Engineering: Theory Applications and Practice*, 10(4), 332-338.

Thisted, A.T. & Thisted, R.A., 1988. *Elements of Statistical Computing*, Florida: CRC Press.

Thoman, D.R., 1972. Linear Order Statistic Estimation for the Two-Parameter Weibull and Extreme-Value Distributions from Type II Progressively Censored Samples. *Technometrics*, 14(3), 679-691.

Thoman, D.R., Bain, L.J. & Antle, E., 1969. Inferences on the Parameters of the Weibull Distribution. *Technometrics*, 11(3), 445-460.

Tiryakioglu, M. & Hudak, D., 2007. On Estimating Weibull Modulus by the Linear Regression Method. *Journal of Materials Science*, 42(24), 10173-10179.

Tobias, P.A. & Trindade, D.C., 1995. *Applied Reliability (Second Edition)*, NY: Van Nostrand Reinhold.

Vittal, S. & Phillips, R., 2007. Uncertainty Analysis of Weibull Estimators for Interval-Censored Data. In *2007 Annual Reliability and Maintainability Symposium*. Orlando, FL.

Wang, F.K. & Keats, J.B., 1995. Improved Percentile Estimation for the Two-parameter Weibull Distribution. *Microelectronics Reliability*, 35(6), 883-892.

Wang, W.D., 2001. Improving the Quality of Field Data Reliability Analysis. In *Proceedings of Annual Reliability and Maintainability Symposium*. Philadelphia, PA.

Wang, W.D., 2004. Refined Rank Regression Method with Censors. *Quality and Reliability Engineering International*, 20(7), 667-678.

Weibull, W., 1951. A Statistical Distribution Function of Wide Applicability. *Journal of Applied Mechanics*, 18, 293-297.

Weibull, W., 1939. A Statistical Theory of the Strength of Materials. *Ingeniörsvetenskapsakademiens Handlingar Nr 151*. Available at: http://www.barringer1.com/wa.htm [Accessed December 8, 2009].

Weibull, W., 1967. The Order Statistics $y_i = \log(z_i^m)$, Their Properties and Use for Parameter Estimation (z = standardized Weibull variate). *USAF AFML-TR=67-161, report AD-818104*. Available at: http://www.barringer1.com/wa.htm [Accessed December 7, 2006].

White, J.S., 1964. Least Squares Unbiased Censored Linear Estimation for the Log Weibull (Extreme Value) Distribution. *Journal of Industrial Mathematical Society*, 14(1), 21-60.

White, J.S., 1969. The Moments of Log-Weibull Order Statistics. *Technometrics*, 11(2), 65-72.

Wu, D.F. & Lu, G.Z., 2004. Improved Estimation of Weibull Parameters with the Linear Regression Method. *Journal of the American Ceramic Society*, 87(9), 1799-1802.

Wu, D.F., Zhou, J.C. & Li, Y.D., 2006. Unbiased Estimation of Weibull Parameters with the Linear Regression Method. *Journal of the European Ceramic Society*, 26(7), 1099-1105.

Wu, E.Y. & Vollertsen, R.P., 2002a. On the Weibull Shape Factor of Intrinsic Breakdown of Dielectric Films and Its Accurate Experimental Determination - Part I: Theory, Methodology, Experimental Techniques. *IEEE Transactions* *on Electron Devices*, 49(12), 2131-2140.

Wu, E.Y. & Vollertsen, R.P., 2002b. On the Weibull Shape Factor of Intrinsic Breakdown of Dielectric Films and Its Accurate Experimental Determination - Part II: Experimental Results and the Effect of Stress Conditions. *IEEE Transactions on Electron Devices*, 49(12), 2141-2150.

Xie, M., Tang, Y. & Goh, T.N., 2002. A Modified Weibull Extension with Bathtub-shaped Failure Rate Function. *Reliability Engineering & System Safety*, 76(3), 279-285.

Xie, M., Yang, Z.L. & Gaudoin, O., 2000. More on the Mis-specification of the Shape Parameter with Weibull-to-Exponential Transformation. *Quality and Reliability Engineering International*, 16(4), 281-290.

Yang, Z.L. & Xie, M., 2003. Efficient Estimation of the Weibull Shape Parameter Based On a Modified Profile Likelihood. *Journal of Statistical Computation and Simulation*, 73(2), 115-123.

Yohai, V.J., 1987. High Breakdown-Point and High Efficiency Estimates for Regression. *Annals of Statistics*, 15(2), 642-665.

Yu, G.H. & Hung, C.C., 2001. A Distribution Free Plotting Position. *Stochastic Environmental Research and Risk Assessment*, 15, 462-476.

Zhang, T.L. & Xie, M., 2007. Failure Data Analysis with Extended Weibull Distribution. *Communications in Statistics-Simulation and Computation*, 36(3), 579-592.

Publications

Zhang, L.F., Xie, M. & Tang, L.C., 2005. A Study of Two Estimation Approaches for Parameters of Two-parameter Weibull Distribution based on WPP. In The Fourth International Conference on Quality and Reliability (ICQR4). Beijing, China.

Zhang, L.F., Xie, M. & Tang, L.C., 2005. On Least Squares Estimators of the Weibull Shape Parameter for Multiply Censored Data. In International Conference on Modeling and Analysis of Semiconductor Manufacturing (MASM2005). Singapore.

Zhang, L.F., Xie, M. & Tang, L.C., 2006. Bias Correction for the Least Squares Estimator of Weibull Shape Parameter with Complete and Censored Data. *Reliability Engineering and System Safety*, 91(8), 930-939.

Zhang, L.F., Xie, M. & Tang, L.C., 2006a. On Weighted Least Squares Estimation for Parameters of the Two-Parameter Weibull Distribution. In The 12th ISSAT International Conference on Reliability and Quality in Design. Chicago, USA.

Zhang, L.F., Xie, M. & Tang, L.C., 2006b. Robust Regression Using Probability Plot for the Estimation of Weibull Shape Parameter. *Quality and Reliability Engineering International*, 22(8), 905-917.

Zhang, L.F., Xie, M. & Tang, L.C., 2007. A Study of Two Estimation Approaches for Parameters of Two-parameter Weibull Distribution based on WPP. *Reliability Engineering and System Safety*, 92(3), 360-368.

Zhang, L.F., Xie, M. & Tang, L.C., 2008. Chapter 3: On Weighted Least Squares Estimation for the Parameters of Weibull Distribution. In *Recent Advances in Reliability and Quality in Design*. Springer Series in Reliability Engineering.

Appendix A

Derivation of Equations (3-8) – (3-10).

Based on the CDF and PDF of the reduced variable Z, i.e.,

$$F(z) = 1 - \exp(-e^z)$$

$$f(z) = \exp(z - e^z)$$

the CDF of the i^{th} order statistic $Z_{(i)}$ $(1 \le i \le n)$ is given by

$$\begin{aligned} F(z_{(i)}) &= i\binom{n}{i}\int_{-\infty}^{z} F^{i-1}(z)(1-F(z))^{n-i} f(z)dz \\ &= i\binom{n}{i}\int_{-\infty}^{z} (1-e^{-e^z})^{i-1}(e^{-e^z})^{n-i} d(1-e^{-e^z}) \end{aligned}$$

and its PDF is

$$\begin{aligned} f(z_{(i)}) &= i\binom{n}{i} F^{i-1}(z)(1-F(z))^{n-i} f(z) \\ &= i\binom{n}{i}(1-e^{-e^z})^{i-1}(e^{-e^z})^{n-i} e^{-e^z} e^z \end{aligned}$$

The mean of $Z_{(i)}$, by definition, can be obtained by

$$E(Z_{(i)}) = \int_{-\infty}^{\infty} zf(z_{(i)})dz = i\binom{n}{i}\int_{-\infty}^{+\infty} z(1-e^{-e^z})^{i-1}(e^{-e^z})^{n-i} e^{-e^z} e^z dz$$

Setting $v = e^z$, so that $z = \ln v, dz = dv/v$, and the above equation becomes

$$\begin{aligned} E(Z_{(i)}) &= i\binom{n}{i}\int_{0}^{+\infty} \ln v \cdot (1-e^{-v})^{i-1} e^{-(n-i+1)v} dv \\ &= i\binom{n}{i}\int_{0}^{+\infty} \ln v \cdot (e^{v}-1)^{i-1} e^{-nv} dv \end{aligned}$$

Making advantage of the binominal theorem, i.e.,

$$(x+a)^n = \sum_{k=0}^{n}\binom{n}{k}x^k a^{n-k}$$

we have

$$(e^v - 1)^{i-1} = \sum_{k=0}^{i-1}\binom{i-1}{k}(-1)^k e^{v(i-1-k)}$$

Thus

$$E(Z_{(i)}) = i\binom{n}{i}\sum_{k=0}^{i-1}\binom{i-1}{k}(-1)^k \int_0^{+\infty} \ln v \cdot e^{(n-i+k+1)v} dv$$

Let $T = (n-i+k+1)t$, after replacing, we have

$$\begin{aligned} E(Z_{(i)}) &= i\binom{n}{i}\sum_{k=0}^{i-1}\binom{i-1}{k}(-1)^k \int_0^{+\infty}\left[\ln T - \ln(n-i+k+1)\right]\cdot \frac{e^{-T}}{n-i+k+1}dT \\ &= i\binom{n}{i}\sum_{k=0}^{i-1}\binom{i-1}{k}(-1)^k \frac{1}{n-i+k+1}\left[\int_0^{+\infty}\ln T \cdot e^{-T} dT - \right. \\ &\qquad \left. \ln(n-i+k+1)\int_0^{+\infty} e^{-T} dT\right] \end{aligned}$$

Since

$\int_0^{+\infty} \ln T \cdot e^{-T} dT = -\gamma = -0.577216$, where γ is the Euler's constant

$\int_0^{+\infty} e^{-T} dT = 1$ and

$i\binom{n}{i}\sum_{k=0}^{i-1}(-1)^k\binom{i-1}{k}\frac{1}{n-i+k+1} = 1$

Finally we have

$$E(Z_{(i)}) = i\binom{n}{i}\cdot\sum_{k=0}^{i-1}\left\{\binom{i-1}{k}(-1)^k \cdot \frac{-\gamma - \ln(n-i+k+1)}{n-i+k+1}\right\}$$

which is Equation (3-8).

Similarly $E(Z_{(i)}^2)$ can be obtained. By definition,

$$E(Z_{(i)}^2) = i\binom{n}{i}\int_{-\infty}^{+\infty} z^2 (1-e^{-e^z})^{i-1} (e^{-e^z})^{n-i} e^{-e^z} e^z dz$$

Replacing e^z by v,

$$\begin{aligned}
E(Z_{(i)}^2) &= i\binom{n}{i}\int_0^{+\infty} \ln^2 v \cdot (e^v - 1)^{i-1} \cdot e^{-nv} dv \\
&= i\binom{n}{i}\sum_{k=0}^{i-1}\binom{i-1}{k}(-1)^k \int_0^{+\infty} \ln^2 v \cdot e^{-(n-i+k+1)v} dv \\
&= i\binom{n}{i}\sum_{k=0}^{i-1}\binom{i-1}{k}(-1)^k \int_0^{+\infty} \left[\ln T - \ln(n-i+k+1)\right]^2 \cdot \frac{e^{-T}}{n-i+k+1} dT \\
&= i\binom{n}{i}\sum_{k=0}^{i-1}\binom{i-1}{k}(-1)^k \frac{1}{n-i+k+1}\Big[\int_0^{+\infty} \ln^2 T \cdot e^{-T} dT + \\
&\quad 2\gamma \ln(n-i+k+1) + \ln^2(n-i+k+1)\Big]
\end{aligned}$$

Since

$$\int_0^{+\infty} \ln^2 T \cdot e^{-T} dT = 1.978112$$

Finally we have

$$E(Z_{(i)}^2) = 1.978112 + i\binom{n}{i}\sum_{k=0}^{i-1}\left\{\binom{i-1}{k}(-1)^k \cdot \frac{2\gamma \ln(n-i+k+1) + \ln^2(n-i+k+1)}{n-i+k+1}\right\}$$

which is Equation (3-9).

The joint density function of two order statistics, $Z_{(i)}$ and $Z_{(j)}$ $(1 \le i < j \le n)$, is given by

$$\begin{aligned}
f(z_i, z_j) &= n(j-i)\binom{n-1}{j-1}\binom{j-1}{i-1}[F(z_i)]^{i-1}[F(z_j) - F(z_i)]^{j-i-1}[1-F(z_j)]^{n-j} f(z_i) f(z_j) \\
&= \binom{n}{j}\binom{j}{i-1}[1-\exp(-e^{z_i})]^{i-1}[\exp(-e^{z_i}) - \exp(-e^{z_j})]^{j-i-1}[\exp(-e^{z_j})]^{n-j} \cdot \\
&\quad e^{z_i}\exp(-e^{z_i}) e^{z_j}\exp(-e^{z_j})
\end{aligned}$$

From the definition,

$$E(Z_{(i)}Z_{(j)}) = \int_{-\infty}^{+\infty}\int_{-\infty}^{z_j} z_i z_j f(z_i, z_j) dz_i dz_j$$

$$= \binom{n}{j}\binom{j}{i-1}\int_{-\infty}^{+\infty}\int_{-\infty}^{z_j} z_i z_j [1-\exp(-e^{z_i})]^{i-1}[\exp(-e^{z_i}) - \exp(-e^{z_j})]^{j-i-1}[\exp(-e^{z_j})]^{n-j} \cdot$$

$$e^{z_i}\exp(-e^{z_i})e^{z_j}\exp(-e^{z_j})$$

Setting $u = e^{z_i}$ and $v = e^{z_j}$ and re-write the above equation,

$$E(Z_{(i)}Z_{(j)}) = \binom{n}{j}\binom{j}{i-1}\int_{0}^{+\infty}\int_{0}^{v} \ln u \ln v [1-e^{-u}]^{i-1}[e^{-u} - e^{-v}]^{j-i-1}[e^{-v}]^{n-j} e^{-u} e^{-v} du dv$$

$$= \binom{n}{j}\binom{j}{i-1}\int_{0}^{+\infty}\int_{0}^{v} \ln u \ln v [1-e^{-u}]^{i-1}[e^{-u} - e^{-v}]^{j-i-1} e^{-u} e^{-(n-j+1)v} du dv$$

which is Equation (3-10).