

**NUMERICAL ISSUES IN ESTIMATION
OF CONTINUOUS PARAMETRIC DISTRIBUTIONS**

**A Thesis Presented to
the Faculty of the School of Engineering and Applied Science
University of Virginia**

**In Partial Fulfillment
of the Requirements for the Degree
Master of Science
in Systems Engineering**

**By
Yiwei Zhang**

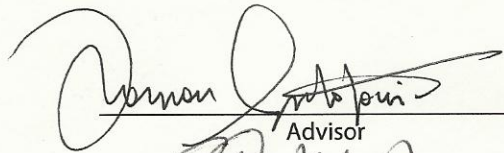
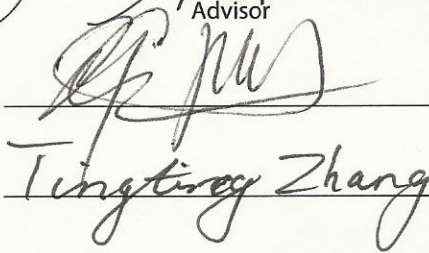
May 2012

APPROVAL SHEET

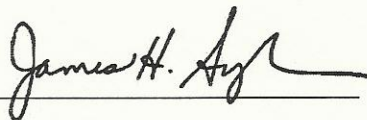
The thesis
is submitted in partial fulfillment of the requirements
for the degree of
Master of Science


AUTHOR

The thesis has been read and approved by the examining committee:


Advisor

Tingting Zhang

Accepted for the School of Engineering and Applied Science:


Dean, School of Engineering and Applied Science

May
2012

ABSTRACT

Continuous variates are used everywhere (almost) in stochastic modeling. This thesis addresses numerical issues arising in the process of estimating a continuous parametric distribution function. It aims to provide a guide to analysts on how to overcome some problems we have encountered. In detail, it (1) applies the uniform method for estimating a one- or two-parameter distribution function from a complete sample; (2) derives the Conditional Empirical Distribution method for estimating distribution function from a censored sample (of any type); (3) illustrates the superiority of the Conditional Empirical Distribution method over the Maximum Likelihood Estimation method; (4) determines the reason for difficulties (unbounded solutions) in optimization of Pareto distribution parameters; (5) demonstrates the fallacy of applying the goodness-of-fit tests meant for discrete distributions, such as the chi-square test, to continuous distributions.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. ATM – 0641572, “New Statistical Techniques for Probabilistic Weather Forecasting”. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

The ensemble data were provided by the Environmental Modeling Center of the National Centers for Environmental Prediction, National Weather Service, National Oceanic and Atmospheric Administration, Camp Springs, Maryland.

I am glad to devote two of my best years to this project. I would like to give great thanks to my advisor, Dr. Roman Krzysztofowicz. The knowledge and experience I have gained from him will have deep impression on my future life.

Great thanks to my mother and father. You are always by my side and support me no matter what my decision is. Home is always my sweetest dream because of you.

Last I want to thank my friends here at University of Virginia. It is your accompany that gives me courage to start a new life at a strange place, and finally chase my dream. I will always remember the days we support each other. Before the assignment due, your accompany supports me through night after night under the yellow warm light.

TABLE OF CONTENTS

ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	iv
Chapter 1 : INTRODUCTION.....	1
1.1 Discrete versus Continuous Distribution.....	1
1.2 Motivation.....	2
1.3 Overview.....	2
1.3.1 Censored Samples.....	2
1.3.2 Optimization with Pareto Distribution.....	3
1.3.3 Fallacies Related to Testing Continuous Distributions.....	3
Chapter 2 : CONCEPTS.....	5
2.1 Density Function and Distribution Function.....	5
2.2 Empirical Distribution Function.....	5
2.3 Parametric Distribution Function.....	7
2.4 Distribution Fitter's Algorithm.....	7
2.5 Catalog of Distributions.....	10
2.5.1 The Weibull Distribution.....	10
2.5.2 The Pareto Distribution.....	11
Chapter 3 : CENSORED SAMPLE METHODS.....	12
3.1 Problem Definition.....	12
3.2 Existing Method.....	13
3.2.1 Preliminaries.....	13
3.2.2 Singly Censored Sample.....	13
3.2.3 MLE for Complete Sample.....	14
3.2.4 MLE for Upper Tail Censored Sample.....	15
3.2.5 MLE for Lower Tail Censored Sample.....	16
3.2.6 MLE for Two Tail Censored Sample.....	18
3.3 New Method.....	19
3.3.1 Preliminaries.....	19
3.3.2 Lower Tail Realizations are Censored.....	20
3.3.3 Upper Tail Realizations are Censored.....	23
3.3.4 Lower and Upper Tail Realizations are Censored.....	26
3.4 Numerical Experiments with New Method.....	27
3.4.1 Motivation.....	27
3.4.2 How to Assess the Lower/Upper Censoring Probability?.....	27
3.4.3 What is the Tolerance of the Method to the Percentage of Missing Realizations?	30
3.4.4 Experimental Results	31
3.5 Summary.....	36

Chapter 4 : COMPARISON OF METHODS.....	37
4.1 Objective.....	37
4.2 Design of Experiments.....	37
4.3 Experimental Results.....	38
4.4 Summary.....	38
Chapter 5 : THE PARETO DISTRIBUTION.....	60
5.1 Introduction.....	60
5.2 Pareto Distribution Function.....	60
5.3 Optimization with Pareto Distribution.....	62
5.4 Insight from the Moments.....	69
5.5 Conclusion.....	74
Chapter 6 : FALLACIES RELATED TO CONTINUOUS DISTRIBUTIONS.....	75
6.1 Chi-Square Test.....	75
6.2 Kolmogorov-Smirnov Test.....	75
6.3 Comparisons of the Tests.....	76
6.4 Conclusion.....	82
APPENDIX : DATA SAMPLES.....	83
A.1 Artificial Samples.....	83
A.2 Random Samples.....	84
A.3 River Stage Samples.....	85
A.4 National Weather Service Samples.....	88
REFERENCES.....	89

CHAPTER 1 : INTRODUCTION

1.1 Discrete versus Continuous Distribution

People tend to quantify their perceptions of interesting objects into mathematical expressions. This quantification facilitates measuring, controlling and forecasting the numeric values of the objects. In mathematics, we call the interesting object a variable. We call the variable a *random variable*, or *variate*, if its observed value is determined by chance (Milton & Arnold, 2004). If a random variable has a countable number of possible realizations, it is called a *discrete random variable*. If the number of possible realizations of a random variable is uncountable, it is called a *continuous random variable*.

Suppose we are measuring the temperature in Charlottesville, Virginia, at sunset. If we repeat the experiment every day for a week, we will get seven realizations of the temperature variable. In January, the temperature can take any value between 18 and 54, in degrees of Fahrenheit. Thus temperature is a continuous variate. The distribution of a continuous variate is called a continuous distribution. Suppose we are conducting another experiment, in which we are counting the number of emails received in one morning. A variable is used to record the number of emails received. If we repeat the experiment several times, we will get several realizations of the variable, each of which is a non-negative integer. Thus this variable is a discrete variate. The distribution of a discrete variate is called a discrete distribution.

In reality, we will meet many continuous variables. This thesis deals with numerical issues arising in modeling and estimation of parametric distributions of continuous random variables.

1.2 Motivation

The National Weather Service (NWS) is concerned with rapid changes of the weather, producing hazardous events such as flood, tornado, and heavy rainfall. To predict the rapid changes, they want to quantify some relevant aspects and construct a model. A predictand is a random variable whose value is to be predicted. For the flood, the predictand is the maximum river stage within some time interval. If the river stage is above the safe level for a certain time, there might be a flood coming. Our motivation is to assist them to quantify the uncertainty and to build a predictive model. More precisely, our job is to model the predictand in terms of a continuous parametric distribution. Then an inference or decision can be made based on the distribution model. For example, the NWS could issue watches and warnings based on this continuous distribution.

However, the procedures of modeling continuous variables are complex. The procedures involve sample acquisition and model fitting. This thesis is dedicated to solving some of the numerical problems encountered during modeling a continuous predictand.

1.3 Overview

1.3.1 *Censored Samples*

Collecting samples is not always a smooth process. Under ideal circumstances, we get all the data points. However, in life testing and natural disaster recording, we tend to finish the experiments with only part of the points acquired. For example, suppose we are trying to collect the lifetimes of the patients after taking certain medicine. The variable recording lifetime enjoys a continuous distribution. However, if the experiment only lasts for ten years, then there is a high probability that some patients in this test are still alive. Thus, we are unable to acquire some

realizations of the variable during the experiment, but only know that they are greater than 10 years. Here, the sample is not complete; it is called a *censored sample*. The number “10” with the unit “year” is called the *censoring threshold*. Chapter 3 describes the Maximum Likelihood Estimation method and our new method to construct a distribution function based on a censored sample. Chapter 4 reports the experiments comparing these methods.

1.3.2 Optimization with Pareto Distribution

Pareto is a continuous distribution with two parameters from the catalog of distributions to be used in probabilistic forecasting. The reason we are curious about it is that the optimization process usually iterates much beyond a satisfactory solution point. This long search wastes the valuable computing time without making any significant improvement in the fit. We want to stop the search early and achieve a balance between an acceptable computing time and a reasonable goodness-of-fit. The optimization process is introduced in Chapter 2 and the Pareto distribution problem is investigated in Chapter 5. The task is to discover whether it is the peculiarity of the Pareto distribution, or it is a shortcoming of the optimization method that leads to a long search in the optimization process. The reason is explained theoretically and demonstrated experimentally.

1.3.3 Fallacies Related to Testing Continuous Distributions

For simplicity of calculation and representation, under many circumstances we transform a task associated with a continuous variable into a discrete form. This is also one situation people may encounter in testing the goodness-of-fit of a hypothesized continuous distribution function. In Chapter 6, we will introduce two goodness-of-fit tests: the Chi-square test and the Kolmogorov-Smirnov test. Those tests aim at testing the null hypothesis that a given set of

observations is drawn from, or “fits,” a specified probability distribution. The Chi-square test is only suitable for discrete variate and K-S test is only suitable for continuous variate. However, people sometimes mistakenly perform the Chi-square test on a continuous variate; this may produce conflicting test results either within Chi-square test itself, or between Chi-square test and K-S test.

CHAPTER 2 : CONCEPTS

This chapter introduces the mathematical techniques and concepts which will be used later. Those concepts include: density function, distribution function, and distribution types. As most of the experiments in Chapters 3 and 4 are based on the distribution fitter software, I also explain the algorithm and methodology of the software.

2.1 Density Function and Distribution Function

A distribution function H of a continuous variate X specifies at every point x the probability of X taking on a value less than or equal to x ; that is $H(x) = P(X \leq x)$, which is also called the nonexceedance probability. Distribution function serves as the first description of a continuous variate. In later chapters, we construct experiments in estimating distribution functions with different methods, and from different types of samples.

A density function h is the derivative of its corresponding distribution function. However, a point's density function value $h(x)$ does not have a practical meaning.

2.2 Empirical Distribution Function

Suppose we have a sample $\{x(n): n = 1, \dots, N\}$ of a continuous random variable X . When the realizations are arranged in ascending order, we denote them as

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(N)},$$

and they constitute the ordered sample $\{x_{(n)}: n = 1, \dots, N\}$.

Our objective is to construct a distribution function $H(x) = P(X \leq x)$ of the variate X . The estimator of this unknown distribution function is constructed from a given sample, without

making any assumptions about the form of H . Such an estimator is called the empirical distribution function (E.D.F.) of X , denoted by \tilde{H} . Let us first define N nonexceedance probabilities p_n , also called the plotting positions, as

$$p_n = P(X \leq x_{(n)}), \quad n = 1, \dots, N. \quad (2.1)$$

Then the E.D.F. is specified by the points

$$\{(x_{(n)}, p_n) : n = 1, \dots, N\}. \quad (2.2)$$

There are multiple ways to specify the plotting positions: standard, Weibull, Meta-Gaussian, etc.

This thesis is based on the meta-Gaussian plotting positions (Krzysztofowicz, 1992, 2012),

where

$$p_n = \left[\left(\frac{N-n+1}{n} \right)^{t_N} + 1 \right]^{-1}. \quad (2.3)$$

The value of t_N varies with the sample size N , and is specified as

$$t_N = 3.0193N^{-1.1018} + 1, \quad 2 \leq N \leq 3; \quad (2.4a)$$

$$t_N = 2.4035N^{-0.9096} + 1, \quad 4 \leq N \leq 5; \quad (2.4b)$$

$$t_N = 2.1408N^{-0.8423} + 1, \quad 6 \leq N \leq 10; \quad (2.4c)$$

$$t_N = 1.9574N^{-0.8039} + 1, \quad 11 \leq N \leq 20000; \quad (2.4d)$$

$$t_N = 1, \quad 20000 < N. \quad (2.4e)$$

2.3 Parametric Distribution Function

Empirical distribution function is a quantized version of the distribution function of a continuous variate. However, to obtain the probability $P(X \leq x)$ at any point x , we need to acquire a completely continuous estimate of the distribution function H .

There are many parametric models for distribution functions. In this thesis, we use a catalog of 28 parametric distribution functions (Krzysztofowicz, 2011a). Each of the functions has two parameters: the scale parameter and the shape parameter. We may also need to determine one or two bounds on the sample space for the purpose of fitting. To estimate the parameters of the distribution function, we go through two stages described in Section 2.4.

To measure the goodness of fit of the distribution function, we define *the maximum absolute difference* (MAD) as the criterion:

$$MAD = \max_{1 \leq n \leq N} [|p_n - H(x_{(n)})|]. \quad (2.5)$$

2.4 Distribution Fitter's Algorithm

The distribution fitter (Krzysztofowicz, 2011a) is FORTRAN based software estimating parametric distributions of continuous variates. The input to the software is a sample of the variate, with the number of realizations ranging from 2 to 20,000. The output from the software is a distribution fitted for the variate based on the input sample. The software itself contains four groups of distributions defined on four types of sample spaces: the unbounded interval, the bounded-below interval, the bounded-above interval, and the bounded open interval. There are several parametric distributions in each group. For example, we have exponential distribution, Weibull distribution, log-Weibull distribution for bounded-below interval sample space. The

distribution parameters are estimated in two stages: first, the estimates are obtained using a least squares method; second, the parameters are optimized using a non-linear method. The algorithm used in the distribution fitter is as follows (Krzysztofowicz, 2011a).

Step 1. Initialize the inputs. For this step, we need a sample of size $2 - 20000$ of a continuous variate. If the sample space has a bound, we also ask for the bound value. Typically, the input needs to specify the type of the sample space.

Step 2. Sort the sample from the smallest value to the largest one. Ask the user to specify whether it is a censored sample or a complete sample. If it is a censored sample, we assess a censoring probability according to the methodology in Chapter 3. Then we calculate unconditional plotting positions for the sample using the meta-Gaussian method.

Step 3. Prune the duplicate realizations in the sample. If a sample has more than two realizations which have the same value, it should go through the pruning process. If the number of the same realizations is odd, only the middle one is kept after pruning. If the number of the same realizations is even, only the middle two realizations are kept after pruning. In this pruning process (i) input is the sample and its plotting positions, and (ii) output is the pruned sample and its corresponding plotting position. It has nothing to do with the distribution type.

Step 4. Initialize the distribution function parameters. Each distribution in the catalog has two parameters α and β to be estimated. The exception is exponential distribution which only has one parameter. The distribution fitter uses the least squares method for the initialization of the distribution parameters. With this method, the parametric distribution function is linearized, and so is the E.D.F. The details are as follows.

A given distribution function is transformed into its inverse function, i.e. the quantile function, denoted by $x = H^{-1}(p)$. The quantile function is linearized and takes the form

$$v = bu + a, \quad (2.6)$$

where v, u, a , and b are functions of x, p, α and β . Each distribution has different equations for v, u, a , and b . An example is shown in Section 2.5.1, which includes the linearized quantile function of the Weibull distribution function.

The linear regression is performed using the least squares method. The estimators are

$$\hat{b} = \frac{\sum_{n=1}^N v_n u_n - N \bar{v} \bar{u}}{\sum_{n=1}^N u_n^2 - N \bar{u}^2}, \quad (2.7a)$$

$$\hat{a} = \bar{v} - \hat{b} \bar{u}, \quad (2.7b)$$

where \bar{v} and \bar{u} are sample means.

The least squares estimates (\hat{a}, \hat{b}) are transformed back to the original coordinates to get the initial parameter value $(\hat{\alpha}, \hat{\beta})$ which are input to Step 5.

Step 5. Perform optimization of the parameters by using the downhill simplex method (Nelder and Mead, 1965) or the gradient method (Luenberger, 1984; Peressini *et al.*, 1993) with the Brent's method (Brent, 1973) for the line search. Calculate the MAD after optimization.

Step 6. Perform Steps 4 and 5 for all the distributions that the user wants to try. Compare their MADs. The distribution with the smallest MAD is the distribution fitting best. Perform the Kolmogorov-Smirnov test (Chapter 6) to evaluate the goodness-of-fit.

2.5 Catalog of Distributions

2.5.1 The Weibull Distribution

Density function

$$h(x) = \frac{\beta}{\alpha} \left(\frac{x-\eta}{\alpha} \right)^{\beta-1} \exp \left[- \left(\frac{x-\eta}{\alpha} \right)^{\beta} \right] \quad (2.8)$$

Distribution function

$$H(x) = 1 - \exp \left[- \left(\frac{x-\eta}{\alpha} \right)^{\beta} \right] \quad (2.9)$$

Quantile function

$$H^{-1}(p) = \alpha [-\ln(1-p)]^{\frac{1}{\beta}} + \eta \quad (2.10)$$

Linearized quantile function $v = bu + a$

$$v = \ln(x - \eta) \quad (2.11)$$

$$u = \ln[-\ln(1-p)] \quad (2.12)$$

$$\beta = 1/b \quad (2.13)$$

$$\alpha = \exp(a) \quad (2.14)$$

Moments

$$E(X) = \alpha \Gamma \left(1 + \frac{1}{\beta} \right) + \eta \quad (2.15)$$

$$VAR(X) = \alpha^2 [\Gamma \left(1 + \frac{2}{\beta} \right) - \Gamma^2 \left(1 + \frac{1}{\beta} \right)] \quad (2.16)$$

2.5.2 The Pareto Distribution

Density function

$$h(x) = \frac{\beta}{\alpha} \left(\frac{\alpha}{x-\eta+\alpha} \right)^{\beta+1} \quad (2.17)$$

Distribution function

$$H(x) = 1 - \left(\frac{\alpha}{x-\eta+\alpha} \right)^{\beta} \quad (2.18)$$

Quantile function

$$H^{-1}(p) = \alpha[(1-p)^{-\frac{1}{\beta}} - 1] + \eta \quad (2.19)$$

Linearized quantile function $v = u + a$

$$v = \ln(x - \eta) \quad (2.20)$$

$$u = \ln[(1-p)^{-\frac{1}{\beta}} - 1] \quad (2.21)$$

$$\alpha = \exp(a) \quad (2.22)$$

Moments

$$E(X) = \frac{\alpha}{\beta-1} + \eta, \quad \text{if } \beta > 1 \quad (2.23)$$

$$VAR(X) = \frac{\beta \alpha^2}{(\beta-2)(\beta-1)^2}, \quad \text{if } \beta > 2 \quad (2.24)$$

CHAPTER 3 : CENSORED SAMPLE METHODS

3.1 Problem Definition

Collecting samples is not always a smooth process. In life testing, not every experiment produces a result. Suppose we are trying to collect the lifetimes of the patients who are taking certain medicine. If the experiment only lasts ten years, we are unable to acquire the lifetime of patients who live longer than ten years. This sample is called a censored sample of the variate lifetime. Since we know every realization is less than “10 years”, the number “10” with the unit “year” is called the *censoring threshold*.

Censored sample is also quite common in data analysis outside the industrial and medical areas. For example, when we apply a thermometer to measure an unknown set of temperatures, we can only get measurements within its range. Temperatures beyond that range will result in an inaccurate measurement or damage to the thermometer. Thus a censored sample is known to be incomplete as some values beyond the range are missing. However, it is still of great value to analyze the censored sample, provided an appropriate method of analysis is used. We are interested in estimation of a continuous distribution function based on a censored sample. An existing method is the Maximum Likelihood Estimation (MLE) method (e.g., Aitkin & Clayton, 1980; Balakrishnan & Kateri, 2008; Bar-Lev, 2004; Chen & Lio, 2010; Cohen, 1975; Engelhardt & Bain, 1973; Jang *et al.*, 2011; Klakattawi *et al.*, 2011; Thomas & Wilson, 1972). A new method is the conditional empirical distribution (CED) method. It is general in that it is applicable to any distribution type. This chapter describes the two methods; chapter 4 compares their performance experimentally.

3.2 Existing Method

3.2.1 Preliminaries

Cohen (1965) describes the Maximum Likelihood Estimation method for estimating parameters of the Weibull distribution from both complete and censored samples. The Weibull distribution is defined in Section 2.5.

According to Cohen, the Weibull density function is

$$f(x) = \left(\frac{\gamma}{\theta}\right) x^{\gamma-1} \exp\left(-\frac{x^\gamma}{\theta}\right); x \geq 0, \gamma > 0, \theta > 0. \quad (3.1)$$

Compared to (2.8), equation (3.1) assigns the shift parameter $\eta = 0$. The other two parameters of equation (3.1) show a relationship with parameters of equation (2.8) as:

$$\beta = \gamma, \alpha = \sqrt[\beta]{\theta}. \quad (3.2)$$

To simplify comparisons, we will just use the parameter symbols in (2.8) to illustrate maximum likelihood estimation (MLE) method.

3.2.2 Singly Censored Sample

Cohen (1965) defined *singly censored sample* by using a life test example. Suppose N specimens are placed under observation. The time of each failure is recorded. At some pre-determined fixed time x_0 or after acquiring a pre-determined fixed number of realizations, the test is terminated. Let N denote the number of specimens, and n denote the number of realizations acquired. In both of these cases, we have the sample x_1, \dots, x_n and the information that $N - n$ specimens survived beyond time x_0 or x_n . If the experiment stops at a fixed time x_0 ,

we get a type I censoring sample. All the realizations of the sample are smaller than or equal to x_0 . If the experiment stops after acquiring a pre-determined number n of realizations, we get a type II censoring sample. All the realizations of the sample are smaller than or equal to the last recorded failure time x_n .

3.2.3 MLE for Complete Sample

This derivation of MLE for complete sample follows Cohen (1965).

Consider a complete sample of size N . Suppose this sample is drawn from a Weibull distribution, thus the likelihood function of the Weibull distribution parameters is

$$L(x_1, \dots, x_N; \alpha, \beta) = \prod_{i=1}^N \left(\frac{\beta}{a^\beta}\right) x_i^{\beta-1} \exp(-x_i^\beta/a^\beta). \quad (3.3)$$

By taking logarithms of (3.3), differentiating with respect to β and a^β , we obtain

$$\begin{cases} \frac{\partial \ln L}{\partial \beta} = \frac{N}{\beta} + \sum_{i=1}^N \ln x_i - \frac{1}{a^\beta} \sum_{i=1}^N x_i^\beta \ln x_i = 0, \\ \frac{\partial \ln L}{\partial a^\beta} = -\frac{N}{a^\beta} + \frac{1}{a^{2\beta}} \sum_{i=1}^N x_i^\beta = 0. \end{cases} \quad (3.4)$$

After substitute one equation into another, we eliminate α and obtain

$$\frac{\sum_{i=1}^N x_i^\beta \ln x_i}{\sum_{i=1}^N x_i^\beta} - \frac{1}{\beta} = \frac{1}{N} \sum_{i=1}^N \ln x_i. \quad (3.5a)$$

This is the equation for parameter β , which can be solved using an iteration method. The solution for α depends on the solution for β , which is

$$\alpha = \left(\frac{1}{N} \sum_{i=1}^N x_i^\beta\right)^{1/\beta}. \quad (3.5b)$$

Equations (3.5a) and (3.5b) constitute the MLE method for a Weibull distribution.

3.2.4 MLE for Upper Tail Censored Sample

When Cohen (1965) talked about the solution for a singly censored sample, he meant an upper tail censored sample. The definition of upper tail censored sample is stated in Section 3.3. The solution is acquired through a procedure similar to that for a complete sample, but is preceded by one step: the collected censored sample of n realizations $\{x_i: i = 1, \dots, n\}$ is extended to a sample of size N by inserting artificial realizations x_{n+1}, \dots, x_N for the $N - n$ survivors whose values are set as follows. In type I censoring, with x_0 fixed, the censoring threshold is $x_U = x_0$, and

$$x_i = x_0, \quad i = n + 1, \dots, N.$$

In type II censoring, the censoring threshold is $x_U = x_n$, and we set

$$x_i = x_n, \quad i = n + 1, \dots, N.$$

Then the likelihood function of the Weibull distribution parameters is

$$L(x_1, \dots, x_N; \alpha, \beta) = \frac{N!}{(N-n)!} \prod_{i=1}^n \left(\frac{\beta}{a^\beta}\right) x_i^{\beta-1} \exp(-x_i^\beta / a^\beta) [1 - F(x_U)]^{N-n}, \quad (3.6)$$

where

$$F(x) = 1 - \exp\left(-\frac{x^\beta}{a^\beta}\right). \quad (3.7)$$

By taking logarithms of (3.6), differentiating with respect to β and a^β , we obtain

$$\begin{cases} \frac{\partial \ln L}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^n \ln x_i - \frac{1}{a^\beta} \sum_{i=1}^N x_i^\beta \ln x_i = 0, \\ \frac{\partial \ln L}{\partial a^\beta} = -\frac{n}{a^\beta} + \frac{1}{a^{2\beta}} \sum_{i=1}^N x_i^\beta = 0. \end{cases} \quad (3.8)$$

After substituting one equation into another, we obtain

$$\frac{\sum_1^N x_i^\beta \ln x_i}{\sum_1^N x_i^\beta} - \frac{1}{\beta} = \frac{1}{n} \sum_1^n \ln x_i. \quad (3.9a)$$

This is the equation for parameter β , which can be solved using an iteration method. The solution for α depends on the solution for β , which is

$$\alpha = \left(\frac{1}{n} \sum_1^N x_i^\beta \right)^{1/\beta}. \quad (3.9b)$$

3.2.5 MLE for Lower Tail Censored Sample

Cohen (1965) only derived the MLE method given an upper tail censored sample. However, in many cases, we need to estimate a distribution function given a lower tail censored sample. Our new method (described in Section 3.3) can deal with both the upper tail and the lower tail censored samples. To complete the maximum likelihood estimation method and to better compare it with our new method, I derive the MLE method for a lower tail censored sample in parallel to the Cohen's method.

Suppose we are given a lower tail censored sample of n realizations $\{x_i: i = 1, \dots, n\}$. A censoring threshold x_L equals either x_0 or x_n , as in Type I or Type II censoring, and is such that $x_L \leq x_i$ for $i = 1, \dots, n$.

Thus the likelihood function of the Weibull distribution parameters becomes

$$L(x_1, \dots, x_N; \alpha, \beta) = \frac{N!}{(N-n)!} \prod_{i=1}^n \left(\frac{\beta}{\alpha^\beta} \right) x_i^{\beta-1} \exp(-x_i^\beta / \alpha^\beta) [F(x_L)]^{N-n}, \quad (3.10)$$

where

$$F(x) = 1 - \exp\left(-\frac{x^\beta}{a^\beta}\right). \quad (3.11)$$

By taking logarithm of (3.10), differentiating with respect to β and a^β , we obtain

$$\frac{\partial \ln L}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^n \ln x_i - \frac{1}{a^\beta} \sum_{i=1}^n x_i^\beta \ln x_i + (N-n) \frac{\frac{1}{a^\beta} e^{-x_L^\beta/a^\beta} x_L^\beta \ln x_L}{1 - e^{-x_L^\beta/a^\beta}} = 0, \quad (3.12)$$

and

$$\frac{\partial \ln L}{\partial a^\beta} = -\frac{n}{a^\beta} + \frac{1}{a^{2\beta}} \sum_{i=1}^n x_i^\beta - (N-n) \frac{\frac{1}{a^{2\beta}} e^{-x_L^\beta/a^\beta} x_L^\beta}{1 - e^{-x_L^\beta/a^\beta}} = 0. \quad (3.13)$$

The task now is to solve the two simultaneous equations (3.12) – (3.13) for α and β . For this purpose, we transform equation (3.13), and get expression

$$1 - e^{-x_L^\beta/a^\beta} = \frac{(N-n)x_L^\beta}{\sum_{i=1}^n x_i^\beta + (N-n)x_L^\beta - n\alpha^\beta}. \quad (3.14a)$$

We can also transform equation (3.13) with β on the left side, and get expression

$$\beta = \frac{\ln\left(-\alpha^\beta \ln\left(1 - \frac{(N-n)x_L^\beta}{\sum_{i=1}^n x_i^\beta + (N-n)x_L^\beta - n\alpha^\beta}\right)\right)}{\ln x_L}. \quad (3.14b)$$

We insert (3.14a) into (3.12) and derive an expression of α^β

$$\alpha^\beta = \frac{\sum_{i=1}^n x_i^\beta (\ln x_i - \ln x_L)}{\frac{n}{\beta} + \sum_{i=1}^n (\ln x_i - \ln x_L)}. \quad (3.15a)$$

We get the solution for α as

$$\alpha = \sqrt[\beta]{\frac{\sum_{i=1}^n x_i^\beta (\ln x_i - \ln x_L)}{\frac{n}{\beta} + \sum_{i=1}^n (\ln x_i - \ln x_L)}}. \quad (3.15 b)$$

Finally we use (3.14b) and (3.15b) to solve for β and α . We insert (3.15b) into (3.14b), to get an equation only with β only. This equation can be solved for β using an iteration method. We employed a graphical method as follows. If we plot the left hand side and right hand side of (3.14b) with various β values, we can find the intersection point as the solution. Figure 3.1 is a plot of left and right hand sides of (3.14b), given sample A2 with one lower realization censored. There is an intersection point at $\beta = 3.07$, and it is the maximum likelihood estimate of the Weibull distribution parameter. Finally, we insert the estimate of β into (3.15b) to obtain the maximum likelihood estimate of α .

From an inspection of equation (3.5a), (3.9a) and (3.14a), we draw a conclusion that the MLE solution for a complete sample is convergent. If we keep the expression of $1/\beta$ on the left side of equation (3.5a), and move the remaining expression to the right side, we can see the left part is strictly decreasing on positive axis and the right part is increasing. The intersection of these two curves guarantees that the solution exists. However, the MLE solution for a censored sample does not necessarily exist. The left hand side and right hand side of the MLE censored method may not intersect. In other words, the optimization problem is not concave.

3.2.6 MLE for Two Tail Censored Sample

As the two tail censored sample case is much more complex, we do not derive the MLE method for it. However, our new method can deal with the two tail censored case, which is explained in Section 3.3.

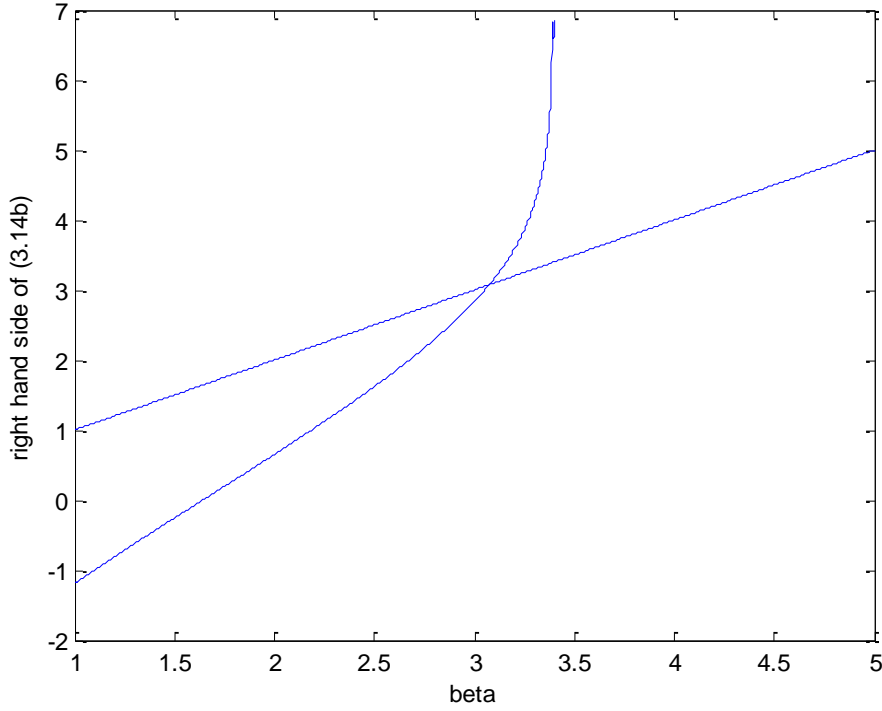


Figure 3.1. Graphical procedure for finding the maximum likelihood estimate of β . The X axis is the left hand side of equation (3.14b), while the Y axis is the right hand side of (3.14b). The data is artificial sample A2 with one lower realization censored.

3.3 New Method

3.3.1 Preliminaries

The CED method (Krzysztofowicz, 2011b) is general in that it can be applied to estimate parameters of any distribution type without the need to perform any derivations as required by any MLE method. Moreover, it is applicable to all three types of censored samples. Whereas the MLE method requires challenging derivations, even for distribution types as simple as Weibull distribution.

In the new method, the censored samples are classified into one of the three types according to the positions of the censored realizations in the whole distribution: (i) lower tail realizations are censored, (ii) upper tail realizations are censored, and (iii) lower and upper tail realizations are censored.

3.3.2 *Lower Tail Realizations are Censored*

We consider a random sample consisting of N realizations. If the points below certain threshold are missing, the sample is said to have lower tail censored. The threshold is called lower censoring threshold, and is denoted by x_L . The censored sample is defined as

$$\{x(n): x_L < x(n), n = 1, \dots, N\}. \quad (3.16)$$

Our purpose is to obtain the plotting positions of a complete sample, from which the censored sample is assumed to be drawn. First, we construct the empirical distribution of X , given the censored sample and the meta Gaussian plotting positions in Section 2.2.

To complete the calculation, lower censoring probability is assessed. Let p_L denote the lower censoring probability. The larger the number of points censored, the larger the lower censoring probability p_L is. In some situations, p_L and x_L can be exchanged.

Given the lower censoring threshold x_L , we can define H_L as the conditional distribution function of X , given that X is greater than the lower censoring threshold x_L . That is,

$$H_L(x) = P(X \leq x | X > x_L). \quad (3.17)$$

In contrast to the conditional distribution function H_L , the unconditional distribution function of X is denoted by H . The relationship between H_L and H is

$$\begin{aligned}
H_L(x) &= P(X \leq x | X > x_L) \\
&= 1 - P(X > x | X > x_L) \\
&= 1 - \frac{P(X > x, X > x_L)}{P(X > x_L)} \\
&= 1 - \frac{P(X > x)}{P(X > x_L)} \\
&= 1 - \frac{1-H(x)}{1-H(x_L)}.
\end{aligned} \tag{3.18}$$

According to the definition, the lower censoring probability p_L is

$$p_L = P(X \leq x_L) = H(x_L). \tag{3.19}$$

Applying Equation (3.19) to Equation (3.18), we get

$$H_L(x) = 1 - \frac{1-H(x)}{1-p_L} = \frac{H(x)-p_L}{1-p_L}. \tag{3.20}$$

Finally, the unconditional distribution function of X is derived:

$$H(x) = (1 - p_L)H_L(x) + p_L. \tag{3.21}$$

When analyzing a censored sample, we use p_n to denote the conditional plotting positions and π_n to denote the unconditional plotting positions. The relationship between them, based on Equation (3.21), can be written as

$$\pi_n = (1 - p_L)p_n + p_L, \quad n = 1, \dots, N. \tag{3.22}$$

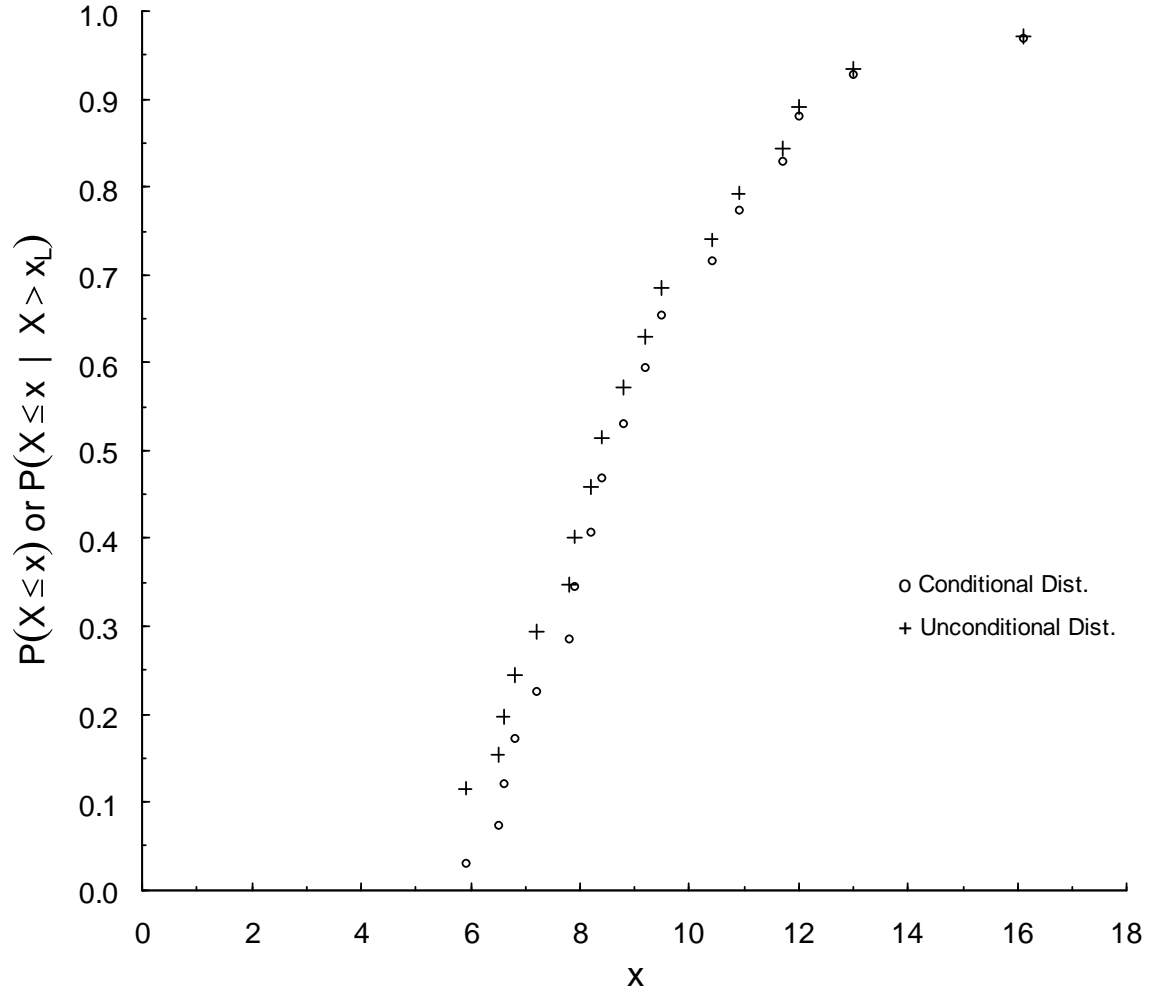


Figure 3.2. The conditional and unconditional empirical distribution functions constructed from a sample with lower tail censored. The sample used is A2 (Appendix A), with lower two realizations censored (out of 20), so that $x_L < 5.9$, and $p_L = 0.0872$. The censoring probability p_L is determined according to Section 3.4.3.

Figure 3.2 presents the conditional empirical distribution function $\{(x_{(n)}, p_n): n = 1, \dots, N\}$ and the corresponding empirical unconditional distribution function $\{(x_{(n)}, \pi_n): n = 1, \dots, N\}$ given a lower tail censored sample of size $N = 18$.

3.3.3 Upper Tail Realizations are Censored

Here is another random sample consisting of N realizations. If the points above certain threshold are missing, the sample is said to have its upper tail censored. The threshold is called upper censoring threshold, and is denoted by x_U . The censored sample is therefore defined as

$$\{x(n): x_U > x(n), n = 1, \dots, N\}. \quad (3.23)$$

For upper tail censored sample, the conditional plotting position p_n enjoys the same definition as in Section 3.2.1. Let p_U denote the upper censoring probability. Similarly, in some situations, p_U and x_U can be switched.

The conditional distribution function of X , given that X is less than x_U , is denoted by H_U . That is,

$$H_U(x) = P(X \leq x | X < x_U). \quad (3.24)$$

The relationship between H_U and H is

$$\begin{aligned} H_U(x) &= P(X \leq x | X < x_U) \\ &= \frac{P(X \leq x, X < x_U)}{P(X < x_U)} \end{aligned}$$

$$\begin{aligned}
&= \frac{P(X \leq x)}{P(X < x_U)} \\
&= \frac{H(x)}{H(x_U)}.
\end{aligned} \tag{3.25}$$

According to the definition, the upper censoring probability p_U is

$$p_U = P(X \leq x_U) = H(x_U). \tag{3.26}$$

Applying Equation (3.26) to Equation (3.25), we get

$$H_U(x) = \frac{H(x)}{p_U}. \tag{3.27}$$

Finally, the unconditional distribution function of X is derived:

$$H(x) = p_U H_U(x). \tag{3.28}$$

When analyzing a censored sample, we use p_n to denote the conditional empirical distribution and π_n to denote the unconditional empirical distribution. The relationship between them, based on Equation (3.28), can be written as

$$\pi_n = p_U p_n, \quad n = 1, \dots, N. \tag{3.29}$$

Figure 3.3 presents the conditional empirical distribution function and the corresponding unconditional distribution function, given an upper tail censored sample of size $N = 18$.

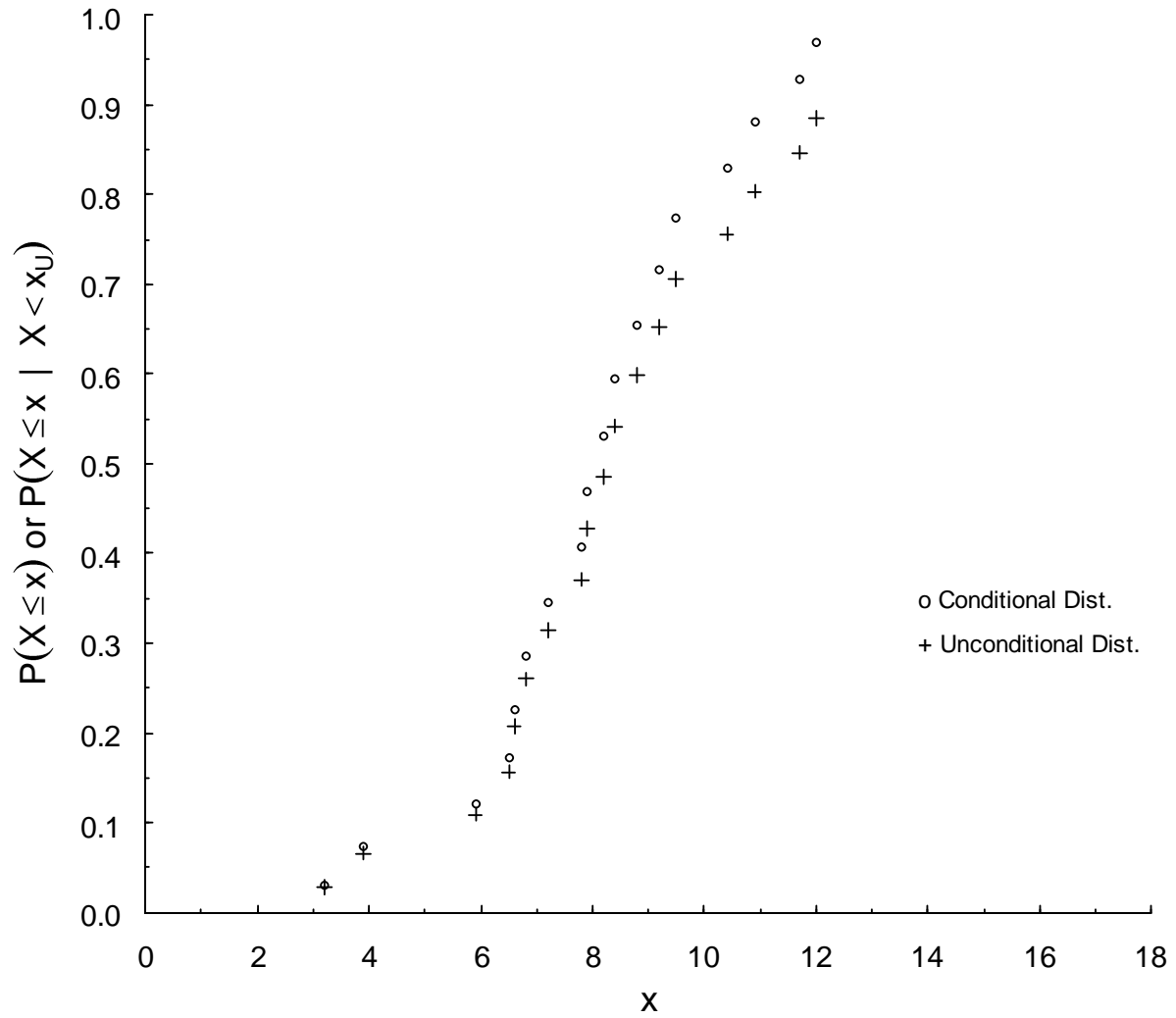


Figure 3.3. The conditional and unconditional empirical distribution functions constructed from a sample with upper tail censored. The sample used is A2 (Appendix A), with upper two realizations censored (out of 20), so that $x_U > 12.0$, and $p_U = 0.9127$. The censoring probability p_U is determined in Section 3.4.3.

3.3.4 Lower and Upper Tail Realizations are Censored

Consider another random sample consisting of N realizations. If the points below certain threshold and above certain threshold are missing, the sample is called two tail censored sample. All the variables and notation are the same as Section 3.3.2 & 3.3.3, except the censored sample

$$\{x(n): x_L < x(n) < x_U, n = 1, \dots, N\}, \quad (3.30)$$

and the conditional distribution function of X , given that X is less than x_U and great than x_L :

$$H_B(x) = P(X \leq x | X > x_L, X < x_U). \quad (3.31)$$

The relationship between H_B and H is

$$\begin{aligned} H_B(x) &= P(X \leq x | X > x_L, X < x_U) \\ &= \frac{P(X \leq x, X > x_L, X < x_U)}{P(X > x_L, X < x_U)} \\ &= \frac{P(X \leq x, X > x_L)}{H(x_U) - H(x_L)} \\ &= \frac{P(X \leq x | X > x_L) P(X > x_L)}{H(x_U) - H(x_L)} \\ &= \frac{[1 - P(X > x | X > x_L)] P(X > x_L)}{H(x_U) - H(x_L)} \\ &= \frac{[1 - \frac{P(X > x)}{P(X > x_L)}] P(X > x_L)}{H(x_U) - H(x_L)} \end{aligned}$$

$$\begin{aligned}
&= \frac{H(x) - H(x_L)}{H(x_U) - H(x_L)} \\
&= \frac{H(x) - p_L}{p_U - p_L}.
\end{aligned} \tag{3.32}$$

Finally, when analyzing a censored sample, the relationship between the conditional and unconditional plotting positions can be written as

$$\pi_n = p_L + (p_U - p_L)p_n. \tag{3.33}$$

Figure 3.4 presents the conditional distribution function and the corresponding unconditional distribution function, given a lower and upper tail censored sample of size $N = 16$.

3.4 Numerical Experiments with New Method

3.4.1 Motivation

In Section 3.3, we describe our method to deal with the three types of censored sample. To test the performance of the algorithm, we design several numerical experiments. The experiments are designed to answer two questions: (i) How to assess the lower/upper censoring probability, given a threshold? (ii) What is the tolerance of the new method to the percentage of missing realizations?

3.4.2 How to Assess the Lower/Upper Censoring Probability?

The three basic components for the simplest input-output system are input, output, and the function transforming input to output. Our experiment is such an input-output system.

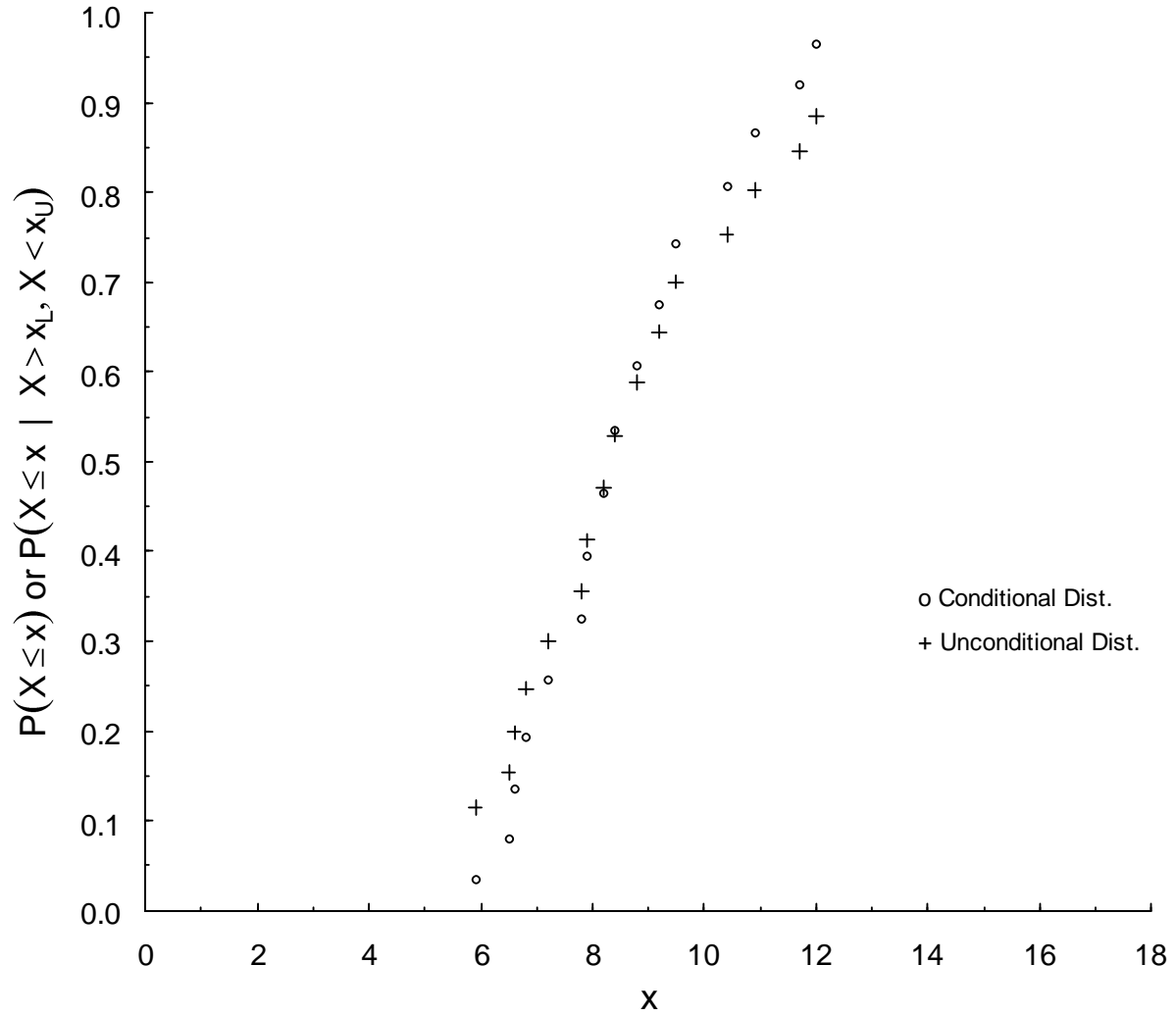


Figure 3.4. The conditional and unconditional empirical distribution functions constructed from a sample with lower and upper tail censored. The sample used is A2 (Appendix A), with the two smallest and the two largest realizations censored (out of 20), so that $x_L < 5.9$, $p_L = 0.0872$, $x_U > 12.0$, and $p_U = 0.9127$. The censoring probability p_L and p_U are determined in Section 3.4.3.

The input to the experiment includes 2 samples, A2, A4 (Appendix A) and a small positive constant ε . The output of the experiment is a list of MAD (Maximum Absolute Difference) between the estimated parametric distribution function and the unconditional empirical distribution function, given different choice of lower or upper censoring probability.

Step 1. Sort the sample. Manually have the sample A2 and A4 censored: lower tail r realizations censored, and upper tail r realizations censored, where “ r ” is a constant chosen from the set $\{0, 1, 2, 3, 4, 5\}$.

Step 2. Set three values of p_L or p_U for each manually censored sample. The choices of lower censoring probability include: $p_r + \varepsilon(p_{r+1} - p_r)$, $p_{r+1} - \varepsilon(p_{r+1} - p_r)$, $0.5(p_r + p_{r+1})$. The choices for upper censoring probability include: $p_{N-r} + \varepsilon(p_{N-r+1} - p_{N-r})$, $p_{N-r+1} - \varepsilon(p_{N-r+1} - p_{N-r})$, $0.5(p_{N-r} + p_{N-r+1})$, where N is the uncensored sample size, 20 in A2 and A4 cases, and constant ε ($0 < \varepsilon < 1$) is chosen arbitrarily as $\varepsilon = 0.1$.

Step 3. Use the values from Step 2 to go through calculations described in Sections 3.3.2 and 3.3.3, getting the unconditional plotting positions π_n .

Step 4. Prune the sample according to step 3 of Section 2.4. Use the unconditional plotting positions as the input to step 4 of Section 2.4. By using the least square estimation method, we initialize the distribution function parameters. Perform optimization according to step 5 of Section 2.4. Through this step, we get our final estimates of distribution parameters. Calculate MAD.

Step 5. Compare the MAD of each choice for lower or upper censoring probability, and choose the censoring probability giving smallest and steadiest MADs.

3.4.3 What is the Tolerance of the Method to the Percentage of Missing Realizations?

The input to the experiment is five samples $N1, N2, N3, N4, N5$ (Appendix A) and lower and upper censoring probabilities p_L or p_U . The output from the experiment is a list of MAD between the estimated parametric distribution function and the unconditional empirical distribution function, given samples with various percentages of the missing realizations.

Step 1. Sort the sample. Manually have the sample $N1, N2, N3, N4$, and $N5$ censored: lower tail r realizations censored, and upper tail r realizations censored, where “ r ” is a constant chosen from the set $\{0,1,2,3,4,5,6,7,8,9,10\}$.

Step 2. According to the experimental result of Section 3.4.2, which is in Section 3.4.4, I choose $p_{r+1} - \varepsilon(p_{r+1} - p_r)$ as the lower censoring probability threshold, $p_{N-r} + \varepsilon(p_{N-r+1} - p_{N-r})$ as the upper censoring threshold for any $r \geq 1$.

Step 3. Go through calculations of Section 3.3.2 or 3.3.3, getting the unconditional plotting positions π_n .

Step 4. Prune the sample according to step 3 of Section 2.4. Use the unconditional plotting positions as the input to step 4 of Section 2.4. By using the least square estimation method, we initialize the distribution function parameters. Perform optimization according to step 5 of Section 2.4. Through this step, we get our final estimates of distribution parameters. Calculate MAD.

Step 5. Relate the MAD of censored sample to the number of missing points. Observe from which point the goodness of fit starts to deteriorate according to MAD.

3.4.4 Experimental Results

Experiment 1. Assessment of the Censoring Probability

We conduct two parallel tests on samples A2 and A4 using the distribution fitter. Table 3.1 and Table 3.2 state the results for lower censoring probability assessment; and Table 3.3 and Table 3.4 show the results for upper censoring probability assessment.

Table 3.1. Experimental results for sample A2 with lower tail censored. Different lower censoring probabilities and number of missing points are applied; the MAD is calculated as the difference between the Weibull distribution function and the unconditional empirical distribution function.

Number of Points Missing	p_L	MAD	Δ MAD	Preferred Method		
				$p_r + \varepsilon(p_{r+1} - p_r)$	$0.5(p_r + p_{r+1})$	$p_{r+1} - \varepsilon(p_{r+1} - p_r)$
0	0	0.0494	0			
1	$p_1 + \varepsilon(p_2 - p_1)$	0.0592	0.0098			
1	$(p_1 + p_2)/2$	0.0575	0.0081		V	
1	$p_2 - \varepsilon(p_2 - p_1)$	0.0692	0.0198			
2	$p_2 + \varepsilon(p_3 - p_2)$	0.0660	0.0166			
2	$(p_2 + p_3)/2$	0.0575	0.0081			
2	$p_3 - \varepsilon(p_3 - p_2)$	0.0532	0.0038			V
3	$p_3 + \varepsilon(p_4 - p_3)$	0.0674	0.0180			
3	$(p_3 + p_4)/2$	0.0570	0.0076		V	
3	$p_4 - \varepsilon(p_4 - p_3)$	0.0590	0.0096			
4	$p_4 + \varepsilon(p_5 - p_4)$	0.0682	0.0188			
4	$(p_4 + p_5)/2$	0.0578	0.0084			
4	$p_5 - \varepsilon(p_5 - p_4)$	0.0572	0.0078			V
5	$p_5 + \varepsilon(p_6 - p_5)$	0.0703	0.0209			
5	$(p_5 + p_6)/2$	0.0599	0.0105			
5	$p_6 - \varepsilon(p_6 - p_5)$	0.0596	0.0102			V
Average	$p_r + \varepsilon(p_{r+1} - p_r)$	0.0662	0.0168	Winning times		
Average	$(p_r + p_{r+1})/2$	0.0579	0.0085			
Average	$p_{r+1} - \varepsilon(p_{r+1} - p_r)$	0.0596	0.0102	0	2	3

Table 3.2. Experimental results for sample A4 with lower tail censored. Different lower censoring probabilities and number of missing points are applied; the MAD is calculated as the difference between the Weibull distribution function and the unconditional empirical distribution function.

Number of Points Missing	p_L	MAD	Δ MAD	Preferred Method		
				$p_r + \varepsilon(p_{r+1} - p_r)$	$0.5(p_r + p_{r+1})$	$p_{r+1} - \varepsilon(p_{r+1} - p_r)$
0	0	0.0610	0			
1	$p_1 + \varepsilon(p_2 - p_1)$	0.0694	0.0084			
1	$(p_1 + p_2)/2$	0.0627	0.0017		v	
1	$p_2 - \varepsilon(p_2 - p_1)$	0.0696	0.0086			
2	$p_2 + \varepsilon(p_3 - p_2)$	0.0777	0.0167			
2	$(p_2 + p_3)/2$	0.0709	0.0099			
2	$p_3 - \varepsilon(p_3 - p_2)$	0.0638	0.0028			v
3	$p_3 + \varepsilon(p_4 - p_3)$	0.0883	0.0273			
3	$(p_3 + p_4)/2$	0.0791	0.0181			
3	$p_4 - \varepsilon(p_4 - p_3)$	0.0696	0.0086			v
4	$p_4 + \varepsilon(p_5 - p_4)$	0.0958	0.0348			
4	$(p_4 + p_5)/2$	0.0846	0.0236			
4	$p_5 - \varepsilon(p_5 - p_4)$	0.0736	0.0126			v
5	$p_5 + \varepsilon(p_6 - p_5)$	0.0999	0.0389			
5	$(p_5 + p_6)/2$	0.0878	0.0268			
5	$p_6 - \varepsilon(p_6 - p_5)$	0.0749	0.0139			v
Average	$p_r + \varepsilon(p_{r+1} - p_r)$	0.0862	0.0252	Winning times		
Average	$(p_r + p_{r+1})/2$	0.0770	0.0160			
Average	$p_{r+1} - \varepsilon(p_{r+1} - p_r)$	0.0703	0.0093	0	1	4

Table 3.3. Experimental results for sample A2 with upper tail censored. Different lower censoring probabilities and number of missing points are applied; the MAD is calculated as the difference between the Weibull distribution function and the unconditional empirical distribution function.

Number of Points Missing	p_U	MAD	Δ MAD	Preferred Method		
				$p_{N-r+1} - \varepsilon(p_{N-r+1} - p_{N-r})$	$(p_{N-r} + p_{N-r+1})/2$	$p_{N-r} + \varepsilon(p_{N-r+1} - p_{N-r})$
0	0	0.0494	0			
1	$p_{20} - \varepsilon(p_{20} - p_{19})$	0.0569	0.0075			
1	$(p_{19} + p_{20})/2$	0.0519	0.0025		V	
1	$p_{19} + \varepsilon(p_{20} - p_{19})$	0.0693	0.0199			
2	$p_{19} - \varepsilon(p_{19} - p_{18})$	0.0634	0.0140			
2	$(p_{18} + p_{19})/2$	0.0557	0.0063			
2	$p_{18} + \varepsilon(p_{19} - p_{18})$	0.0544	0.0050			V
3	$p_{18} - \varepsilon(p_{18} - p_{17})$	0.0634	0.0140			
3	$(p_{17} + p_{18})/2$	0.0609	0.0115			
3	$p_{17} + \varepsilon(p_{18} - p_{17})$	0.0595	0.0101			V
4	$p_{17} - \varepsilon(p_{17} - p_{16})$	0.0775	0.0281			
4	$(p_{16} + p_{17})/2$	0.0674	0.0180			
4	$p_{16} + \varepsilon(p_{17} - p_{16})$	0.0569	0.0075			V
5	$p_{16} - \varepsilon(p_{16} - p_{15})$	0.0811	0.0317			
5	$(p_{15} + p_{16})/2$	0.0704	0.0210			
5	$p_{15} + \varepsilon(p_{16} - p_{15})$	0.0593	0.0099			V
Average	$p_{N-r+1} - \varepsilon(p_{N-r+1} - p_{N-r})$	0.0685	0.0191	Winning times		
Average	$(p_{N-r} + p_{N-r+1})/2$	0.0613	0.0119			
Average	$p_{N-r} + \varepsilon(p_{N-r+1} - p_{N-r})$	0.0599	0.0105	0	1	4

Table 3.4. Experimental results for sample A4 with upper tail censored. Different lower censoring probabilities and number of missing points are applied; the MAD is calculated as the difference between the Weibull distribution function and the unconditional empirical distribution function.

Number of Points Missing	p_U	MAD	Δ MAD	Preferred Method		
				$p_{N-r+1} - \varepsilon(p_{N-r+1} - p_{N-r})$	$(p_{N-r} + p_{N-r+1})/2$	$p_{N-r} + \varepsilon(p_{N-r+1} - p_{N-r})$
0	0	0.0610	0			
1	$p_{20} - \varepsilon(p_{20} - p_{19})$	0.0709	0.0099			
1	$(p_{19} + p_{20})/2$	0.0665	0.0055			
1	$p_{19} + \varepsilon(p_{20} - p_{19})$	0.0615	0.0004			V
2	$p_{19} - \varepsilon(p_{19} - p_{18})$	0.0837	0.0227			
2	$(p_{18} + p_{19})/2$	0.0777	0.0167			
2	$p_{18} + \varepsilon(p_{19} - p_{18})$	0.0699	0.0089			V
3	$p_{18} - \varepsilon(p_{18} - p_{17})$	0.0946	0.0336			
3	$(p_{17} + p_{18})/2$	0.0859	0.0248			
3	$p_{17} + \varepsilon(p_{18} - p_{17})$	0.0770	0.0159			V
4	$p_{17} - \varepsilon(p_{17} - p_{16})$	0.1036	0.0426			
4	$(p_{16} + p_{17})/2$	0.0935	0.0325			
4	$p_{16} + \varepsilon(p_{17} - p_{16})$	0.0819	0.0208			V
5	$p_{16} - \varepsilon(p_{16} - p_{15})$	0.1087	0.0477			
5	$(p_{15} + p_{16})/2$	0.0957	0.0347			
5	$p_{15} + \varepsilon(p_{16} - p_{15})$	0.0836	0.0225			V
Average	$p_{N-r+1} - \varepsilon(p_{N-r+1} - p_{N-r})$	0.0923	0.0313	Winning times		
Average	$(p_{N-r} + p_{N-r+1})/2$	0.0838	0.0228			
Average	$p_{N-r} + \varepsilon(p_{N-r+1} - p_{N-r})$	0.0748	0.0137		0	5

According to the MADs listed in Tables 3.1-3.4, we prefer the $p_{r+1} - \varepsilon(p_{r+1} - p_r)$ for lower tail censoring, and $p_{N-r} + \varepsilon(p_{N-r+1} - p_{N-r})$ for upper tail censoring as the lower/upper censoring probability, for the following three reasons. Let's define the choice above as the first choice, and the middle point as the second choice. First of all, the mean of MAD or Δ MAD of first choice wins three experiments, and it only loses one experiment, ranking in the middle. Second, the difference of means between the first choice and the second choice is obvious for three experiments, usually around 0.01. Finally, for each experiment we run five censored samples. The winning times show an overwhelming vote the first choice.

Experiment 2. The Tolerance of the Estimation Method to the Percentage of Missing Realizations

We conduct five parallel tests on samples N1, N2, N3, N4, and N5 using the distribution fitter. Table 4.16 to Table 4.20 show the experimental results based on each sample.

Ideally, we would see a “jump” of MAD at certain point in each of the experiments. However, the experimental result shows that the jump is not obvious, and sometimes the jump is erratic and we cannot decide from which point the distribution does not fit. The reason for this problem is probably the nature of the method itself, or the nature of the sample.

However, if we define the initial MAD to be the MAD when 0 point is missing, we can calculate an increment of MAD as the number of points missing rises. When we set the tolerance of increment to 0.005 approximately, we get Table 3.5 with the points where the distribution fit turns “bad”. In short, we call the point when the increment of MAD compared to the initial MAD exceeds approximately 0.005 a turning point.

The result shows that on average the turning occurs at about 7.2/62, i.e., where about 11.6% of realizations are censored. The results are analyzed based on Tables 4.16 – 4.20 with the rows called “CED Method”. As the MAD is not strictly increasing with the number of point missing, there is not an absolute turning point for a specific sample. For example, the upper censored sample experiment of N2 has a large MAD when number of missing points is 2. Though the MAD of this experiment decreases after number of missing points is 2, we choose the turning point at 2.

Table 3.5. Summary of increment of MAD for experiment 3.4.3.

Sample Name	Initial MAD	Type of Censoring	Turning Point	Turning Point MAD
N1	0.0566	Lower Tail Censored	2 point missing	0.0666
N1	0.0566	Upper Tail Censored	10 points missing	0.0610
N2	0.0569	Lower Tail Censored	10 points missing	0.0577
N2	0.0569	Upper Tail Censored	2 points missing	0.0642
N3	0.0798	Lower Tail Censored	8 points missing	0.0844
N3	0.0798	Upper Tail Censored	6 points missing	0.0839
N4	0.0647	Lower Tail Censored	8 points missing	0.0761
N4	0.0647	Upper Tail Censored	6 points missing	0.0734
N5	0.0388	Lower Tail Censored	10 point missing	0.0423
N5	0.0388	Upper Tail Censored	10 points missing	0.0397
Average			7.2 points missing	

3.5 Summary

This chapter introduces the MLE method and the CED method for constructing a distribution function from a censored sample. We define three types of censoring. For the MLE method, we rewrite Cohen (1965) method for the upper censoring case, and derive the method for the lower censoring case. For the CED method, we derive the process and explain with plots on how to apply it to each of the three censoring cases.

CHAPTER 4 : COMPARISON OF METHODS

4.1 Objective

In this chapter, we compare the two methods, MLE and CED, experimentally. We will apply the two methods simultaneously to 20 groups of samples. These samples are of different sizes, from different sources, and from different sample spaces. The experiments are intended to produce distribution functions fitted to the samples using each of the two methods. To test the goodness-of-fit, the experiments apply the K-S test and calculate the MAD. We want to draw a conclusion about which method performs better.

4.2 Design of Experiments

To verify the solution from MLE and to compare MLE with CED, we use four groups of samples: Artificial Samples A1, A2, A3, and A4; Random Samples R1, R2, R3, R4 and R5; River Stage Samples S1, S2, S3, S4, S5 and S6; National Weather Service Samples N1, N2, N3, N4, and N5 (Appendix). Each sample is used as an input for estimating the parameters of a Weibull distribution. Take A2 as an example. We apply each of the two methods six times. First time, we use A2 as a complete sample. Second time, we eliminate A2's largest realization to create a censored sample. Sixth time, we eliminate A2's five largest realizations. Table 4.2 displays the experimental results for A2.

The results of the experiment with 20 groups of samples, Tables 4.1 – 4.20, show in total 208 paired comparisons between MLE and CED. The column MAD is calculated as the difference between the estimated Weibull distribution and the unconditional empirical distribution function constructed from the complete sample.

In designing the experiment, according to Section 3.2.2, we assume type II censoring for the MLE, and thus use x_n to extend the censored sample from size n to size N .

The random samples R1 to R5 are drawn randomly from five Weibull distribution functions with quite different shapes. They correspond to these values of parameters: $\alpha = 1, \beta = 0.5$; $\alpha = 2, \beta = 1$; $\alpha = 4, \beta = 2$; $\alpha = 6, \beta = 3$; and $\alpha = 8, \beta = 4$. Tables 4.1 – 4.4 show the experimental comparisons of MLE and CED using artificial sample group. Tables 4.5 – 4.9 show the experimental comparisons of MLE and CED using random sample group. Tables 4.10 – 4.15 show the results for the river stage sample group. Tables 4.16 – 4.20 show the results for the National Weather Service sample group. The summary of Tables 4.1 – 4.20 is shown in Table 4.21.

4.3 Experimental Results

From all the tables below, we can draw the following conclusions: (i) Both methods have a trend that MAD increases with the number of points missing. (ii) The MLE method performs better when it is applied to upper tail censored sample than when it is applied to lower tail censored sample. (iii) The K-S test statistic orders the estimated distributions consistently with MAD. Where MAD is larger, K-S statistic is larger. (iv) MLE method for both tail censored sample is too complex. CED method deals with this problem in a simpler manner. (v) In total, CED performs better on both MAD and K-S statistics. In addition, the MLE method can only be applied to positive realizations as it requires a $\ln(\cdot)$ computation of each realization. However, CED method is not restricted in that way.

4.4 Summary

This chapter compares the MLE method and the CED method experimentally. It answers the question as to which method performs better. It also explains the restrictions of the MLE method.

Table 4.1. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample A1 (complete sample size N=5).

Number of Points Missing	Type of Censoring	CED Method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	7.8533	2.3230	0.0474	0.1916	0.2
1	Lower Tail	7.8131	2.1021	0.0586	0.1624	0.2
2	Lower Tail	7.7440	3.0897	0.1136	0.2598	0.2
		Average		0.0732	0.2046	
0	Upper Tail	7.8533	2.3230	0.0474	0.1916	0.2
1	Upper Tail	8.0918	2.0423	0.0836	0.1695	0.2
2	Upper Tail	9.3852	1.5436	0.2081	0.2837	0.2
		Average		0.1130	0.2149	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	7.3099	2.1155	0.1095	0.2095	0.2
1	Lower Tail	5.8317	1.1012	0.2480	0.3112	0.2
2	Lower Tail	7.6215	1.5730	0.0972	0.1912	0.2
		Average		0.1516	0.2373	
0	Upper Tail	7.3099	2.1155	0.1095	0.2095	0.2
1	Upper Tail	7.3383	1.9932	0.1079	0.2079	0.2
2	Upper Tail	7.7025	1.7854	0.0803	0.1788	0.2
		Average		0.0992	0.1987	

Table 4.2. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample A2 (complete sample size N=20).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	7.3886	2.9857	0.0494	0.0950	0.2
1	Lower Tail	7.4876	2.6157	0.0778	0.0820	0.2
2	Lower Tail	7.3859	2.9008	0.0572	0.0897	0.2
3	Lower Tail	7.3854	2.8437	0.0626	0.0860	0.2
4	Lower Tail	7.4070	2.8467	0.0607	0.0841	0.2
5	Lower Tail	7.4136	2.8246	0.0624	0.0820	0.2
Average				0.0617	0.0865	
0	Upper Tail	7.3886	2.9857	0.0494	0.0950	0.2
1	Upper Tail	7.4675	2.9203	0.0567	0.0831	0.2
2	Upper Tail	7.4305	2.9011	0.0539	0.0855	0.2
3	Upper Tail	7.4341	2.8465	0.0588	0.0815	0.2
4	Upper Tail	7.3363	2.9552	0.0559	0.0980	0.2
5	Upper Tail	7.3150	2.9745	0.0558	0.1012	0.2
Average				0.0551	0.0907	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	7.5471	2.3572	0.1016	0.1059	0.2
1	Lower Tail	7.7158	2.3625	0.0898	0.0940	0.2
2	Lower Tail	7.9892	2.5739	0.1207	0.1275	0.2
3	Lower Tail	8.1689	2.5747	0.1414	0.1482	0.2
4	Lower Tail	8.3366	2.4901	0.1569	0.1637	0.2
5	Lower Tail	8.5421	2.4303	0.1756	0.1824	0.2
Average				0.1310	0.1370	
0	Upper Tail	7.5471	2.3572	0.1016	0.1059	0.2
1	Upper Tail	7.4397	2.4999	0.0934	0.0977	0.2
2	Upper Tail	7.4210	2.5096	0.0937	0.0980	0.2
3	Upper Tail	7.4981	2.4213	0.0979	0.1021	0.2
4	Upper Tail	7.4218	2.4816	0.0967	0.1009	0.2
5	Upper Tail	7.4121	2.4817	0.0974	0.1016	0.2
Average				0.0967	0.1010	

Table 4.3. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample A3 (complete sample size N=5).

Number of Points Missing	Type of Censoring	CED Method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	5.0008	1.1311	0.0600	0.2042	0.2
1	Lower Tail	5.0937	0.9731	0.0739	0.1674	0.2
2	Lower Tail	4.8614	1.4108	0.1097	0.2559	0.2
		Average		0.0812	0.2092	
0	Upper Tail	5.0008	1.1311	0.0600	0.2042	0.2
1	Upper Tail	5.3158	0.9962	0.0885	0.1820	0.2
2	Upper Tail	7.6231	0.7107	0.1896	0.2620	0.2
		Average		0.1127	0.2161	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	4.4946	0.9541	0.1083	0.2083	0.2
1	Lower Tail	5.8317	1.1012	0.1233	0.2257	0.2
2	Lower Tail	7.6215	1.5730	0.2771	0.3400	0.2
		Average		0.1696	0.2580	
0	Upper Tail	4.4946	0.9541	0.1083	0.2083	0.2
1	Upper Tail	4.2645	0.9572	0.1268	0.2268	0.2
2	Upper Tail	4.9869	0.8285	0.0799	0.1799	0.2
		Average		0.1050	0.2050	

Table 4.4. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample A4 (complete sample size N=20).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	4.6794	1.0914	0.0610	0.0948	0.2
1	Lower Tail	4.7381	1.0277	0.0746	0.0870	0.2
2	Lower Tail	4.8079	1.0480	0.0670	0.0936	0.2
3	Lower Tail	4.8585	1.0616	0.0659	0.0997	0.2
4	Lower Tail	4.8981	1.0681	0.0698	0.1036	0.2
5	Lower Tail	4.9240	1.0618	0.0699	0.1037	0.2
Average				0.0680	0.0971	
0	Upper Tail	4.6794	1.0914	0.0610	0.0948	0.2
1	Upper Tail	4.7151	1.0951	0.0642	0.0979	0.2
2	Upper Tail	4.6141	1.1007	0.0661	0.0981	0.2
3	Upper Tail	4.5297	1.1090	0.0733	0.1058	0.2
4	Upper Tail	4.4839	1.1094	0.0775	0.1099	0.2
5	Upper Tail	4.4639	1.1075	0.0793	0.1117	0.2
Average				0.0702	0.1030	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	4.6955	1.1125	0.0666	0.1004	0.2
1	Lower Tail	4.8891	1.0934	0.0748	0.1086	0.2
2	Lower Tail	5.1369	1.0934	0.0894	0.1232	0.2
3	Lower Tail	5.4603	1.1358	0.1166	0.1504	0.2
4	Lower Tail	5.7120	1.0830	0.1197	0.1506	0.2
5	Lower Tail	6.0473	1.0558	0.1429	0.1590	0.2
Average				0.1017	0.1320	
0	Upper Tail	4.6955	1.1125	0.0666	0.1004	0.2
1	Upper Tail	4.6937	1.1038	0.0646	0.0984	0.2
2	Upper Tail	4.6680	1.1077	0.0637	0.0975	0.2
3	Upper Tail	4.7310	1.0836	0.0627	0.0965	0.2
4	Upper Tail	4.4960	1.1489	0.0760	0.1097	0.2
5	Upper Tail	4.5139	1.1400	0.0741	0.1079	0.2
Average				0.0680	0.1017	

Table 4.5. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample R1 (complete sample size N=20).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	1.2803	0.5876	0.0526	0.0983	0.2
1	Lower Tail	1.2822	0.5747	0.0574	0.0939	0.2
2	Lower Tail	1.3129	0.5856	0.0552	0.0942	0.2
3	Lower Tail	1.3390	0.5922	0.0609	0.0999	0.2
4	Lower Tail	1.3592	0.5922	0.0632	0.1022	0.2
5	Lower Tail	1.3718	0.5866	0.0621	0.1011	0.2
Average				0.0586	0.0983	
0	Upper Tail	1.2803	0.5876	0.0526	0.0983	0.2
1	Upper Tail	1.3286	0.5793	0.0542	0.0932	0.2
2	Upper Tail	1.3049	0.5776	0.0551	0.0924	0.2
3	Upper Tail	1.3344	0.5669	0.0580	0.0883	0.2
4	Upper Tail	1.1375	0.6043	0.0733	0.1190	0.2
5	Upper Tail	1.1187	0.6064	0.0759	0.1217	0.2
Average				0.0615	0.1022	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	1.5351	0.5509	0.0874	0.1008	0.2
1	Lower Tail	1.6940	0.5550	0.1065	0.1160	0.2
2	Lower Tail	1.8403	0.5421	0.1242	0.1272	0.2
3	Lower Tail	2.0065	0.5217	0.1423	0.1453	0.2
4	Lower Tail	2.2591	0.5251	0.1651	0.1680	0.2
5	Lower Tail	2.5815	0.5463	0.1918	0.1948	0.2
Average				0.1362	0.1420	
0	Upper Tail	1.5351	0.5509	0.0874	0.1008	0.2
1	Upper Tail	1.4483	0.5822	0.0710	0.1069	0.2
2	Upper Tail	1.4343	0.5839	0.0687	0.1063	0.2
3	Upper Tail	1.5128	0.5598	0.0832	0.1029	0.2
4	Upper Tail	1.2022	0.6519	0.0805	0.1263	0.2
5	Upper Tail	1.3012	0.6185	0.0679	0.1069	0.2
Average				0.0765	0.1084	

Table 4.6. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample R2 (complete sample size N=20).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	1.9013	1.7665	0.0831	0.1288	0.2
1	Lower Tail	1.8983	1.7656	0.0841	0.1292	0.2
2	Lower Tail	1.9126	1.7869	0.0866	0.1289	0.2
3	Lower Tail	1.9221	1.7980	0.0897	0.1284	0.2
4	Lower Tail	1.9276	1.8006	0.0916	0.1277	0.2
5	Lower Tail	1.9307	1.7689	0.0930	0.1238	0.2
Average				0.0880	0.1278	
0	Upper Tail	1.9013	1.7665	0.0831	0.1288	0.2
1	Upper Tail	1.9153	1.7139	0.0886	0.1207	0.2
2	Upper Tail	1.9028	1.7010	0.0891	0.1215	0.2
3	Upper Tail	1.9003	1.6372	0.0962	0.1148	0.2
4	Upper Tail	1.8721	1.8231	0.0937	0.1394	0.2
5	Upper Tail	1.8846	1.7775	0.0871	0.1328	0.2
Average				0.0896	0.1263	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	1.9518	1.4696	0.1024	0.1109	0.2
1	Lower Tail	2.0062	1.4268	0.1172	0.1211	0.2
2	Lower Tail	2.0821	1.4216	0.1366	0.1405	0.2
3	Lower Tail	2.1902	1.5076	0.1659	0.1698	0.2
4	Lower Tail	2.2706	1.4633	0.1835	0.1875	0.2
5	Lower Tail	2.3712	1.4512	0.2056	0.2095	0.2
Average				0.1519	0.1566	
0	Upper Tail	1.9518	1.4696	0.1024	0.1109	0.2
1	Upper Tail	1.9691	1.4266	0.1074	0.1114	0.2
2	Upper Tail	1.9693	1.4193	0.1075	0.1122	0.2
3	Upper Tail	1.9910	1.3864	0.1132	0.1172	0.2
4	Upper Tail	1.8353	1.5877	0.1198	0.1308	0.2
5	Upper Tail	1.7436	1.7215	0.1378	0.1498	0.2
Average				0.1147	0.1221	

Table 4.7. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample R3 (complete sample size N=20).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	3.8627	2.3390	0.0632	0.1076	0.2
1	Lower Tail	3.8491	2.2789	0.0711	0.1050	0.2
2	Lower Tail	3.8628	2.2466	0.0736	0.1002	0.2
3	Lower Tail	3.8650	2.1594	0.0839	0.0951	0.2
4	Lower Tail	3.8797	2.1575	0.0827	0.0945	0.2
5	Lower Tail	3.8924	2.1642	0.0807	0.0932	0.2
		Average		0.0759	0.0993	
0	Upper Tail	3.8627	2.3390	0.0632	0.1076	0.2
1	Upper Tail	3.8777	2.3276	0.0655	0.1042	0.2
2	Upper Tail	3.8489	2.3316	0.0652	0.1093	0.2
3	Upper Tail	3.8439	2.3100	0.0681	0.1084	0.2
4	Upper Tail	3.8604	2.2624	0.0720	0.1018	0.2
5	Upper Tail	3.7191	2.4320	0.0915	0.1386	0.2
		Average		0.0709	0.1117	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	4.0305	2.3143	0.0974	0.1233	0.2
1	Lower Tail	4.0994	2.2440	0.1083	0.1242	0.2
2	Lower Tail	4.2069	2.2722	0.1302	0.1412	0.2
3	Lower Tail	4.2949	2.1868	0.1415	0.1525	0.2
4	Lower Tail	4.4186	2.2040	0.1639	0.1749	0.2
5	Lower Tail	4.5407	2.1232	0.1779	0.1889	0.2
		Average		0.1365	0.1508	
0	Upper Tail	4.0305	2.3143	0.0974	0.1233	0.2
1	Upper Tail	4.0352	2.2865	0.0973	0.1207	0.2
2	Upper Tail	4.0508	2.2486	0.0989	0.1184	0.2
3	Upper Tail	4.0304	2.2723	0.0958	0.1184	0.2
4	Upper Tail	4.0353	2.2580	0.0962	0.1174	0.2
5	Upper Tail	3.8373	2.5398	0.0808	0.1274	0.2
		Average		0.0944	0.1209	

Table 4.8. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample R4 (complete sample size N=20).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	6.4523	3.4669	0.0477	0.0937	0.2
1	Lower Tail	6.4456	3.4373	0.0505	0.0907	0.2
2	Lower Tail	6.4699	3.4817	0.0505	0.0965	0.2
3	Lower Tail	6.4904	3.4774	0.0544	0.0982	0.2
4	Lower Tail	6.5030	3.4296	0.0582	0.0957	0.2
5	Lower Tail	6.5055	3.3046	0.0624	0.0859	0.2
		Average		0.0540	0.0935	
0	Upper Tail	6.4523	3.4669	0.0477	0.0937	0.2
1	Upper Tail	6.4756	3.4001	0.0541	0.0907	0.2
2	Upper Tail	6.4561	3.3609	0.0532	0.0857	0.2
3	Upper Tail	6.4556	3.2431	0.0602	0.0777	0.2
4	Upper Tail	6.4513	3.3127	0.0568	0.0824	0.2
5	Upper Tail	6.4949	3.1986	0.0638	0.0761	0.2
		Average		0.0560	0.0844	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	6.4871	3.3401	0.0580	0.0870	0.2
1	Lower Tail	6.5734	3.2755	0.0753	0.0900	0.2
2	Lower Tail	6.6827	3.2750	0.0947	0.0998	0.2
3	Lower Tail	6.8207	3.4050	0.1172	0.1217	0.2
4	Lower Tail	6.9279	3.2789	0.1377	0.1406	0.2
5	Lower Tail	7.0612	3.2219	0.1607	0.1636	0.2
		Average		0.1073	0.1171	
0	Upper Tail	6.4871	3.3401	0.0580	0.0870	0.2
1	Upper Tail	6.5067	3.2583	0.0640	0.0822	0.2
2	Upper Tail	6.5276	3.1945	0.0696	0.0789	0.2
3	Upper Tail	6.5487	3.1408	0.0747	0.0777	0.2
4	Upper Tail	6.3237	3.5708	0.0670	0.1127	0.2
5	Upper Tail	6.3292	3.5477	0.0651	0.1108	0.2
		Average		0.0664	0.0916	

Table 4.9. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample R5 (complete sample size N=20).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	7.7031	4.8667	0.0461	0.0840	0.2
1	Lower Tail	7.6949	4.8034	0.0503	0.0794	0.2
2	Lower Tail	7.7167	4.8243	0.0489	0.0840	0.2
3	Lower Tail	7.7316	4.8586	0.0526	0.0882	0.2
4	Lower Tail	7.7463	4.8781	0.0561	0.0916	0.2
5	Lower Tail	7.7620	4.7999	0.0589	0.0900	0.2
		Average		0.0522	0.0862	
0	Upper Tail	7.7031	4.8667	0.0461	0.0840	0.2
1	Upper Tail	7.7566	4.6080	0.0557	0.0789	0.2
2	Upper Tail	7.6089	5.0930	0.0541	0.0901	0.2
3	Upper Tail	7.6089	5.0260	0.0530	0.0887	0.2
4	Upper Tail	7.5924	5.0259	0.0547	0.0925	0.2
5	Upper Tail	7.6013	4.9571	0.0550	0.0890	0.2
		Average		0.0531	0.0872	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	7.6838	4.0654	0.0958	0.0988	0.2
1	Lower Tail	7.8571	4.5785	0.0768	0.0926	0.2
2	Lower Tail	7.9511	4.5836	0.0965	0.1075	0.2
3	Lower Tail	8.0599	4.6683	0.1206	0.1316	0.2
4	Lower Tail	8.1679	4.7002	0.1432	0.1542	0.2
5	Lower Tail	8.2712	4.5410	0.1583	0.1693	0.2
		Average		0.1152	0.1257	
0	Upper Tail	7.6838	4.0654	0.0958	0.0988	0.2
1	Upper Tail	7.7405	3.8419	0.1031	0.1060	0.2
2	Upper Tail	7.6321	4.1528	0.0973	0.1002	0.2
3	Upper Tail	7.6746	4.0227	0.0998	0.1027	0.2
4	Upper Tail	7.6818	3.9913	0.1008	0.1038	0.2
5	Upper Tail	7.6467	4.0502	0.1017	0.1047	0.2
		Average		0.0998	0.1027	

Table 4.10. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample S1 (complete sample size N=50).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	2.3094	1.1830	0.1190	0.1373	0.2
2	Lower Tail	2.3006	1.1759	0.1206	0.1348	0.2
4	Lower Tail	2.3149	1.1894	0.1207	0.1393	0.2
6	Lower Tail	2.3244	1.1974	0.1235	0.1421	0.2
8	Lower Tail	2.3328	1.2064	0.1263	0.1449	0.2
10	Lower Tail	2.3373	1.2316	0.1323	0.1509	0.2
Average				0.1237	0.1416	
0	Upper Tail	2.3094	1.1830	0.1190	0.1373	0.2
2	Upper Tail	2.3287	1.1731	0.1232	0.1370	0.2
4	Upper Tail	2.3153	1.1653	0.1213	0.1339	0.2
6	Upper Tail	2.3108	1.1527	0.1215	0.1305	0.2
8	Upper Tail	2.3185	1.1313	0.1244	0.1268	0.2
10	Upper Tail	2.3417	1.0984	0.1308	0.1291	0.2
Average				0.1233	0.1324	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	1.9829	1.3648	0.1845	0.1914	0.05
2	Lower Tail	2.0292	1.3284	0.1733	0.1801	0.05
4	Lower Tail	2.0825	1.2917	0.1613	0.1682	0.1
6	Lower Tail	2.1509	1.2769	0.1464	0.1533	0.15
8	Lower Tail	2.2260	1.2581	0.1312	0.1453	0.2
10	Lower Tail	2.3341	1.3598	0.1584	0.1771	0.05
Average				0.1592	0.1692	
0	Upper Tail	1.9829	1.3648	0.1845	0.1914	0.05
2	Upper Tail	1.9889	1.3458	0.1830	0.1899	0.05
4	Upper Tail	2.0012	1.3252	0.1801	0.1869	0.05
6	Upper Tail	2.0163	1.3033	0.1765	0.1834	0.05
8	Upper Tail	2.0097	1.3077	0.1781	0.1850	0.05
10	Upper Tail	2.0326	1.2854	0.1728	0.1797	0.05
Average				0.1792	0.1861	

Table 4.11. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample S2 (complete sample size N=46).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	1.8217	1.7357	0.1213	0.1442	0.2
2	Lower Tail	1.8165	1.7303	0.1231	0.1443	0.2
4	Lower Tail	1.8216	1.7437	0.1219	0.1449	0.2
6	Lower Tail	1.8290	1.7478	0.1231	0.1444	0.2
8	Lower Tail	1.8316	1.7514	0.1240	0.1444	0.2
10	Lower Tail	1.8315	1.7441	0.1240	0.1439	0.2
Average				0.1229	0.1444	
0	Upper Tail	1.8217	1.7357	0.1213	0.1442	0.2
2	Upper Tail	1.8377	1.6824	0.1284	0.1383	0.2
4	Upper Tail	1.8291	1.6844	0.1255	0.1394	0.2
6	Upper Tail	1.8279	1.6559	0.1270	0.1371	0.2
8	Upper Tail	1.7773	2.0244	0.1405	0.1634	0.15
10	Upper Tail	1.7840	1.9935	0.1390	0.1619	0.15
Average				0.1303	0.1474	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	2.1820	1.2216	0.2237	0.2220	0.01
2	Lower Tail	2.2468	1.1972	0.2360	0.2343	0.01
4	Lower Tail	2.3146	1.1619	0.2474	0.2457	<0.01
6	Lower Tail	2.4033	1.1421	0.2622	0.2604	<0.01
8	Lower Tail	2.5158	1.1467	0.2813	0.2796	<0.01
10	Lower Tail	2.6801	1.2525	0.3171	0.3154	<0.01
Average				0.2613	0.2596	
0	Upper Tail	2.1820	1.2216	0.2237	0.2220	0.01
2	Upper Tail	2.2039	1.1889	0.2273	0.2256	0.01
4	Upper Tail	2.1589	1.2406	0.2193	0.2176	0.01
6	Upper Tail	2.0624	1.3533	0.1987	0.1970	0.05
8	Upper Tail	1.6804	2.0869	0.1463	0.1693	0.1
10	Upper Tail	1.6915	2.0505	0.1451	0.1681	0.1
Average				0.1934	0.1999	

Table 4.12. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample S3 (complete sample size N=44).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	3.2192	1.0889	0.0526	0.0687	0.2
2	Lower Tail	3.2928	1.0357	0.0592	0.0749	0.2
4	Lower Tail	3.3200	1.0483	0.0564	0.0721	0.2
6	Lower Tail	3.3289	1.0634	0.0568	0.0735	0.2
8	Lower Tail	3.3494	1.0649	0.0590	0.0757	0.2
10	Lower Tail	3.3619	1.0643	0.0600	0.0767	0.2
Average				0.0573	0.0736	
0	Upper Tail	3.2192	1.0889	0.0526	0.0687	0.2
2	Upper Tail	3.2328	1.0885	0.0532	0.0699	0.2
4	Upper Tail	3.1942	1.0900	0.0554	0.0680	0.2
6	Upper Tail	3.1600	1.0932	0.0590	0.0683	0.2
8	Upper Tail	3.1306	1.0924	0.0627	0.0702	0.2
10	Upper Tail	3.1102	1.0998	0.0643	0.0718	0.2
Average				0.0579	0.0695	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	3.2803	1.1243	0.0653	0.0820	0.2
2	Lower Tail	3.3878	1.0949	0.0690	0.0857	0.2
4	Lower Tail	3.5108	1.0620	0.0724	0.0891	0.2
6	Lower Tail	3.6763	1.0522	0.0865	0.1001	0.2
8	Lower Tail	3.8734	1.0569	0.1064	0.1158	0.2
10	Lower Tail	4.1047	1.0871	0.1285	0.1387	0.2
Average				0.0880	0.1019	
0	Upper Tail	3.2803	1.1243	0.0653	0.0820	0.2
2	Upper Tail	3.2995	1.1025	0.0625	0.0792	0.2
4	Upper Tail	3.2407	1.1377	0.0642	0.0810	0.2
6	Upper Tail	3.2054	1.1576	0.0648	0.0815	0.2
8	Upper Tail	3.1809	1.1675	0.0643	0.0810	0.2
10	Upper Tail	3.1311	1.1885	0.0633	0.0800	0.2
Average				0.0641	0.0808	

Table 4.13. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample S4 (complete sample size N=46).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	2.7175	1.4452	0.0667	0.0863	0.2
2	Lower Tail	2.7256	1.4128	0.0719	0.0814	0.2
4	Lower Tail	2.7557	1.4077	0.0700	0.0841	0.2
6	Lower Tail	2.8009	1.3765	0.0720	0.0837	0.2
8	Lower Tail	2.8887	1.2950	0.0808	0.0805	0.2
10	Lower Tail	2.9311	1.2572	0.0853	0.0850	0.2
Average				0.0745	0.0835	
0	Upper Tail	2.7175	1.4452	0.0667	0.0863	0.2
2	Upper Tail	2.7207	1.4484	0.0677	0.0873	0.2
4	Upper Tail	2.6948	1.4547	0.0697	0.0851	0.2
6	Upper Tail	2.6714	1.4613	0.0741	0.0857	0.2
8	Upper Tail	2.6585	1.4634	0.0767	0.0883	0.2
10	Upper Tail	2.6411	1.6496	0.0916	0.1112	0.2
Average				0.0744	0.0907	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	2.9975	1.2626	0.0788	0.0827	0.2
2	Lower Tail	3.1062	1.2642	0.0780	0.0937	0.2
4	Lower Tail	3.2124	1.2498	0.0940	0.1004	0.2
6	Lower Tail	3.3832	1.3164	0.1158	0.1289	0.2
8	Lower Tail	3.5041	1.2928	0.1331	0.1335	0.2
10	Lower Tail	3.6387	1.2603	0.1510	0.1491	0.2
Average				0.1085	0.1147	
0	Upper Tail	2.9975	1.2626	0.0788	0.0827	0.2
2	Upper Tail	2.9528	1.3141	0.0716	0.0885	0.2
4	Upper Tail	2.9156	1.3485	0.0717	0.0913	0.2
6	Upper Tail	2.8792	1.3824	0.0743	0.0939	0.2
8	Upper Tail	2.8869	1.3725	0.0732	0.0928	0.2
10	Upper Tail	2.7944	1.4409	0.0754	0.0950	0.2
Average				0.0742	0.0907	

Table 4.14. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample S5 (complete sample size N=40).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	4.8955	2.2137	0.0614	0.0861	0.2
2	Lower Tail	4.8827	2.2040	0.0640	0.0867	0.2
4	Lower Tail	4.9079	2.2188	0.0633	0.0851	0.2
6	Lower Tail	4.9253	2.2158	0.0655	0.0829	0.2
8	Lower Tail	4.9407	2.1938	0.0658	0.0793	0.2
10	Lower Tail	4.9483	2.1443	0.0635	0.0759	0.2
Average				0.0639	0.0827	
0	Upper Tail	4.8955	2.2137	0.0614	0.0861	0.2
2	Upper Tail	4.9224	2.1891	0.0628	0.0809	0.2
4	Upper Tail	4.8986	2.1839	0.0645	0.0831	0.2
6	Upper Tail	4.8952	2.1630	0.0673	0.0817	0.2
8	Upper Tail	4.8447	2.2040	0.0685	0.0909	0.2
10	Upper Tail	4.8358	2.2007	0.0699	0.0916	0.2
Average				0.0657	0.0857	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	5.1347	2.0003	0.0732	0.0867	0.2
2	Lower Tail	5.2452	1.9565	0.0819	0.0953	0.2
4	Lower Tail	5.4137	2.0029	0.1058	0.1192	0.2
6	Lower Tail	5.5671	1.9947	0.1212	0.1346	0.2
8	Lower Tail	5.7326	1.9749	0.1355	0.1489	0.2
10	Lower Tail	5.9308	2.0077	0.1580	0.1714	0.15
Average				0.1126	0.1260	
0	Upper Tail	5.1347	2.0003	0.0732	0.0867	0.2
2	Upper Tail	5.0344	2.1563	0.0753	0.0887	0.2
4	Upper Tail	5.0771	2.0931	0.0750	0.0884	0.2
6	Upper Tail	5.0970	2.0674	0.0751	0.0885	0.2
8	Upper Tail	5.0249	2.1319	0.0717	0.0852	0.2
10	Upper Tail	5.0114	2.1410	0.0708	0.0842	0.2
Average				0.0735	0.0870	

Table 4.15. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample S6 (complete sample size N=34).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	3.6179	2.1340	0.0914	0.1216	0.2
2	Lower Tail	3.6129	2.0246	0.1037	0.1163	0.2
4	Lower Tail	3.6350	1.9627	0.1098	0.1224	0.2
6	Lower Tail	3.6307	1.8117	0.1306	0.1433	0.2
8	Lower Tail	3.6295	1.7435	0.1408	0.1535	0.2
10	Lower Tail	3.6220	1.6476	0.1564	0.1691	0.2
Average				0.1221	0.1377	
0	Upper Tail	3.6179	2.1340	0.0914	0.1216	0.2
2	Upper Tail	3.6324	2.1110	0.0921	0.1172	0.2
4	Upper Tail	3.6068	2.1194	0.0929	0.1221	0.2
6	Upper Tail	3.6214	2.0755	0.0969	0.1159	0.2
8	Upper Tail	3.6237	2.0656	0.0979	0.1146	0.2
10	Upper Tail	3.3469	2.5482	0.1671	0.1973	0.1
Average				0.1064	0.1315	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	3.7870	2.1625	0.1253	0.1363	0.2
2	Lower Tail	3.8861	2.1364	0.1403	0.1514	0.2
4	Lower Tail	3.9835	2.0584	0.1487	0.1598	0.2
6	Lower Tail	4.1083	2.0077	0.1656	0.1735	0.2
8	Lower Tail	4.2582	1.9751	0.1870	0.1915	0.15
10	Lower Tail	4.4325	1.9376	0.2093	0.2138	0.05
Average				0.1627	0.1711	
0	Upper Tail	3.7870	2.1625	0.1253	0.1363	0.2
2	Upper Tail	3.8203	2.0847	0.1242	0.1353	0.2
4	Upper Tail	3.8343	2.0541	0.1238	0.1349	0.2
6	Upper Tail	3.8956	1.9721	0.1268	0.1373	0.2
8	Upper Tail	3.8612	2.0085	0.1241	0.1351	0.2
10	Upper Tail	3.4529	2.5404	0.1511	0.1813	0.2
Average				0.1292	0.1434	

Table 4.16. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample N1 (complete sample size N=62).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	5.3255	5.8195	0.0566	0.0747	0.2
2	Lower Tail	5.3238	5.5101	0.0666	0.0785	0.2
4	Lower Tail	5.3305	5.4816	0.0671	0.0790	0.2
6	Lower Tail	5.3273	5.2093	0.0773	0.0892	0.2
8	Lower Tail	5.3315	5.2310	0.0761	0.0880	0.2
10	Lower Tail	5.3349	5.2590	0.0748	0.0866	0.2
		Average		0.0698	0.0826	
0	Upper Tail	5.3255	5.8195	0.0566	0.0747	0.2
2	Upper Tail	5.3349	5.7895	0.0582	0.0709	0.2
4	Upper Tail	5.3284	5.7917	0.0572	0.0729	0.2
6	Upper Tail	5.3263	5.7791	0.0577	0.0732	0.2
8	Upper Tail	5.3209	5.7872	0.0579	0.0751	0.2
10	Upper Tail	5.3150	5.7018	0.0610	0.0742	0.2
		Average		0.0581	0.0735	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	5.4147	5.3951	0.0701	0.0771	0.2
2	Lower Tail	5.4402	5.3029	0.0792	0.0799	0.2
4	Lower Tail	5.4706	5.2423	0.0897	0.0891	0.2
6	Lower Tail	5.4996	5.1282	0.0990	0.0984	0.2
8	Lower Tail	5.5342	5.0519	0.1101	0.1094	0.2
10	Lower Tail	5.5737	5.0140	0.1227	0.1221	0.2
		Average		0.0951	0.0960	
0	Upper Tail	5.4147	5.3951	0.0701	0.0771	0.2
2	Upper Tail	5.4193	5.3278	0.0717	0.0776	0.2
4	Upper Tail	5.4220	5.2907	0.0726	0.0787	0.2
6	Upper Tail	5.4036	5.4492	0.0661	0.0764	0.2
8	Upper Tail	5.3466	5.9221	0.0672	0.0799	0.2
10	Upper Tail	5.3646	5.7754	0.0665	0.0792	0.2
		Average		0.0690	0.0782	

Table 4.17. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample N2 (complete sample size N=62).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	5.2844	6.7304	0.0569	0.0722	0.2
2	Lower Tail	5.2770	6.6795	0.0608	0.0723	0.2
4	Lower Tail	5.2820	6.7082	0.0583	0.0721	0.2
6	Lower Tail	5.2837	6.7139	0.0577	0.0719	0.2
8	Lower Tail	5.2850	6.7107	0.0575	0.0716	0.2
10	Lower Tail	5.2871	6.7498	0.0577	0.0722	0.2
Average				0.0582	0.0721	
0	Upper Tail	5.2844	6.7304	0.0569	0.0722	0.2
2	Upper Tail	5.2975	6.6347	0.0642	0.0677	0.2
4	Upper Tail	5.2927	6.6056	0.0626	0.0679	0.2
6	Upper Tail	5.2907	6.5603	0.0625	0.0671	0.2
8	Upper Tail	5.2824	6.6384	0.0610	0.0704	0.2
10	Upper Tail	5.2828	6.6420	0.0608	0.0704	0.2
Average				0.0613	0.0693	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	5.4037	5.3607	0.1236	0.1212	0.2
2	Lower Tail	5.4385	5.3807	0.1360	0.1336	0.2
4	Lower Tail	5.4767	5.4306	0.1495	0.1471	0.1
6	Lower Tail	5.5089	5.3827	0.1613	0.1589	0.05
8	Lower Tail	5.5460	5.3791	0.1746	0.1722	0.05
10	Lower Tail	5.5845	5.3692	0.1883	0.1859	0.01
Average				0.1556	0.1532	
0	Upper Tail	5.4037	5.3607	0.1236	0.1212	0.2
2	Upper Tail	5.3990	5.3993	0.1215	0.1191	0.2
4	Upper Tail	5.3571	5.8617	0.1003	0.0979	0.2
6	Upper Tail	5.3459	5.9719	0.0944	0.0920	0.2
8	Upper Tail	5.3263	6.1635	0.0838	0.0814	0.2
10	Upper Tail	5.3206	6.2174	0.0806	0.0782	0.2
Average				0.1007	0.0983	

Table 4.18. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample N3 (complete sample size N=62).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	5.4705	6.2108	0.0798	0.0962	0.2
2	Lower Tail	5.4661	6.1685	0.0818	0.0940	0.2
4	Lower Tail	5.4715	6.2067	0.0799	0.0962	0.2
6	Lower Tail	5.4759	6.2591	0.0827	0.0995	0.2
8	Lower Tail	5.4787	6.2849	0.0844	0.1012	0.2
10	Lower Tail	5.4815	6.2955	0.0855	0.1023	0.2
Average				0.0824	0.0982	
0	Upper Tail	5.4705	6.2108	0.0798	0.0962	0.2
2	Upper Tail	5.4804	6.1527	0.0845	0.0958	0.2
4	Upper Tail	5.4753	6.1270	0.0833	0.0936	0.2
6	Upper Tail	5.4732	6.0778	0.0839	0.0909	0.2
8	Upper Tail	5.4879	5.8405	0.0951	0.0947	0.2
10	Upper Tail	5.5010	5.6458	0.1044	0.1041	0.2
Average				0.0885	0.0959	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	5.3953	6.3163	0.1114	0.1251	0.2
2	Lower Tail	5.4165	6.1697	0.1024	0.1141	0.2
4	Lower Tail	5.4424	6.0772	0.0918	0.1024	0.2
6	Lower Tail	5.4729	6.0487	0.0845	0.0899	0.2
8	Lower Tail	5.5151	6.3298	0.0943	0.1111	0.2
10	Lower Tail	5.5554	6.5938	0.1137	0.1305	0.2
Average				0.0997	0.1122	
0	Upper Tail	5.3953	6.3163	0.1114	0.1251	0.2
2	Upper Tail	5.4068	6.1215	0.1065	0.1173	0.2
4	Upper Tail	5.4011	6.1842	0.1089	0.1206	0.2
6	Upper Tail	5.4091	6.0906	0.1055	0.1159	0.2
8	Upper Tail	5.4175	6.0152	0.1021	0.1115	0.2
10	Upper Tail	5.4350	5.8748	0.0953	0.1046	0.2
Average				0.1050	0.1158	

Table 4.19. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample N4 (complete sample size N=62).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	5.2567	6.3276	0.0647	0.0874	0.2
2	Lower Tail	5.2563	6.2143	0.0694	0.0844	0.2
4	Lower Tail	5.2613	6.2253	0.0679	0.0837	0.2
6	Lower Tail	5.2672	6.1897	0.0683	0.0814	0.2
8	Lower Tail	5.2772	5.9666	0.0761	0.0774	0.2
10	Lower Tail	5.2844	5.8582	0.0796	0.0809	0.2
Average				0.0710	0.0825	
0	Upper Tail	5.2567	6.3276	0.0647	0.0874	0.2
2	Upper Tail	5.2690	6.2586	0.0670	0.0830	0.2
4	Upper Tail	5.2795	6.1518	0.0676	0.0778	0.2
6	Upper Tail	5.2818	6.3113	0.0734	0.0827	0.2
8	Upper Tail	5.2769	6.1067	0.0700	0.0770	0.2
10	Upper Tail	5.2492	6.2862	0.0677	0.0878	0.2
Average				0.0684	0.0826	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	5.4088	5.4381	0.0846	0.0939	0.2
2	Lower Tail	5.4422	5.4469	0.0949	0.1043	0.2
4	Lower Tail	5.4784	5.4778	0.1068	0.1162	0.2
6	Lower Tail	5.5119	5.4521	0.1153	0.1247	0.2
8	Lower Tail	5.5482	5.4428	0.1250	0.1344	0.2
10	Lower Tail	5.5797	5.3073	0.1276	0.1369	0.15
Average				0.1090	0.1184	
0	Upper Tail	5.4088	5.4381	0.0846	0.0939	0.2
2	Upper Tail	5.4112	5.3982	0.0837	0.0931	0.2
4	Upper Tail	5.3928	5.5822	0.0852	0.0946	0.2
6	Upper Tail	5.3584	5.8877	0.0855	0.0949	0.2
8	Upper Tail	5.3840	5.6635	0.0855	0.0949	0.2
10	Upper Tail	5.3384	6.0253	0.0837	0.0931	0.2
Average				0.0847	0.0941	

Table 4.20. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method applied to censored sample N5 (complete sample size N=62).

Number of Points Missing	Type of Censoring	CED method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	5.3795	5.5032	0.0388	0.0576	0.2
2	Lower Tail	5.4008	5.2781	0.0440	0.0559	0.2
4	Lower Tail	5.4107	5.2912	0.0431	0.0550	0.2
6	Lower Tail	5.4106	5.3261	0.0421	0.0540	0.2
8	Lower Tail	5.4171	5.3142	0.0421	0.0540	0.2
10	Lower Tail	5.4181	5.3045	0.0423	0.0542	0.2
Average				0.0421	0.0551	
0	Upper Tail	5.3795	5.5032	0.0388	0.0576	0.2
2	Upper Tail	5.3922	5.4873	0.0399	0.0584	0.2
4	Upper Tail	5.3815	5.5021	0.0388	0.0573	0.2
6	Upper Tail	5.3754	5.5090	0.0398	0.0590	0.2
8	Upper Tail	5.3733	5.5048	0.0404	0.0596	0.2
10	Upper Tail	5.3917	5.4453	0.0397	0.0563	0.2
Average				0.0395	0.0580	
Number of Points Missing	Type of Censoring	MLE method			K-S stat	Significance level
		α	β	MAD		
0	Lower Tail	5.4073	5.7943	0.0566	0.0751	0.2
2	Lower Tail	5.4315	5.6924	0.0560	0.0744	0.2
4	Lower Tail	5.4583	5.5962	0.0558	0.0743	0.2
6	Lower Tail	5.4892	5.5311	0.0575	0.0762	0.2
8	Lower Tail	5.5306	5.6429	0.0693	0.0926	0.2
10	Lower Tail	5.5662	5.6048	0.0740	0.1052	0.2
Average				0.0615	0.0830	
0	Upper Tail	5.4073	5.7943	0.0566	0.0751	0.2
2	Upper Tail	5.4090	5.7453	0.0547	0.0732	0.2
4	Upper Tail	5.4095	5.7294	0.0541	0.0726	0.2
6	Upper Tail	5.4190	5.6427	0.0517	0.0701	0.2
8	Upper Tail	5.4259	5.5877	0.0502	0.0687	0.2
10	Upper Tail	5.4568	5.3704	0.0448	0.0633	0.2
Average				0.0520	0.0705	

Table 4.21. Conditional Empirical Distribution (CED) method and Maximum Likelihood Estimation (MLE) method comparison summary.

Type of Censoring	Number of Samples	Method	Number of times better	Average MAD	Average K-S stat
None	20	CED	20	0.0660	0.1065
		MLE	0	0.1012	0.1219
		Difference	20	-0.0352	-0.0154
Lower Tail	94	CED	93	0.0768	0.1018
		MLE	1	0.1351	0.1473
		Difference	92	-0.0583	-0.0455
Upper Tail	94	CED	81	0.0781	0.1023
		MLE	13	0.0961	0.1142
		Difference	68	-0.0181	-0.0119
Sum of Lower and Upper Tail	188	CED	174	0.0774	0.1020
		MLE	14	0.1156	0.1307
		Difference	160	-0.0382	-0.0287

CHAPTER 5 : THE PARETO DISTRIBUTION

5.1 Introduction

Pareto is a distribution with two parameters from the catalog of distributions to be used in probabilistic forecasting. The reason we are curious about it is that when we want to fit the Pareto distribution to a sample using distribution fitter software, the optimization process sometimes iterates much beyond a satisfactory solution point. The scale parameter can increase from the initial value of 10 to the final optimized value of 7,000, with the MAD improvement less than 0.01. This optimization process is described in Section 2.4.

Figure 5.1 shows a 3 -dimensional graph of the MAD surface for the Pareto distributions. It shows the surface of MAD with scale parameter varying from 3,000 to 3,600, and shape parameter varying from 500 to 800. There is a valley on the surface which represents the direction MAD descends fastest. However, the slope of the valley is almost 0 in the region of the figure, which means the decrease of MAD is subtle as the values of the parameters increase. In this chapter, we discuss more specific properties of this distribution, and give a feasible explanation for the long optimization process of parameters.

5.2 Pareto Distribution Function

Pareto distribution is abbreviated as $PA(\alpha, \beta, \eta)$. We optimize the scale parameter α and the shape parameter β to achieve a smaller MAD. The shift parameter η is specified as one input. Parameters should satisfy the following constraints: $\alpha > 0, \beta > 0, -\infty < \eta < +\infty$. Below we recall from Section 2.5.2 the equations we will use in this chapter.

Distribution function

$$H(x) = 1 - \left(\frac{\alpha}{x - \eta + \alpha} \right)^\beta \quad (5.1)$$

Moments

$$E(X) = \frac{\alpha}{\beta - 1} + \eta, \quad \text{if } \beta > 1 \quad (5.2)$$

$$VAR(X) = \frac{\beta \alpha^2}{(\beta - 2)(\beta - 1)^2}, \quad \text{if } \beta > 2 \quad (5.3)$$

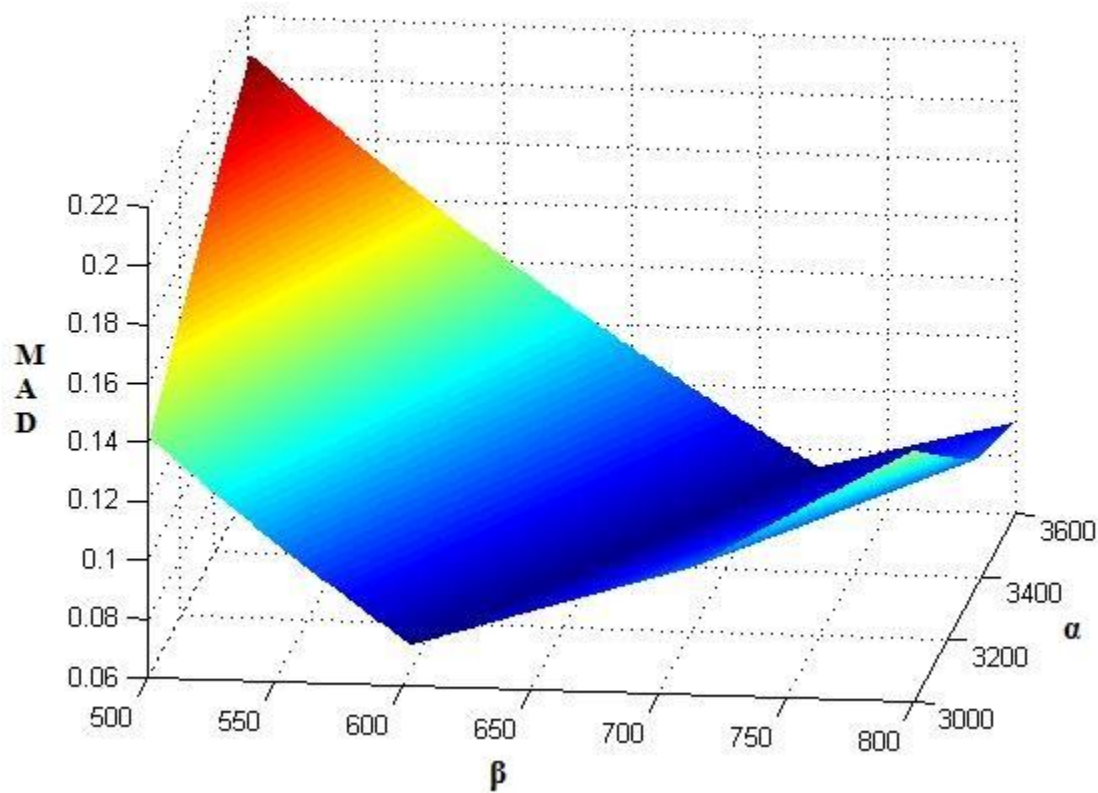


Figure 5.1. The MAD surface for the Pareto distributions with the scale parameter in the interval (3000,3600) and the shape parameter in the interval (500,800).

5.3 Optimization with Pareto Distribution

Section 2.4 describes our method of estimating the parameters of a distribution function. Two stages produce the initial estimates using the Least Squares Method, and the optimized estimates using the Uniform Distance Method. When optimizing the parameters of Pareto distribution, we observe a long search by the optimization algorithm, without too much improvement in the goodness-of-fit. In this section, we use samples A3 and A4 to illustrate this problem. Figures 5.2 and 5.3 show the Pareto distribution functions fitted to samples A3 and A4. We list the Least Squares estimates and the Uniform Distance estimates of parameters in both of the figures. Comparing the two estimates, we observe a huge increase between the input and the output of the optimization process.

To know what happens in the process of the parameters increasing from a scale of ten to a scale of thousand, we conduct the following experiments based on samples A3 and A4.

For sample A3, we choose its Pareto scale parameter in a set bounded by 1 and 7791, with step size 10. For each value of the scale parameter, we search for the best shape parameter which gives the minimum MAD. The searching for the shape parameter also ranges from 1 to 7791, with a smaller step size 1. The output of the experiment is $\frac{7791-1}{10} + 1 = 780$ pairs of Pareto distribution function's parameters and their corresponding MAD. Figure 5.4 shows the 2D plot of the MAD versus scale parameter, and Figure 5.5 shows the MAD versus shape parameter. We perform a similar experiment on sample A4, with Pareto scale parameter ranging from 1 to 3396 in step size 5, and Pareto shape parameter ranging from 1 to 3396 in step size 1. Figure 5.6 and Figure 5.7 are the 2D plots of the MAD versus a parameter.

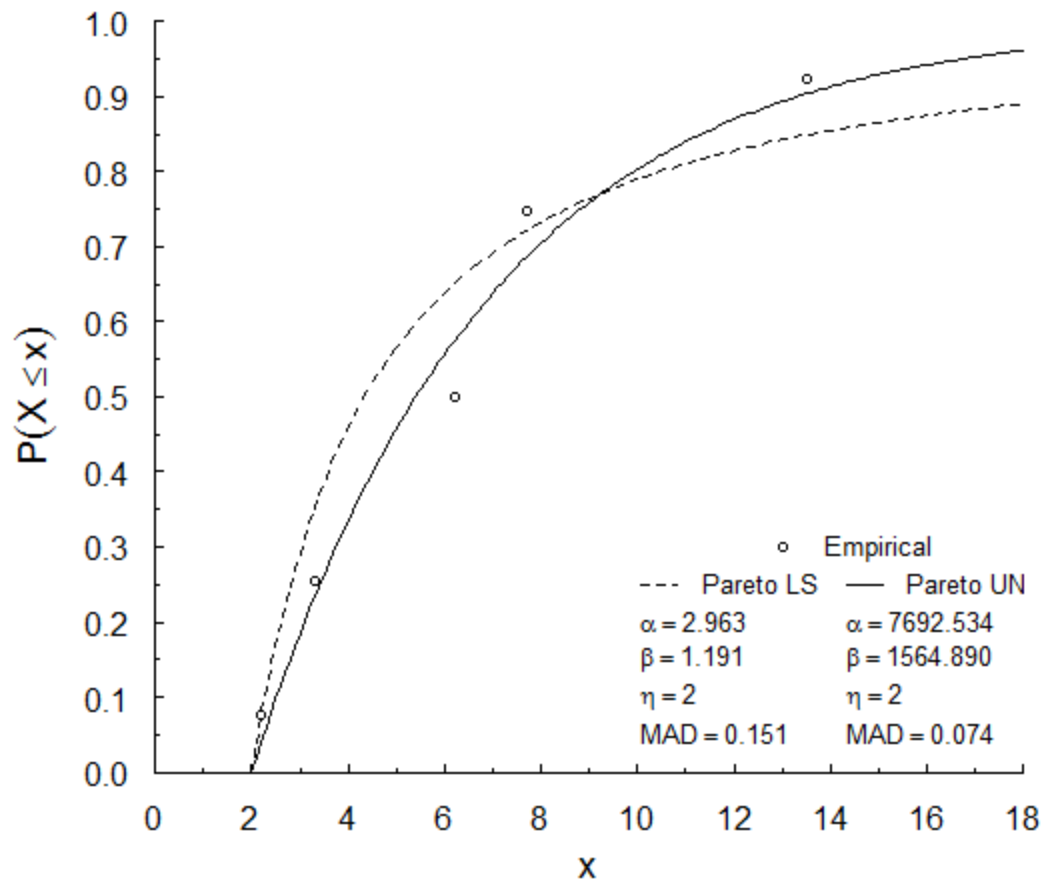


Figure 5.2. The Pareto distribution function fitted to Sample A3, with the initial fit from Least Squares Method and the optimized fit from the Uniform Distance Method. The left set of parameters is the result from Least Squares Method. The right set of parameters is the result from the Uniform Distance Method.

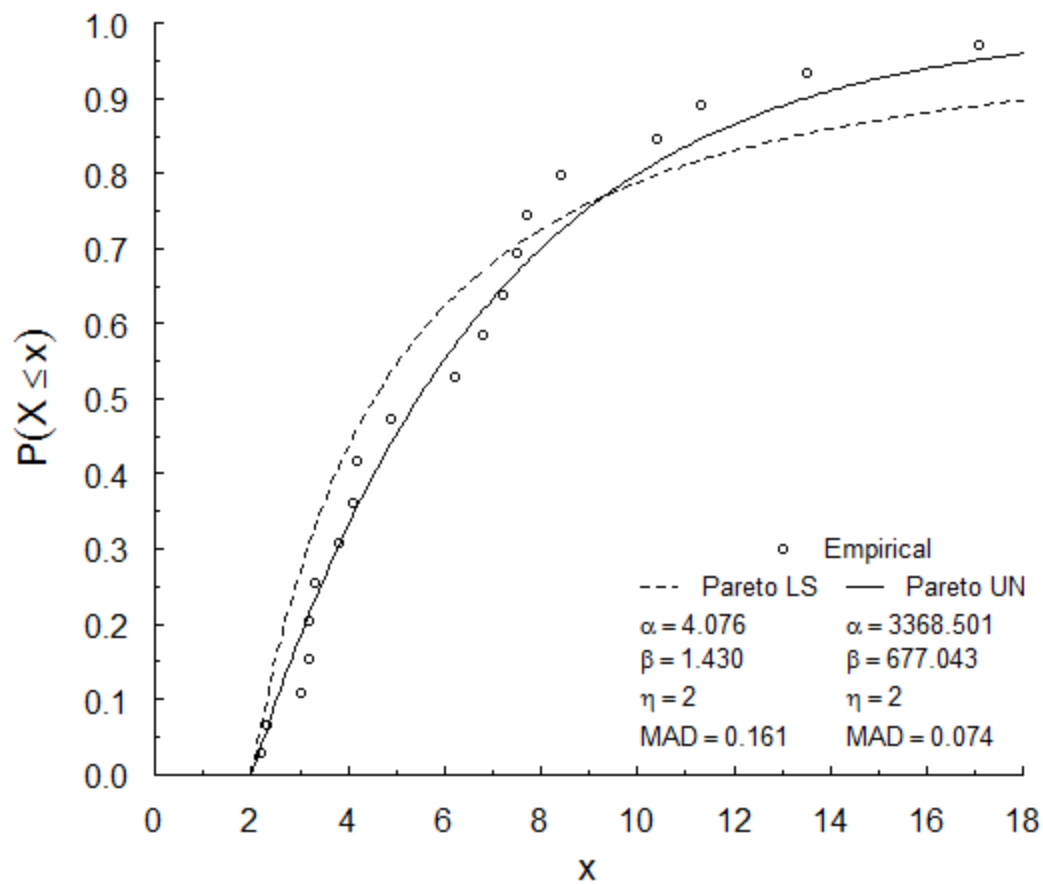


Figure 5.3. The Pareto distribution function fitted to Sample A4, with the initial fit from Least Squares Method and the optimized fit from the Uniform Distance Method. The left set of parameters is the result from Least Squares Method. The right set of parameters is the result from the Uniform Distance Method.

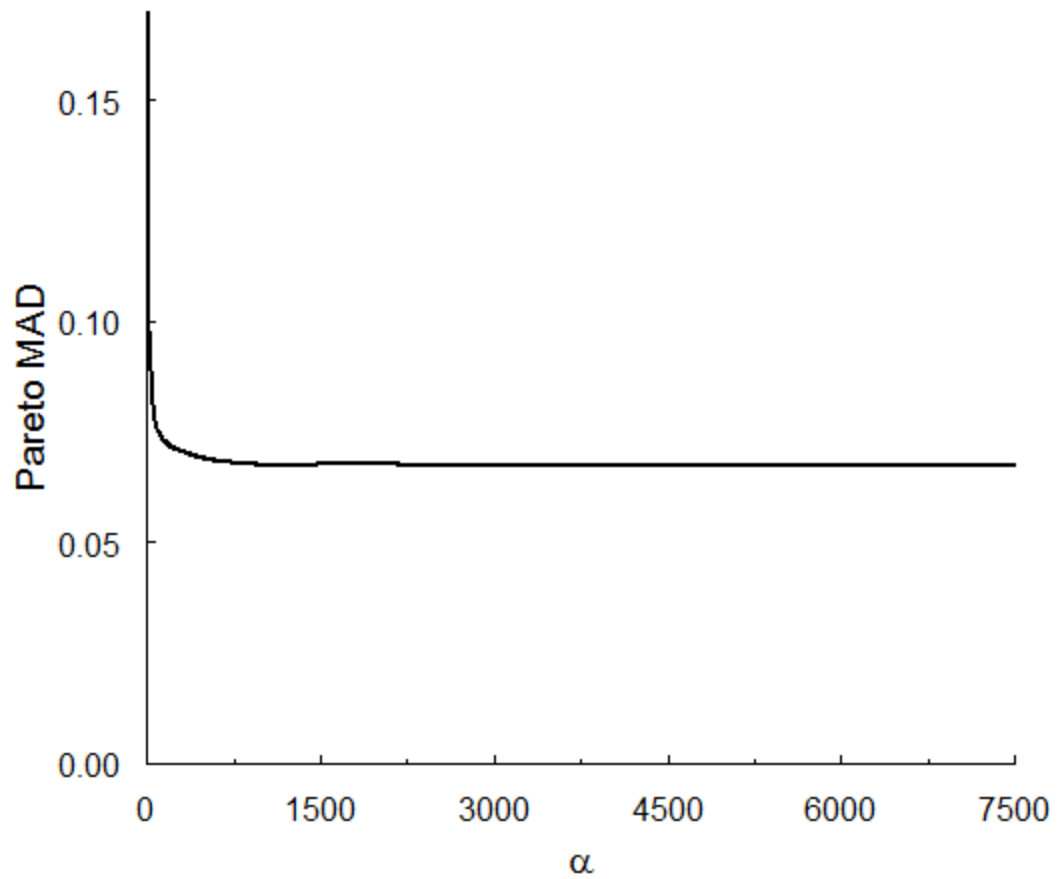


Figure 5.4. The MAD from the best-fit Pareto distribution's shape parameter versus the scale parameter, based on sample A3.

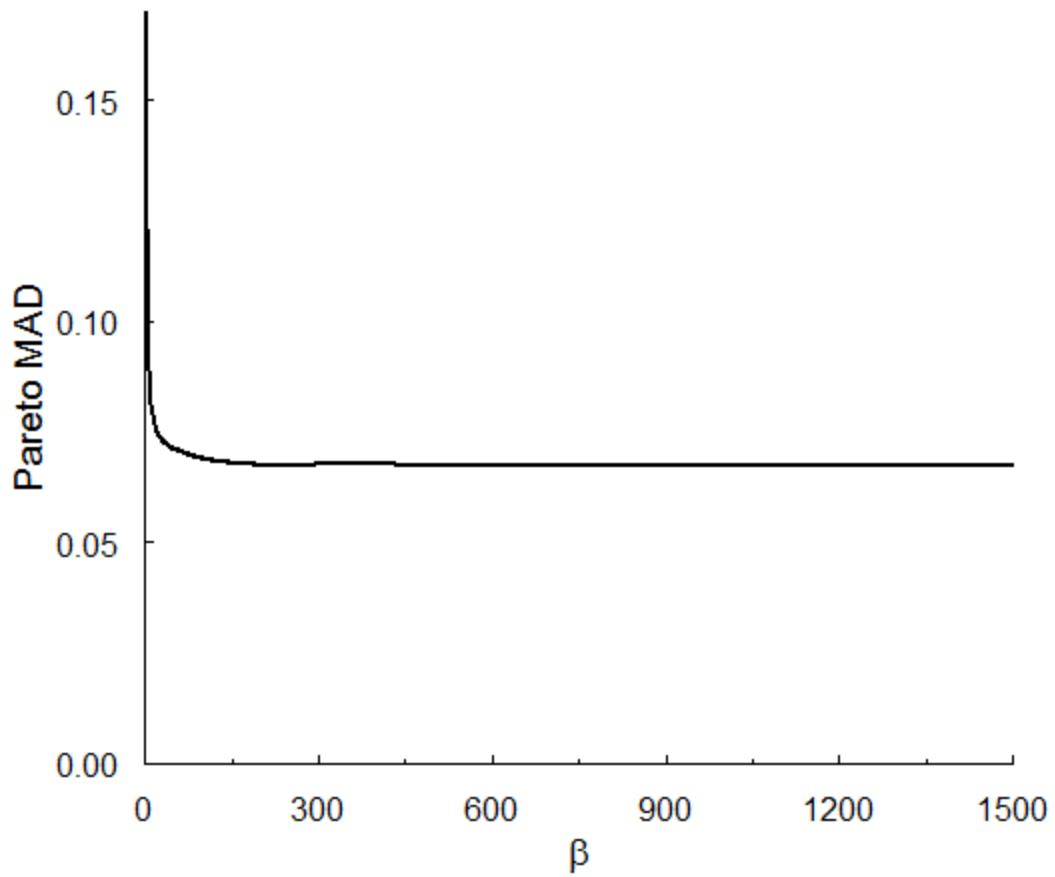


Figure 5.5. The MAD from the best-fit Pareto distribution's scale parameter versus the shape parameter, based on sample A3.

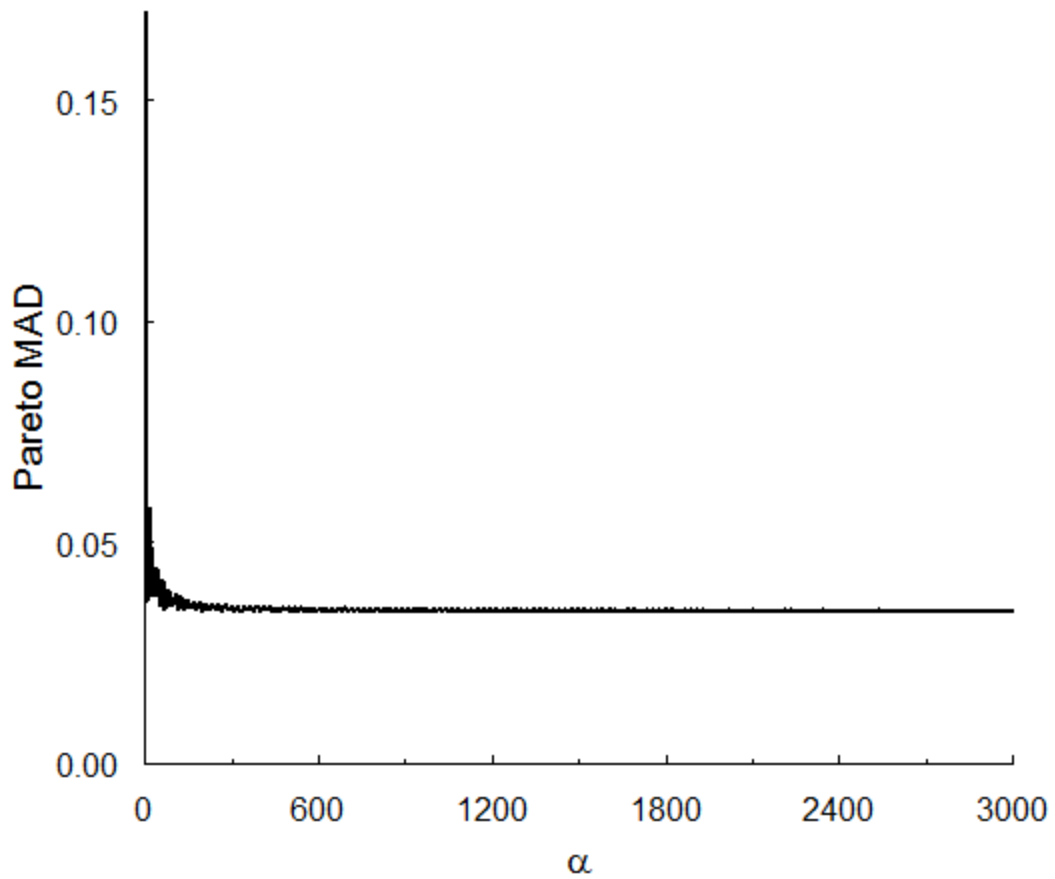


Figure 5.6. The MAD from the best-fit Pareto distribution's shape parameter versus the scale parameter, based on sample A4.

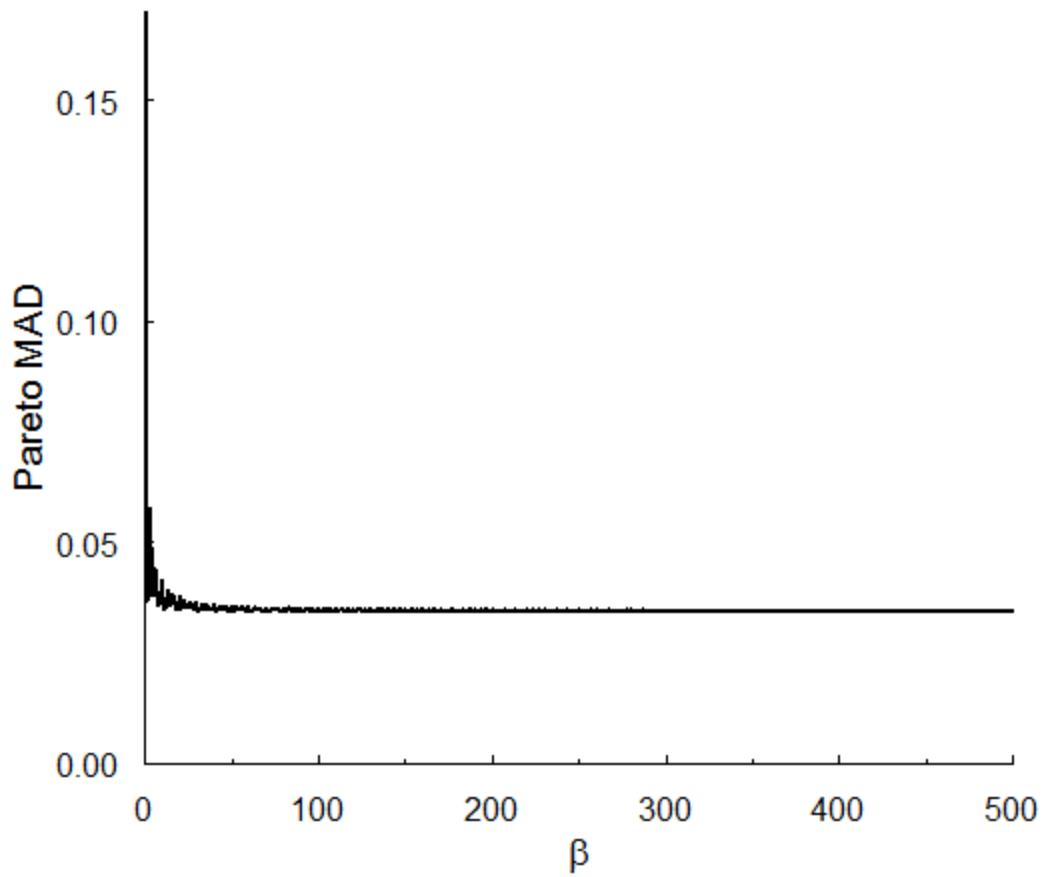


Figure 5.7. The MAD from the best-fit Pareto distribution's scale parameter versus the shape parameter, based on sample A4.

From Figures 5.4 – 5.7 we draw a conclusion that though the MAD keeps decreasing as α and β increase, the major improvement in MAD occurs before β reaches 100. This means the continued search beyond $\beta = 100$ should be cut off so that we can save time without much sacrifice of the goodness-of-fit. We explain the cause from an aspect of moments in next section.

5.4 Insight from the Moments

The mean and the variance of a Pareto variate are given by expressions (5.2) and (5.3). If we use m to denote the mean, and s^2 to denote the variance, then

$$m = E(X) = \frac{\alpha}{\beta - 1} + \eta, \quad \text{if } \beta > 1,$$

and

$$s^2 = \text{VAR}(X) = \frac{\beta \alpha^2}{(\beta - 2)(\beta - 1)^2} = \frac{\beta}{\beta - 2} \cdot (m - \eta)^2, \quad \text{if } \beta > 2.$$

Thus

$$\left(\frac{s}{m - \eta} \right)^2 = \frac{\beta}{\beta - 2}. \quad (5.4)$$

If we denote the coefficient of variation by c , then

$$c^2 = \left(\frac{s}{m - \eta} \right)^2 = \frac{\beta}{\beta - 2}, \quad \beta > 2, c > 1. \quad (5.5)$$

If $c > 1$, we can apply the method of moments to estimate the parameters:

$$\beta = \frac{2c^2}{c^2 - 1}, \quad (5.6)$$

$$\alpha = (m - \eta)(\beta - 1). \quad (5.7)$$

According to (5.5), when $\beta = 3, c = 1.7321$; when $\beta = 30, c = 1.0351$; when $\beta = 300, c = 1.0034$. As β increases, c decreases. However, no matter how large β is, c remains greater than 1. Sample A3's mean is $m = 6.58$, variance is $s^2 = 15.85$, and its coefficient of variation is $c = 0.87 < 1$. When we are given an initial estimate of β ($\beta > 1$) from Least Squares Method, the optimization process will continue increasing β to achieve smaller c , though c will never go below 1. The cause of the continued searching is that the algorithm wants to reach a “ c ” which can never be reached.

When $0 < \beta \leq 1$, the mean and variance are infinite. According to the definition, the coefficient of variation can not be calculated. When $1 < \beta \leq 2$, the variance is infinite. Thus the coefficient of variation is infinite. When $\beta > 2$, both the mean and variance are defined. Thus the coefficient of variation can be calculated and is greater than 1. Thus we postulate that Pareto distribution is only suitable for those samples whose coefficient of variation is larger than 1. To demonstrate this postulate, we do the following experiments.

We replace sample A3's largest realization by a number ranging from 8.5 to 25, increasing it by 0.1 each time. This change increases the sample coefficient of variation from 0.68 to 1.99. Simultaneously, the least squares estimates of β increases from 0 to 5. To get a wider range of β , we also replace the largest realization with some larger value, 200 and 2000. Figures 5.8, 5.9 and 5.10 show fits for three of those samples. According to our postulate, the optimization process should work as long as $c > 1$. From this test, we do not get this turning point at $c = 1$. However, we demonstrate the existence of this point when the largest realization is increased to 22 and $c = 1.13$. When $c > 1.13$, the optimization process never searches too far away. When $c < 1.13$, the optimization searches until the value of β is in thousands, with only one exception

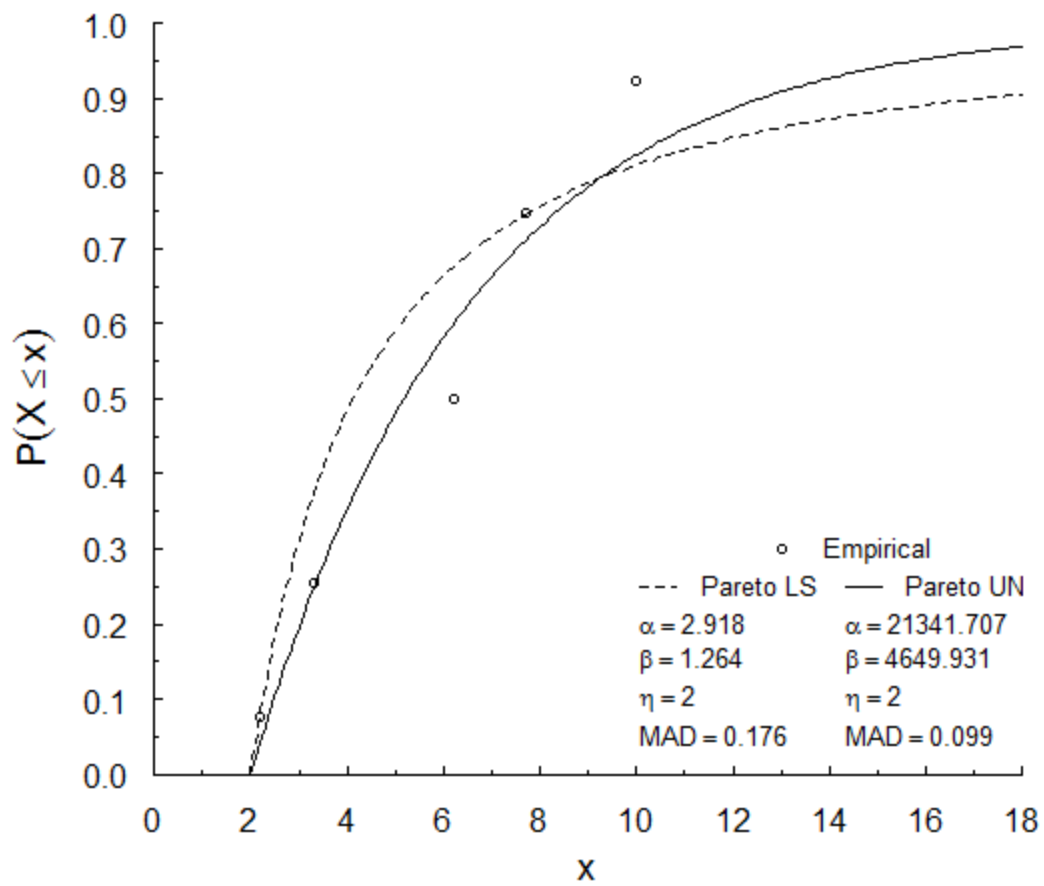


Figure 5.8. Pareto distribution fitted to sample $\{2.2, 3.3, 6.2, 7.7, 10.0\}$. The sample coefficient of variation is $c = 0.7343$. The coefficient of variation under the Pareto distribution with $\beta = 4649.931$ is $c = 1.0002$.

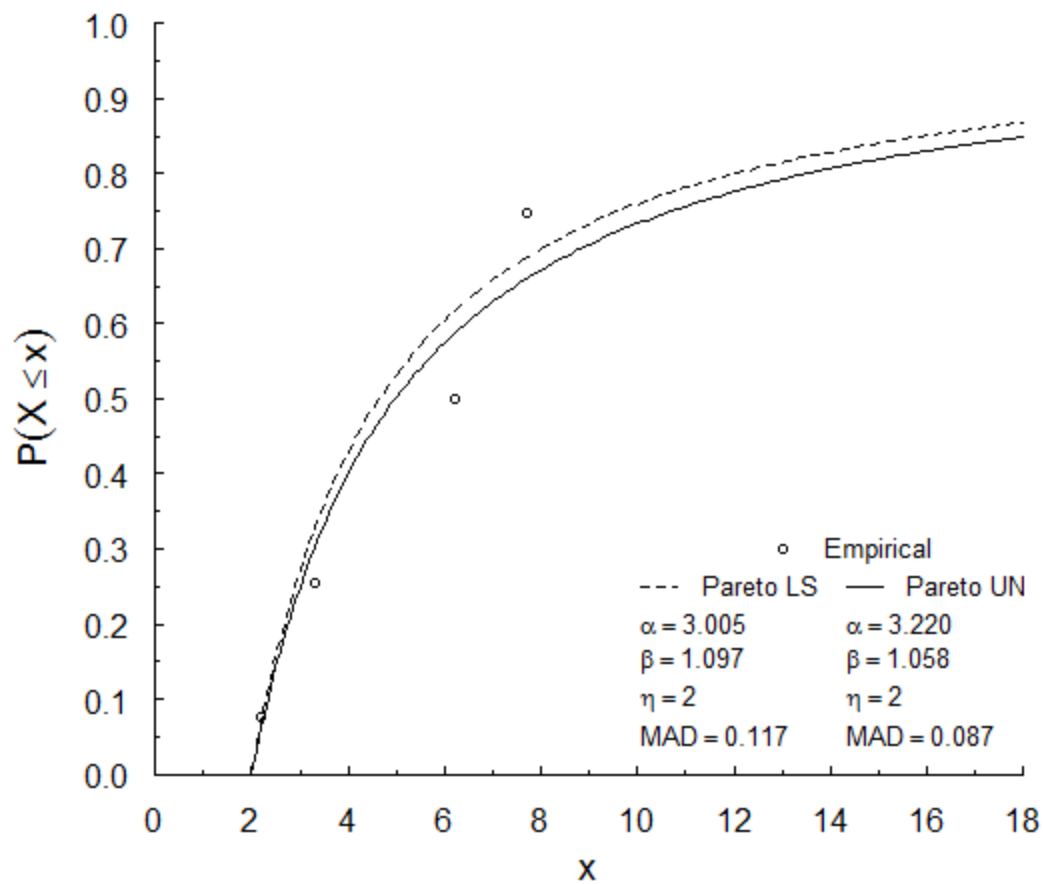


Figure 5.9. Pareto distribution fitted to sample $\{2.2, 3.3, 6.2, 7.7, 22.0\}$. The sample coefficient of variation is $c = 1.1364$. The coefficient of variation under the Pareto distribution with $\beta = 1.058$ does not exist.

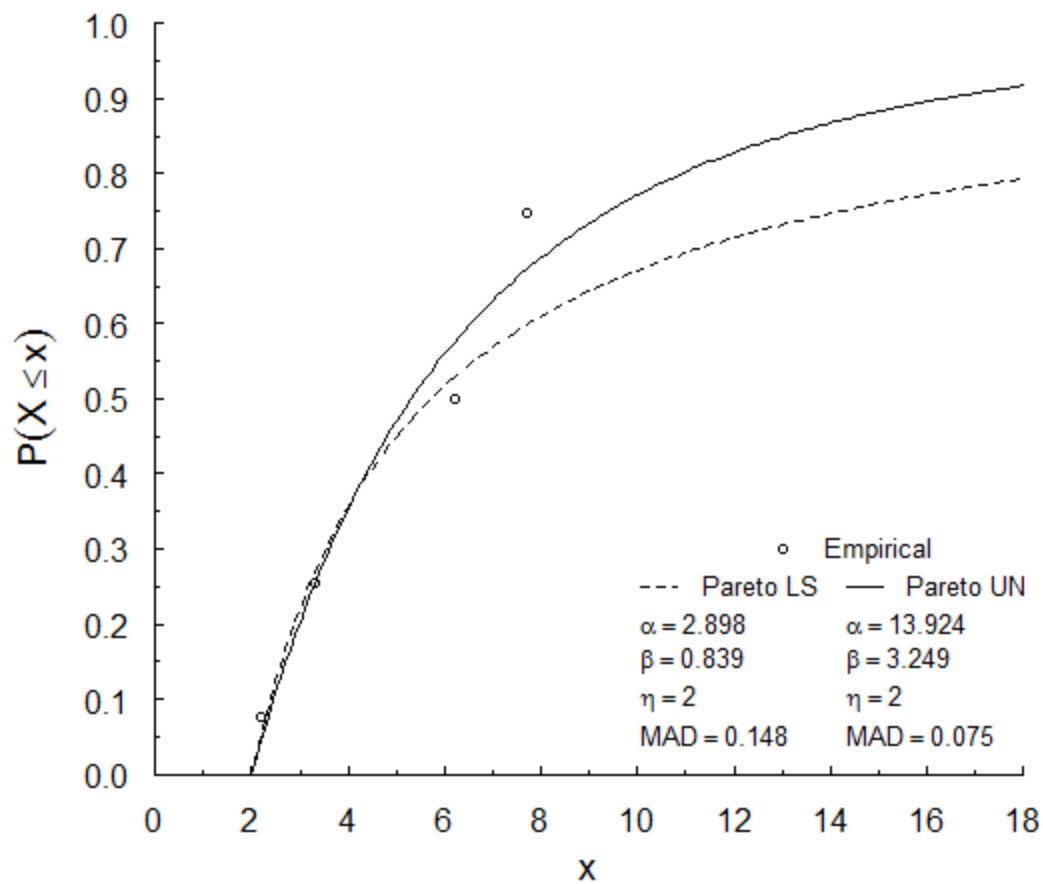


Figure 5.10. Pareto distribution fitted to sample $\{2.2, 3.3, 6.2, 7.7, 200.0\}$. The sample coefficient of variation is $c = 1.8644$. The coefficient of variation under the Pareto distribution with $\beta = 3.249$ is $c = 1.6128$.

when the largest realization is replaced by 8.8. This difference between theoretic value and experimental value is probably due to the algorithm of the optimization and the variability in the sample estimate of c . From a theoretic perspective, our postulate that Pareto distribution is only suitable for sample whose $c > 1$ works.

5.5 Conclusion

This chapter introduces the specifications of the Pareto distribution. It talks about the problems we have encountered with this distribution during the optimization of its parameters, and gives a possible explanation of the cause of the problem. At last, we design an experiment to demonstrate our postulate.

CHAPTER 6 : FALLACIES RELATED TO CONTINUOUS DISTRIBUTIONS

6.1 Chi- Square Test

Chi-square test aims at testing the null hypothesis that a given sample of realizations is drawn from, or “fits,” a specified probability distribution of a discrete variate X representing K classes or categories (Milton & Arnold, 2004). Let (X_1, X_2, \dots, X_k) be a multinomial random vector with parameters n, p_1, p_2, \dots, p_k . For large n , the random variable

$$D_1 = \sum_{i=1}^k \frac{(X_i - np_i)^2}{np_i} \quad (6.1)$$

follows approximately chi-square distribution with $k - 1$ degrees of freedom.

When applying to real samples, the chi-square statistic can be expressed as

$$D_1 = \sum_{i=1}^k \frac{[(\text{observed frequency}) - (\text{expected frequency})]^2}{\text{expected frequency}}. \quad (6.2)$$

If this statistic is large enough, we reject the null hypothesis that a given sample is drawn from a specified distribution.

6.2 Kolmogorov-Smirnov Test

Kolmogorov-Smirnov test, in short, K-S test, is a goodness of fit test for continuous variates (Lindgren, 1976). If the K-S statistic is large enough, the null hypothesis that variate X has the distribution function H is rejected. We are given a sample of size N with sample space (η_L, η_U) . The test employs an empirical distribution function constructed as

$$\check{H}(x) = \frac{n}{N}, \quad \text{for } x_{(n)} \leq x \leq x_{(n+1)}, \quad n = 0, 1, \dots, N, \quad (6.3)$$

where $x_{(0)} = \eta_L$, and $x_{(N+1)} = \eta_U$.

The Kolmogorov-Smirnov statistic is thus defined as

$$D_2 = \sup_x |\check{H}(x) - H(x)|. \quad (6.4)$$

As $\check{H}(x)$ is a stepwise function and $H(x)$ is a continuous function, D_2 can be calculated as

$$D_2 = \max \left\{ \max_{1 \leq n \leq N} \left| \frac{n-1}{N} - H(x_{(n)}) \right|, \max_{1 \leq n \leq N} \left| \frac{n}{N} - H(x_{(n)}) \right| \right\}, \quad n = 1, \dots, N. \quad (6.5)$$

The hypothesized function H is accepted if D_2 is smaller than a constant c .

6.3 Comparisons of the Tests

The K-S test has an advantage over the chi-square test for its stability: for the same input sample and the hypothesized distribution function, the result of K-S test is always the same (e.g., when each of several statisticians applies the test). For the chi-square test, the sample space is divided into several class intervals. However, different starting points and length of class intervals may lead to significantly different chi-square test results (Benjamin and Cornell, 1970). The chi-square test also fails to recognize the order of realizations within each class interval, while the K-S test really cares about the order (Lindgren, 1976). We want to conclude in this chapter that K-S test is suitable for continuous variate, while the chi-square test is suitable only for discrete variate.

To illustrate the above conclusion, I created the following example. Suppose we are applying the chi-square test to a continuous variate X ; we first need to partition the sample space of X into

class intervals. We are given a sample of X as

$$\{0.1, 1, 3, 4, 5, 5, 12, 13, 14, 14, 21, 23, 24, 24.5, 24.9\}.$$

If we partition this sample space in two different manners, the histograms of the variate X could have two different shapes as shown in Figure 6.1. The upper histogram has an interval width 10, and the lower histogram has an interval width 5. They both start from 0. If the upper histogram is used, the probability of X falling within the interval $(0,10]$ is $\frac{6}{6+4+5} = 0.4$. The three class intervals are $(0,10]$, $(10,20]$, and $(20,30]$. If the lower histogram is used, the probability of X falling within the interval $(5, 10]$ is 0, given the class interval width 5. This information is not included in the upper histogram, which even misleads us to believe the probability of X falling within $(5, 10]$ is not 0. We call it information distortion generated from the rough discretization of continuous sample space. However, the problem is not solved if we narrow the class interval. Also, the intervals could be translated to generate different histograms. Thus chi-square test may result in conflicting conclusions with the same input because of the different discretizations of the continuous sample space.

To explain more clearly, we input the above sample of X to the distribution fitter, outputting the hypothesized *kappa* distribution function

$$H(x) = \left[1 + \left(\frac{x-\eta}{\alpha} \right)^{-\beta} \right]^{-1/\beta}, \quad (6.6)$$

with $\alpha = 21.08$, $\beta = 2.95$, and $\eta = 0$.

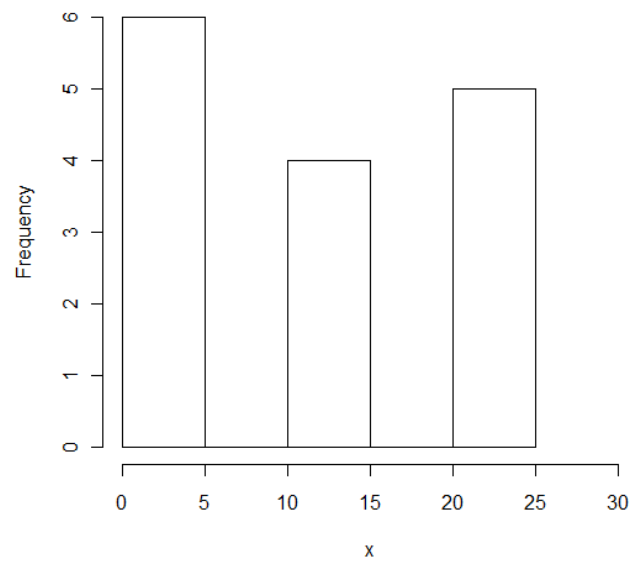
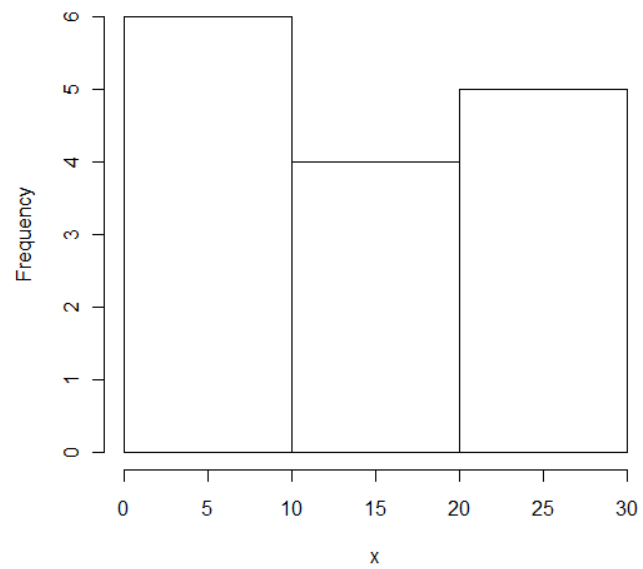


Figure 6.1. Two histograms of variate X constructed from the same sample of size 15, with the class intervals of width 10 (upper) and 5 (lower).

Figure 6.2 shows two density histograms and the kappa density function corresponding to the above distribution function. The density of each class interval is the frequency of that interval divided by the product of the sample size and the interval width. The two plots of Figure 6.2 have the same density functions, but different density histograms. The upper histogram is misleading about the density of X at intervals $(5,10]$, $(15,20]$, and $(25,30]$. We can draw a conclusion from this figure that it is not proper to apply chi-square test to continuous distribution.

According to equation (6.2), I list the expected frequency and observed frequency in Table 6.1. From it, we can tell that the left class intervals and the right class intervals generate quite different chi-square statistics. According to the table of cumulative chi-square distribution, D_1 from the right class intervals has a p value smaller than 0.005, while D_1 from left class intervals has a p value 0.1. The K-S test statistic is $D_2 = 0.16$, which means for the sample of size 15, we do not reject the model at a significant level 0.2. Figure 6.3 shows the estimated kappa distribution function and the K-S test statistic.

Table 6.1. Comparison of the chi-square test results with two different discretizations, and the comparison of the chi-square test with the K-S test.

Class Interval	Expected frequency	Observed frequency	Density	Class Interval	Expected frequency	Observed frequency	Density
(0,10]	0.45	0.4	0.040	(0,5]	0.23	0.4	0.080
(10,20]	0.31	0.27	0.027	(5,10]	0.22	0	0
(20,30]	0.14	0.33	0.033	(10,15]	0.19	0.27	0.54
				(15,20]	0.12	0	0
				(20,25]	0.09	0.33	0.066
				(25,30]	0.05	0	0
$D_1 = 4.18$				$D_1 = 18.06$			
$D_2 = 0.16$							

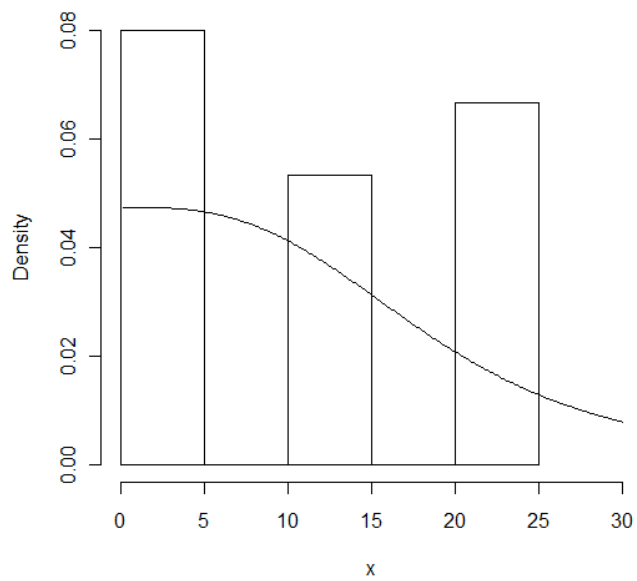
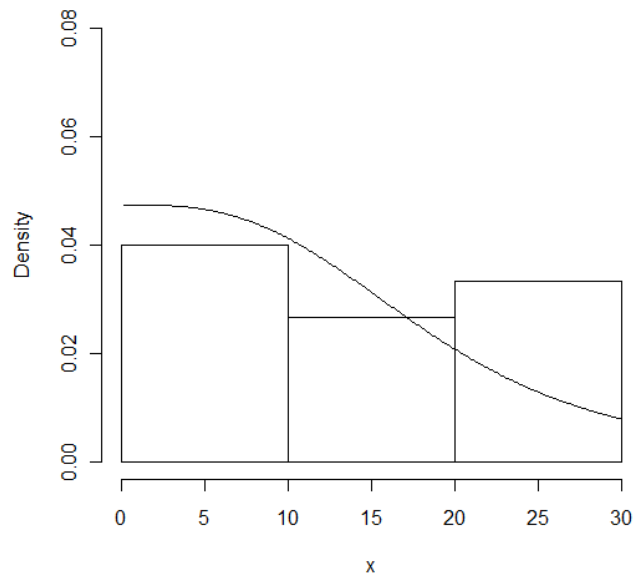


Figure 6.2. Two density histograms of variate X constructed from sample of size 15, and the kappa density function.

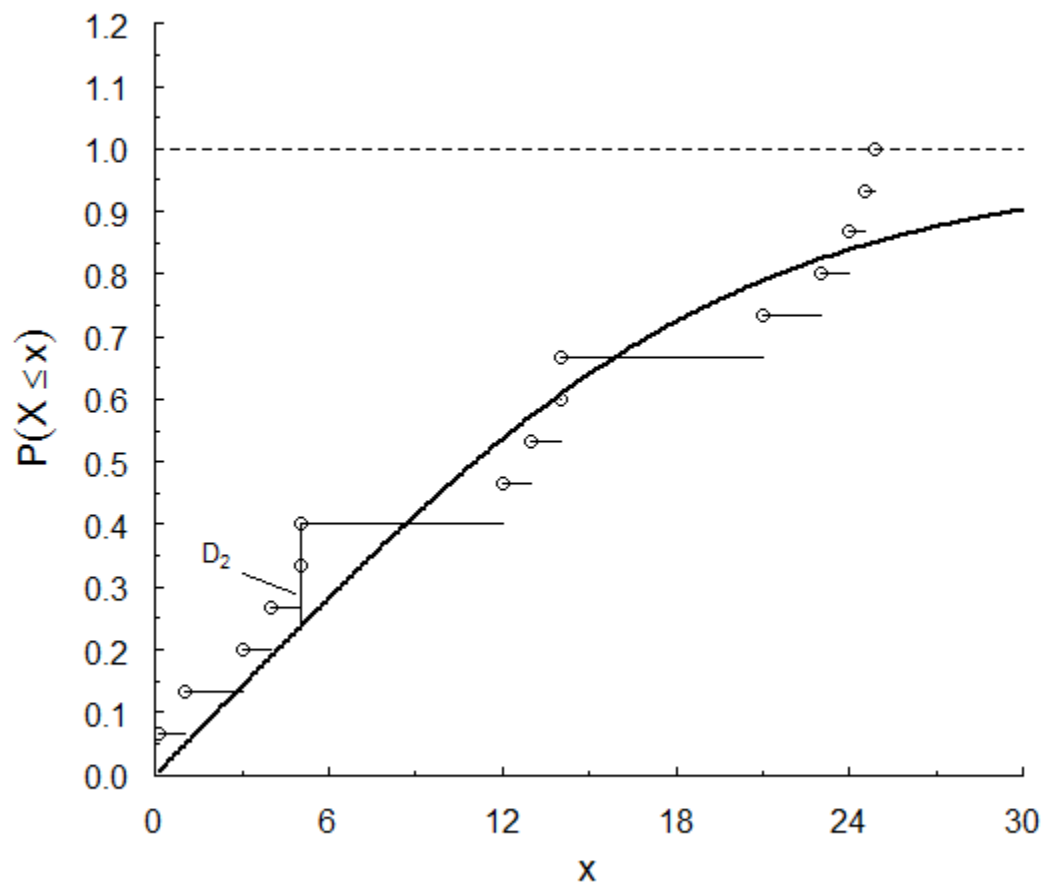


Figure 6.3. Applying K-S test to the estimated kappa distribution function.

The above experiment means this: if we conduct the chi-square test and discretize the sample space of X using the right class intervals in Table 6.1, we reject the hypothesized distribution function H at the significance level of 0.05 as the p value is so small. If we discretize the sample space of X using the left class intervals in Table 6.1, we accept the hypothesized function H at the significance level of 0.05. This is a contradiction within the chi-square test. There is also a contradiction between the chi-square test and the K-S test. This experiment supports our conclusion.

5.4 Conclusion

This chapter compares two goodness-of-fit tests: the chi-square test and the Kolmogorov-Smirnov test. We demonstrate that the chi-square test should only be applied to a discrete distribution, or it may generate conflicting results when applied to the continuous distribution.

APPENDIX : DATA SAMPLES

We use 20 samples which are of four categories in this thesis. They are 4 artificial samples, 5 random samples, 6 river stage samples, and 5 National Weather Service samples.

A.1 Artificial Samples

Artificial samples were created by me mainly for the purpose of comparing the behaviors of the least square estimation method and the uniform distance estimation method. These artificial samples come from bounded sample space(2,18). Samples A1 and A3 are of size 5. Samples A2 and A4 are of size 20.

Table A.1. Artificial samples A1 – A4.

A1	A2	A3	A4
3.7	3.2	2.2	2.2
6.2	3.9	3.3	2.3
9.1	5.9	6.2	3
10.5	6.5	7.7	3.2
12.9	6.6	13.5	3.2
	6.8		3.3
	7.2		3.8
	7.8		4.1
	7.9		4.2
	8.2		4.9
	8.4		6.2
	8.8		6.8
	9.2		7.2
	9.5		7.5
	10.4		7.7
	10.9		8.4
	11.7		10.4
	12		11.3
	13		13.5
	16.1		17.1

A.2 Random Samples

Random samples were generated randomly from five Weibull distribution functions of quite different shapes. The lower bound of the sample space is 0. The sample size is 20. The Weibull distribution functions are specified by scale parameter and shape parameter: $\alpha = 1, \beta = 0.5$; $\alpha = 2, \beta = 1$; $\alpha = 4, \beta = 2$; $\alpha = 6, \beta = 3$; $\alpha = 8, \beta = 4$. For example, sample R1 is randomly drawn from the Weibull distribution function with $\alpha = 1, \beta = 0.5$. And R5 is drawn randomly from the Weibull distribution function with $\alpha = 8, \beta = 4$.

Table A.2. Random samples R1 – R5.

R1	R2	R3	R4	R5
0.0041	0.2363	1.3299	2.4327	1.5198
0.0377	0.2662	1.3685	2.7857	4.4243
0.0507	0.3679	1.8907	3.2523	4.9683
0.0532	0.6563	1.8994	3.9967	5.5580
0.0952	0.6980	2.2077	4.0476	5.9556
0.2166	0.7942	2.2600	4.2287	6.0343
0.2286	0.9721	2.3139	4.6919	6.1255
0.2349	1.4065	2.4765	5.3593	6.3134
0.4741	1.4349	2.8007	5.4015	6.9399
0.7236	1.5273	3.2848	5.8610	7.0722
0.9838	1.6295	3.4657	6.0626	7.1214
1.1689	1.7237	3.5989	6.2545	7.4106
1.3494	1.8431	3.6202	6.3861	7.5359
1.5990	1.8547	4.3005	6.7732	8.0675
2.1579	1.9938	4.3955	6.8212	8.1082
2.2578	2.4327	5.1380	7.0865	8.4912
5.2286	3.2955	5.4351	8.1170	8.8028
5.7036	3.5405	5.9169	8.4314	9.0083
8.3783	4.0487	6.3086	8.8126	10.1863
23.2001	4.5556	7.0824	9.3795	10.2475

A.3 River Stage Samples

River stage samples are composed of the river stage [ft] daily records at Eldred, PA, from year 1992 to year 1996. Sample S1 records the river stage at Eldred in the period September 1 – 10 from 1992 to 1996; sample S2 is the record in the period September 11 – 20 from 1992 to 1996; sample S3 is the record in the period September 21 – 30; sample S4 is the record in the period November 1 – 10; sample S5 is the record in the period November 11 – 20; sample S6 is the record in the period November 21 – 30. Theoretically, each sample should be of size $10 \text{ (days/year)} \times 5 \text{ (years)} = 50$. However, as some records are missing, the sample size is smaller than or equal to 50. As September represents a warm season, the water flow is low. Thus the lower bound of sample space of S1, S2 and S3 is 1.25 [ft]. As November represents a cool season, the water flow is high. Thus the lower bound of sample space of S4, S5 and S6 is 3.45 [ft].

Table A.3. River stage samples S1 – S6.

S1	S2	S3	S4	S5	S6
1.46	1.52	1.53	3.57	4.4	4.3
1.46	1.57	1.56	3.64	4.5	4.9
1.47	1.58	1.56	3.8	4.7	5
1.48	1.58	1.58	3.8	4.9	5.1
1.49	1.6	1.59	3.9	5.4	5.1
1.51	1.64	1.64	4.1	5.7	5.3
1.54	1.65	1.7	4.4	5.8	5.3
1.56	1.65	1.84	4.42	6.1	5.34
1.58	1.78	1.87	4.5	6.1	5.5
1.65	1.96	2.2	4.55	6.3	5.54
1.8	2.2	2.3	4.57	6.58	5.68
1.81	2.3	2.4	4.6	6.6	5.7
1.85	2.4	2.5	4.7	6.6	5.7
1.85	2.4	2.5	4.7	6.68	5.9

1.87	2.5	2.5	4.71	6.7	5.9
1.89	2.6	2.6	4.72	6.71	6.05
1.93	2.65	2.6	4.72	6.8	6.18
1.97	2.7	2.7	5.21	7.16	6.3
2.2	2.7	2.8	5.3	7.2	6.32
2.9	2.7	3.5	5.4	7.2	6.62
2.93	2.8	3.6	5.4	7.5	6.7
3.1	2.8	3.62	5.6	7.8	6.76
3.23	2.8	3.8	5.7	7.97	7
3.3	2.8	3.9	5.9	8	7.3
3.3	2.9	4.06	5.9	8.3	8.1
3.4	2.9	4.1	6.1	8.44	8.39
3.4	3	4.32	6.1	8.6	8.5
3.4	3	4.7	6.2	8.7	8.9
3.4	3	4.77	6.2	9.3	9.03
3.4	3	5	6.3	9.37	9.1
3.4	3	5.09	6.5	9.58	9.5
3.4	3	5.29	6.5	9.8	9.6
3.5	3	5.4	6.7	10.4	9.9
3.6	3.04	5.5	6.9	10.5	10.2
3.6	3.1	5.94	7	10.7	
3.61	3.2	6.2	7.1	11	
3.7	3.3	6.6	7.5	11.3	
3.8	3.3	7	7.8	11.4	
3.8	5.01	7.1	8.2	13.4	
3.9	5.23	8.32	8.2	15.3	
4	5.61	9.31	8.2		
4.1	6.54	10.8	9.2		
4.3	6.83	12.5	9.9		
4.5	8.23	12.7	10.7		
4.8	8.53		12.1		
4.8	9.24		15.3		
5					
5.3					
5.8					
6.5					

A.4 National Weather Service Samples

National Weather Service samples are the records of heights of the 500 hPa isobar. National Weather Service divides the globe into $144 \times 73 = 10512$ grid points evenly. Imagine our globe is a two dimensional field. The longitude is divided into 144 segments, and each segment is $360^\circ \div 144 = 2.5^\circ$ wide. The latitude is divided into 72 segments, and each segment is $180^\circ \div (73 - 1) = 2.5^\circ$ high. The variate is the height of the 500 hPa isobar obtained at a grid point [ft] through analysis using a numerical weather prediction model and representing an estimate of the “true” value. Each sample of the variate contains 62 realizations measured daily from July 1 to August 31 in 2009. We have 10512 samples in total for each of the 10512 grid points. In addition, each sample is standardized. Samples N1 – N5 are five of them with the lower bound -5.

Table A.4. National Weather Service samples N1 – N5.

N1	N2	N3	N4	N5
0.5156	2.0846	-1.7737	-1.4936	-0.3025
1.0259	0.2964	-0.6574	-0.9403	-0.1915
1.4418	1.0749	-0.3784	0.512	0.3634
1.9049	1.3912	-0.6749	1.2036	0.3634
1.4985	0.4667	-1.9307	-0.2487	0.2524
0.8653	0.2235	-2.0528	-0.6637	0.0304
0.2131	-0.1171	-2.0005	-0.8712	-0.9683
0.3832	-0.3118	-2.0353	-1.217	-0.7464
0.232	-0.5794	-1.8609	-0.8712	-1.0793
-0.8738	-0.6645	-1.0237	-0.0413	1.1402
-0.6564	-0.2509	-0.6923	-0.3871	-0.3025
-0.2216	0.5154	-0.6574	-0.5254	0.1414
-0.9683	1.7805	-0.8667	-0.0413	1.1402
0.0997	2.3886	-1.0586	-1.0786	1.2512
0.0808	2.4616	-1.3551	0.3045	1.917
0.9503	1.4885	-2.0353	2.0335	1.4731
0.6006	0.114	-0.797	1.9643	0.8073
-0.0515	0.2235	0.0751	1.5494	0.5853
0.1375	0.4059	0.7205	0.9269	1.2512

0.9787	1.2817	1.1391	-0.5254	0.2524
1.8671	0.9655	1.5053	-2.3235	1.1402
2.2546	1.1358	1.453	-2.3927	0.5853
1.9994	0.4424	0.9298	-0.4562	0.0304
1.7348	-0.1415	0.4937	0.9269	0.0304
0.9409	-0.0198	0.267	0.8578	0.1414
0.3265	-0.8105	0.6507	0.7886	-0.8574
0.1186	-0.2631	0.703	1.2036	-1.5232
-0.1649	-0.5064	0.7205	1.0653	-1.0793
0.3927	-0.555	0.1449	0.4428	-0.0805
1.0259	0.0532	0.6507	0.2354	-0.0805
1.1582	0.9533	0.6681	1.2036	-0.4135
0.9881	-0.1293	1.2263	1.6877	0.3634
0.6857	-0.3361	0.9298	0.7886	-1.1903
0.6101	-0.3847	0.8774	0.4428	-1.6342
0.1091	0.5519	0.6856	0.5812	-0.3025
-0.7037	0.5032	0.5286	-0.0413	-0.0805
-0.4012	0.0897	-0.1691	-0.5254	0.3634
-0.3918	0.3329	-0.0121	-0.3871	1.1402
-0.1271	0.2721	-0.8319	0.512	1.5841
0.336	0.2843	-1.4772	2.3793	0.9182
0.336	-0.3604	-0.7098	1.5494	0.5853
-0.0988	0.2843	-0.6051	0.5812	0.9182
-0.4107	0.6127	-0.3086	0.3737	1.0292
-0.8171	-0.1779	0.3367	-0.0413	1.806
-1.2613	-0.3969	0.7902	0.097	1.3621
-1.3086	-0.2266	0.1274	-0.7329	0.8073
-1.6205	-0.6645	0.4239	-1.7702	-0.5244
-1.3275	-1.1754	0.7379	-0.8712	-1.6342
-1.1762	-1.2484	0.8949	-0.2487	-1.1903
-1.2235	-0.9686	0.703	-1.2861	-1.3013
-1.4976	-0.9564	0.546	-0.5254	-1.7452
-1.318	-0.1293	0.7379	0.1662	-1.8561
-0.7415	-1.8809	1.5751	-0.3179	-1.6342
-0.7037	-0.2388	1.0344	-0.802	0.0304
-0.5714	-0.6037	1.1216	-0.5945	-1.0793
-1.1006	-1.8809	0.86	-0.6637	-0.7464
-1.0345	-3.0365	0.3193	-0.5254	-0.4135
-0.5146	-1.3335	-0.2388	-0.6637	0.3634
-0.5052	0.1991	-0.047	-0.3179	0.6963
-0.9021	-1.0538	0.4065	-0.5945	0.9182
-1.4976	-1.5768	0.4414	0.1662	-0.9683
-1.6205	0.1018	0.8251	0.4428	-1.8561

REFERENCES

- Aitkin, M. & Clayton, D. (1980). The Fitting of Exponential, Weibull and Extreme Value Distributions to Complex Censored Survival Data Using GLIM. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, **29**(2), 156-163.
- Balakrishnan, N. & Kateri, M. (2008). On the Maximum Likelihood Estimation of Parameters of Weibull Distribution Based on Complete and Censored Data. *Statistics and Probability Letters*, **78**, 2971-2975.
- Bar-Lev, S. K. (2004). Likelihood-Based Inference for the Shape Parameter of a Two Parameter Weibull Distribution. *Lifetime Data Analysis*, **10**(3), 293-308.
- Benjamin, J.R. & Cornell, C.A. (1970). *Probability, Statistics, and Decision for Civil Engineers*, McGraw-Hill, Inc.
- Brent, R.P. (1973). *Algorithms for Minimization without Derivatives*, Chapter 5, Prentice Hall, Englewood Cliffs, New Jersey.
- Chen, D.G. & Lio, Y.L. (2010). Parameter Estimations For Generalized Exponential Distribution Under Progressive Type-I Interval Censoring. *Computational Statistics and Data Analysis*, **54**, 1581-1591.
- Cohen, A.C. (1965). Maximum Likelihood Estimation in the Weibull Distribution Based on Complete and on Censored Samples. *Technometrics*, **7**(4), 579-588.
- Cohen, A.C. (1975). Multi-Censored Sampling in the Three Parameter Weibull Distribution. *Technometrics*, **17**(3), 347-351.

- Engelhardt, M. & Bain, L.J. (1973). Some Complete and Censored Sampling Results for the Weibull or Extreme-Value Distribution. *Technometrics*, **15**(3), 541-549.
- Jang, D., Park, J. & Kim, C. (2011). Estimation of the Scale Parameter of the Half-logistic Distribution With Multiply Type II Censored Sample. *Journal of the Korean Statistical Society*, article in Press.
- Klakattawi, H.S., Baharith, L.A. & AL-Dayian, G.R. (2011). Bayesian and Non Bayesian Estimations on the Exponentiated Modified Weibull Distribution for Progressive Censored Sample. *Communications in Statistics- Simulation and Computation*, **40**, 1291-1309.
- Krzysztofowicz, R. (1992). A Bayesian Estimator of an Empirical Distribution, Working paper. University of Virginia.
- Krzysztofowicz, R. (2011a). *Distribution fitter for NCEP: Statistical Methodology*. University of Virginia.
- Krzysztofowicz, R. (2011b). Personal communication.
- Krzysztofowicz, R. (2012). *Probabilistic Forecasts and Optimal Decisions, Packet for SYS 3060*. University of Virginia.
- Lindgren, B. W. (1976). *Statistical Theory, 3rd edition*. Macmillan, New York.
- Luenberger, D.G. (1984). *Linear and Nonlinear Programming*, 2nd edition. Addison-Wesley, Reading, Massachusetts.
- Milton, J.S. & Arnold, J.C. (2004). *Introduction to Probability and Statistics*, McGraw-Hill, NY.

- Nelder, J. A. & Mead, R. (1965). A Simplex Method for Function Minimization. *Computer Journal*, **7**(4), 308-313.
- Peressini, A., Sullivan, F. & Uhl, J. (1993). *The Mathematics of Nonlinear Programming*, Springer Verlag, New York.
- Thomas, D.R. & Wilson, W.M. (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.