MODIFIED WEIBULL DISTRIBUTIONS IN RELIABILITY ENGINEERING

JIANG HONG

NATIONAL UNIVERSITY OF SINGAPORE

2010

MODIFIED WEIBULL DISTRIBUTIONS IN RELIABILITY ENGINEERING

JIANG HONG
(B.S., USTC)

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
DEPARTMENT OF INDUSTRIAL AND SYSTEMS
ENGINEERING
NATIONAL UNIVERSITY OF SINGAPORE

2010

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my supervisor Professor Xie Min, for his guidance on my study and research at NUS. I am grateful for his understanding, encouragement and patience, without which this thesis would not have been possible.

I would also like to thank my supervisor Professor Tang Loon Ching, for his motivation and support on this work. Prof Tang is always accessible and willing to answer my questions and provide constructive comments.

I am delightful to interact with Prof Goh Thong Ngee, Prof Ang Beng Wah, Prof Poh Kim Leng, Prof Lee Loo Hay, Dr Ng Szu Hui and Dr Chai Kah Hin, by attending their classes and having their inspirations on related research topics.

Also, I would like to say thanks to department officers and technologists, especially Ms. Ow Lai Chun and Mr. Victor Cheo Peng Yim, for the assistance they always timely provide, which ensures my research life in NUS is smooth and rewarding.

My friends at the Computing Lab made it a harmonious and comfort place to work. Their sincerity wins my friendship and gratitude.

Last but not least, my deepest gratitude goes to my parents for their unflagging care and support, and the love from the bottom of my heart belongs to my beloved wife Wu Yanping and our son to be born.

TABLE OF CONTENTS

SUMMARY

This thesis concerns the modeling of the Weibull family to lifetime data, studies the statistical properties of the distributions, and considers the parameter estimation based on a complete or censored sample. Related issues such as model selection, evaluating mean residual life and burn-in time are addressed as well.

In our research, the modified Weibull distribution and odd Weibull distribution are studied. As an important step in Weibull analysis, model characterization provides insight into the properties and applicability to model data of the distributions. For the distributions considered, we describe the important statistics and distribution functions, both in analytical and numerical ways.

Parameter estimation is crucial for the model to be built and is often a difficult problem, especially for distributions with more than 2 parameters. In this thesis, maximum likelihood estimation is studied in detail. Several techniques regarding this estimation method are proposed to simplify computation, which help look into the existence and uniqueness properties of the estimators. Another estimation method called Markov chain Monte Carlo is used to estimate the parameters of the modified Weibull distribution and is found to outperform MLE in several aspects when the prior is independent generalized uniform and the size of sample data is small. A graphic parameter estimation method is proposed for the odd Weibull distribution. The method is especially useful when the shape parameters are negative.

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

FFL	Failure Free Life
ML	Maximum likelihood
MLE	Maximum likelihood estimation/estimator/estimate (singular or plural)
MPE	Maximum probability estimation/estimator/estimate (singular or plural)
MPS	Maximum Product of Spacing
LSE	Least squares estimation/estimator/estimate (singular or plural)
MCMC	Markov chain Monte Carlo
MCMCE	Markov chain Monte Carlo estimation/estimator/estimate
WPP	Weibull probability plot
TTT	Total time on test
PDF	Probability density function
CDF	Cumulative distribution function
SF	Survival Function
FRF	Failure rate function
MRL	Mean residual life function

Chapter 1. Introduction

Probabilistic and stochastic models have been derived and used to model randomness of real life problems, such as the Bernoulli distribution to model winning times in a gamble and the Poisson distribution to model arrivals of buses in a crossing during a time interval. Ever since the introduction of the Weibull distribution by Professor Waloddi Weibull (Weibull, 1939) and the fitting of the distribution to some field data (Weibull, 1951), the Weibull distribution has been extensively studied and applied to model physical attributes of systems or parts of systems, especially failure times.

Using Weibull analysis techniques to investigate the life mechanism of a system starts with gathering failure data of the system in concern, exploring the data, finding a suitable probabilistic distribution, possibly a Weibull related distribution, to model the data, estimating the model parameters, and finally making descriptions of the unknown or future life behavior of the system.

The reason that the Weibull distribution is favored as a good alternative for modeling life data mainly relies on of its flexibility. It can exhibit three different kinds of failure rates – increasing, constant and decreasing – which are the elementary components of any real life failure rate. Failure rate evaluates the proneness of a system to fail as time goes by, so it is often an important indicator which attracts attentiveness. An increasing failure rate suggests a "better new than old" life mode for a system or that the system is currently within its wear-out period of the life cycle. A constant failure rate means that the system is "as good as new" or it is undergoing a period when failures only come from random events rather than systematic change of the system quality. A decreasing failure rate hints a "better old than new" life mode for the system or that failures result from "infant mortality" and the failure rate decreases since defective items are moved out from the population.

However, in many cases, the life behavior of mechanic or electronic systems cannot be suitably described by a monotonic failure rate. Instead, some other patterns of failure rate such as upside-down unimodal shape and bathtub or "U" shape are frequently encountered. Bathtub shaped failure rate is very common among the life modes of modern systems, such as computer processors. A typical bathtub curve composes of three phases: the first part is monotonically decreasing, known as infant mortality; the second part is constant at a relatively low level, known as random failure period; and the last part is monotonically increasing, known as wear-out period. When the system exhibits a unimodal or bathtub shaped failure rate, the Weibull distribution is not able to adequately model the life behavior. In such case, more sophisticated models are needed.

A simple generalization of the Weibull distribution can be done by model mixture (Mendenhall and Hader, 1958; Kao, 1959; Castet and Saleh, 2009), risk competing (David, 1970), model multiplying (Jiang and Murthy, 1995(1), 1997), or Weibull sectioning (Kao, 1959; Mann *et al*., 1974). Compared to these manipulations involving more than one Weibull distribution, in recent years, a few extensions, of the Weibull distribution have been proposed and applied to life time data analysis, such as inverse Weibull (Drapella, 1993), exponentiated Weibull (Mudholkar and Srivastava, 1993, 1995), generalized Weibull (Mudholkar and Kollia, 1994; Mudholkar *et al*., 1996), additive Weibull (Xie *et al*. 1996), extended Weibull (Marshall and Olkin, 1997; Nandi and Dewan, 2010), Weibull extension (Xie *et al*., 2002), modified Weibull (Lai *et al*., 2003), odd Weibull (Cooray, 2006), and flexible Weibull (Bebbington *et al*., 2007(1)), etc. Except the inverse Weibull, these newly proposed models commonly have 3 model parameters, with one additional parameter to the traditional 2-parameter Weibull distribution, and because of their non-piecewise and non-log-piecewise properties, parameters of these models based on complete or censored sample data are able to be estimated in a statistical point of view. All these generalization models of the Weibull distribution, together with the traditional 2 or 3-parameter Weibull, form a family named the "Weibull family", and all these models are called in a joint name "Weibull models" (Murthy *et al*., 2004(1)).

In using the Weibull models to model system life, a very important step is to estimate the model parameters based on a sample data. Except for methods which are universally used for all statistical distributions, such as maximum likelihood estimation (MLE) (Cohen, 1965; Lemon, 1975; Yang and Xie, 2003; Tang *et al*., 2003; Ng, 2005; Carta and Ramirez, 2007; Yang and Lin, 2007; Balakrishnan and Kateri, 2008; Jiang *et al*., 2010; Krishnamoorthy *et al*., 2009; Tan, 2009), Bayesian estimation (Nassar and Eissa, 2004; Kaminskiy and Krivtsov, 2005; Pang *et al*., 2005; Singh *et al*., 2005; Banerjee and Kundu, 2008; Gupta *et al*., 2008; Jiang *et al*., 2008(1); Kundo, 2008; Zhao *et al*., 2008; Touw, 2009), moment estimation (White, 1969; Cran, 1988; Rekkas and Wong (2005), Cao, 2005; Gaeddert and Annamalai, 2005; Nadarajah and Gupta, 2005; Merganič and Sterba, 2006; Nadarajah and Kotz, 2007; Carrasco *et al*., 2008), and percentile estimation (Dubey, 1967; Wang and Keats, 1995; Chen, 2004; Marks, 2005; Cao and McCarty, 2006; Chen and Chen, 2009), a graphic method called WPP (Weibull probability plot) is very popular for Weibull models. Contrasting to the other estimation methods stated above, as a graphic realization of least squares estimation (LSE), WPP is easy to implement and hence is appreciated among practitioners. Early contributions of this method track back to Weibull (1951), and Benard and Bos-Levenbach (1953). Recent discussions of WPP and LSE can be found in Hossain and Howlader (1996), Lu *et al*. (2004), Zhang *et al*. (2006)(1), Zhang *et al*. (2006)(2), Zhang *et al*. (2007), Jiang *et al*. (2008)(2), Jukić *et al*.

(2008), Tiryakioglu and Hudak (2008), Cousineau (2009), Marković *et al*. (2009), Bhattacharya and Bhattacharjee (2009), etc.

The Weibull distribution often cited by researchers is the 3-parameter Weibull distribution, while the "standard Weibull distribution" (page 10, Murthy *et al*., 2004[(2)]), with the location parameter equal to 0, is the 2-parameter special case. However, there are usually no rigorous different notations for the two distributions, because if the location parameter is known, the 3-parameter Weibull distribution can be shifted horizontally to the 2-parameter Weibull distribution, and as such many authors do not intentionally use "2-parameter" or "3-parameter" to discriminate the two distributions in their works, as long as no confusion will be caused. In the rest of the thesis, "the Weibull distribution" specifically denotes the 2-parameter Weibull distribution, unless otherwise stated.

1.1 Modeling of the Weibull Models to Life Data

The Weibull models, including the Weibull distribution and the generalizations of the Weibull distribution, are useful for modeling life data with different failure rates. As stated in Murthy *et al*. (2004)[(1)], a typical empirical modeling process involves three steps:

1. Model selection;
2. Estimation of model parameters;

3. Model validation description.

The model selection step is important as one requires a thorough preliminary analysis of the data and good understanding of the candidate models so that he is able to able to find out the most appropriate model to fit the data and model the life mode of the system.

Effective model selection is composed of two sides, data side and model side. On the data side, one usually carries out a preliminary analysis with the data, including computing a few sample statistics and drawing some different plots to measure the variability and pattern of the data. TTT (total time on test) and WPP are such useful tools for the Weibull models. According to Barlow and Campo (1975) and Bergman and Klefsjo (1984), the shape of the failure rate curve of the data uniquely determines the shape of the empirical TTT plot, and thus from the TTT plot one can know whether a model with a monotonic, unimodal or bathtub shaped failure rate is suitable for the data. The other plot WPP was originally developed for the Weibull distribution, but has since been used for all Weibull models. WPP makes a simple transformation on the data and the empirical probability, detects the discrepancy of the sample data against the Weibull distribution, and obtains estimates of the parameters through trial-and-error (if the discrepancy is small enough) or assist selecting a model from rest of the Weibull family (if the discrepancy is large).

On the model side, one needs to get a clear picture of the statistical characteristics of the candidate models to decide which of the models are suitable for modeling the given sample data and how the models can be used for the purpose of application, including estimation and prediction. Besides the basic statistics such as mean, variance and modes, characteristics of statistical models include the shapes of probability density function (PDF), failure rate function (FRF) and mean residual life (MRL), as well as some statistical inference procedures and goodness-of-fit tests. For the Weibull models, FRF (Murthy *et al*., 2004[(1)]) and MRL (Lai *et al*., 2004; Xie *et al*., 2004) are useful pattern indicators. FRF figures the risk of immediate failure at any time and if relating the shape of it to the empirical TTT plot of sample data, the appropriateness of modeling using the distribution can be roughly verified. Compared to FRF, MRL summarizes the trend of residual life, and has special importance if remaining using time of the system is of interest or in actuarial study where human life expectancy is crucial to life insurance policies.

Lai *et al*. (2003) proposed the modified Weibull distribution by introducing another shape parameter to the traditional Weibull distribution. The distribution has an advantage of being able to model bathtub shaped failure data, and the model parameters can be estimated easily based on WPP. Lai *et al*. (2004) studied the shapes of FRF and MRL of the distribution and claimed that the "model is very flexible for modeling different reliability situations". In the other research paper focusing on the relationship between FRF and MRL of several generalizations of the Weibull distribution, Xie *et al*. (2004) delved in the change points of the two functions and calculated the length of the flat portion of FRF under different parameter sets. Regarding parameter estimation, Ng (2005) discussed ML estimation and confidence intervals of the modified Weibull parameters for progressively type-2 censored samples, and concluded that MLE performs better than LSE based on a simulation study. In Bebbington *et al*. (2008), the authors obtained the estimate of the turning point of MRL via first estimating the model parameters using MLE method. Carrasco *et al.* (2008) proposed a regression model considering this modified Weibull distribution. Despite the volume of available works on the modified Weibull distribution, an overall statistical characterization which is useful for application and referencing is still lacking. In the first part of this thesis, a systematic study of the statistical characteristics and parameter estimation procedures is carried out.

As a newly proposed generalization of the Weibull distribution, the odd Weibull (Cooray, 2006) has been shown to be able to exhibit monotonic, unimodal and bathtub shaped failure rate. Another favorable merit of the model is that when its FRF is bathtub shaped, the second portion of curve could be quite flat and long, which is a good property in application. However, its complicated form of PDF makes ML estimation of the model parameters not stable, sometimes even unreachable. In such case, a good graphic method can help find acceptable estimates of the model parameters, for application or starting point of further investigation. In this thesis, a statistical characterization of the odd Weibull distribution is done, and a graphic parameter estimation method is proposed to replace WPP.

1.2 Observed Fisher Information Matrix and Uniqueness of MLE

The Fisher information, firstly introduced by R. A. Fisher in the 1920s, is the amount of information in a single sample about the unknown parameters of the distribution, or the likelihood function. When considering estimation of the model parameters, from the Cramer-Rao inequality, the inverse of the Fisher information matrix is the lower bound of the error variance of the unbiased estimators of the parameters of the given distribution, and is the asymptote of the variance-covariance matrix of MLE of the model parameters under some regularity conditions. However, for many statistical distributions, the calculation of the Fisher information matrix could be quite troublesome because of the complexity of PDF and the high dimensionality of the parameter vector. In such case, the Fisher information matrix is usually replaced by its approximate at the MLE point, the Observed Fisher Information matrix, which is the inverse matrix of the minus second derivatives of the log-likelihood function. Compared to the Fisher information matrix, the Observed Fisher Information matrix is relatively easy to calculate and meaningful in real application (Efron and Hinkley, 1978).

Gertsbakh and Kagan (1999) proved that the Weibull distribution can be characterized by the Fisher information lack-of-memory property in type-1 censored data. Zheng (2001) obtained a similar result in type-2 censored data case by expressing the Fisher information matrix of the Weibull distribution using FRF. Zheng and Park (2004) extended the result to multiply censored and progressively censored data. Gupta and Kundu (2006) compared the Fisher information matrix of the generalized exponential (GE) and Weibull distributions for complete and type-1 censored data, observed that due to right censoring the loss of information of the Weibull distribution is much more than the GE model, and concluded that for some data sets if the asymptotic variances of the median estimators and the average asymptotic variances are of interest, the GE distribution is preferred to the Weibull distribution. Borzadaran *et al*. (2007) derived entropy, variance, Fisher information, and analog of the Fisher information for some Weibull known families, including the Weibull family, and set up links between the measures for the families.

An issue related to ML estimation of the model parameters of statistical distributions is the existence and uniqueness of the estimators for a given sample data. A simple transformation on the likelihood equations of the Weibull distribution was used in Farnum and Booth (1997) to prove the existence and uniqueness of MLE of the model parameters. Wang and Fei (2003) proved in a tampered failure rate model, MLE of the shape parameter of the Weibull distribution exists and is unique. Mittal and Dahiya (1989) showed that MLE do not always exist for the truncated Weibull distribution. The MLE of the log-logistic parameters for right censored sample data were proven to uniquely exist in Gupta *et al*. (1999), and the result was generalized to grouped data case in Zhou *et al*. (2007). A similar result was obtained for the Normal distribution in Balakrishnan and Mi (2003), and in Mi (2006) the discussion was even extended to location-scale distributions for complete and partially grouped data.

Existing literature on the Fisher information of the Weibull distribution mainly focuses on the relationship between the Fisher information matrix and the Weibull distribution properties. The description of the matrix and the calculation involved in approaching MLE of parameters of the Weibull models remain untouched. In addition, although the existence and uniqueness of several 2-parameter distributions have been studied, no research work is available for multi-parameter distributions, such as 3-parameter generalizations of the Weibull distribution. A study taking into account the calculation of the elements of the Observed Fisher Information matrix and the relationship between this matrix and the property of MLE of the parameters of some Weibull models would be worthwhile. In this thesis, a technique of simplifying the calculation involved in the Observed Fisher

Information matrix and the accompanying application in proving the existence and uniqueness of MLE will be narrated and illustrated.

1.3 Bayesian Estimation and MCMC Algorithm

Bayesian theory suggests inferring truth of the probability of a statistical model by updating information in light of new observations on the base of a prior. Following this theory, Bayesian estimation of parameters of statistical distributions involves a prior of the parameters and a posterior with data information added in.

Bayesian estimation for the scale and shape parameters of the Weibull distribution was developed in Canavos and Tsokos (1973) by assuming independent prior distributions. The authors compared the performance of the Bayesian estimators and MLE through a simulation study and found that MSE (mean squared error) of Bayesian estimators are significantly smaller than those of MLE. For the 3-parameter Weibull distribution, Smith and Naylor (1987) pointed that ML estimation are not stable in the sense that small changes in the likelihood may correspond to large changes in the parameters, while the choice of priors does not make much influence on the Bayesian estimates as long as the priors are flat enough. Because of the mathematical intractability of the posterior expectations of the parameters of the 3-parameter Weibull distribution, Sinha and Sloan (1988) proposed the use of Bayesian linear estimators to approximate. Tsionas (2002) considered Bayesian estimation of the parameters and reliability function of the Weibull mixture distribution. Nassar and Eissa (2004) and Singh *et al*. (2005) discussed the problem of Bayesian parameter estimation under LINEX loss functions for the exponentiated Weibull distribution. Touw (2009) presented a study on Bayesian estimation for parameters of mixed Weibull models.

In many cases, when PDF of the statistical distribution is complex, obtaining the Bayesian estimates of the model parameters by direct calculating the posterior expectations is very time consuming or coarse, e.g. when the sample size is large and the posterior PDF of the parameters are so steep that integration over the parameter space is subject to substantial error. In such case, an algorithm called MCMC (Markov chain Monte Carlo) is useful. MCMC methodology provides a convenient and efficient way to sample from a high dimensional distribution, and obtain estimates of the parameters from the Markov chain formed. Following MCMC algorithm, Green *et al*. (1994) modeled tree diameter data with the 3-parameter Weibull distribution and indicated the advantage of MCMC to MLE that the former guarantees a positive estimate for the location parameter while the latter does not. Pang *et al*. (2001) dealt wind speed data with the 3-parameter Weibull distribution using MCMC techniques and highlighted the flexibility of the method that any quantity of interest regarding the distribution or parameters can be easily processed under the frame. Bayesian estimation via MCMC sampling of the coefficient of variation for the 3-parameter Weibull distribution was studied in Pang *et al*. (2005). Gong (2006) estimated mixed Weibull distribution parameters using SCEM-UA method adopting MCMC theory, and showed that the estimates of the parameters are more accurate than MLE for the automotive data. Gupta *et al*. (2008) used MCMC method to estimate the model parameters of the Weibull extension distribution. As an application in clinical study, Zhao *et al.* (2008) constructed Bayesian model for the Weibull distribution and used MCMC simulation method to estimate the model parameters.

Despite the advantage of the Bayesian estimation stated in the literature, for the Weibull models except the traditional 2 or 3-parameter Weibull, this estimation method is not extensively used. In this thesis, Bayesian estimation of the parameters of the modified Weibull distribution is studied by adopting MCMC theory, and the estimation performance is compared with MLE.

1.4 Research Objective

The main purpose of this study is to develop a systematic statistical analysis, including parametric characterization and parameter estimation, of the modified Weibull distribution, which is a very useful generalization of the Weibull distribution. In addition, the odd Weibull distribution, another 3-parameter generalization of the Weibull distribution, will be studied in detail and applied to real life data analysis. As the most wildly used member of the Weibull family, the 3-parameter Weibull distribution always has intricacy in its parameter estimation. As such, a detailed literature survey of the available estimation methods will be done, and a discussion on the maximum probability estimation (MPE) for the distribution will be initiated.

Chapter 2 presents the general background of the Weibull models and some related topics such as properties of the models, application to life data and parameter estimation methods.

The modified Weibull has both the Weibull distribution and type-1 extreme value distribution as special cases, and is able to model increasing, decreasing, constant and bathtub shaped failure rate data. Several aspects of the distribution have been covered by researchers, but a comprehensive statistical analysis of the distribution is still lacking. In Chapter 3, a systematic structural analysis of the distribution is carried out and some interesting issues related to the modeling of the distribution to life data are explored. We also included the discussion of MPE of the parameters of the 3-parameter Weibull distribution as a section of this chapter.

When analyzing the properties of MLE of the parameters of statistical distributions, the Observed Fisher Information matrix, which is the approximate of the Fisher information matrix at the MLE point, is usually seen as the variance-covariance matrix of MLE. For the 3-parameter generalizations of the Weibull distribution, the calculation of the Observed Fisher Information matrix and related issues are seldom considered. However, as a matter of fact, a study of the simplification of this matrix does not only save calculation time, but also shed light to the variability of MLE of the parameters, as well as help look into the existence and uniqueness properties of the estimates. Chapter 4 conducts a general study of the Observed Fisher Information matrix for a class of distributions and the application of the result to the modified Weibull distribution to prove the existence and uniqueness of MLE of the model parameters for complete or progressively type-2 censored data. Using the techniques proposed, a study of the existence and uniqueness properties of the MLE of the modified Weibull distribution is carried out. The two properties are important because they ensure that usual optimization methods are able to locate the estimates that maximize the log-likelihood function, and statistical inferences can be drawn from the fact that the estimates are asymptotically normally distributed. To get the results, the parameter space is slacked before the analysis, and the non-negativity constraints are re-imposed afterwards.

Chapter 5 provides a Bayesian estimation of the parameters of the modified Weibull distribution. Bayesian methods have been shown in the literature to have some preferable qualities as compared to the MLE for the Weibull parameters. In this chapter, Gibbs sampler, as one of the MCMC simulation methods, is used to produce the Bayesian estimators of the model parameters. To overcome the difficulty in sampling from the posterior conditional distributions, a technique called adaptive rejection sampling is applied. The Bayesian estimators obtained in this way are compared with MLE, and they are shown to have smaller MSE than their counterparts.

After the study of modified Weibull distribution is completed, the properties of another recently proposed model, the odd Weibull distribution, are investigated in Chapter 6. A detailed statistical characterization of the distribution is done. WPP parameter estimation is carried out and shown to perform well. Burn-in and useful period related issues are discussed.

Chapter 7 concludes current research works and discusses some possible future research topics.

Chapter 2. Literature Review

2.1 Weibull Models

As quoted in Murthy *et al*. (2004)[1], basically there are two different approaches used for life data modeling, theory based modeling and empirical modeling. As the name stands, theory based modeling has the assumption that the mechanism of the system in research is known thoroughly or partially so that a theory based model which fits the life mode of the system can be built. However, due to the complicated manufacturing procedures of modern units and their multi-layer structures, mathematically and physically precise models for their life modes are impossible or very hard to construct. In such case, empirical modeling is useful for researchers to develop a suitable model for the system based on the information included in a given sample of data, or help look into the operation mechanism of the system so that a theory based model can be formed.

Empirical data modeling involves an explorative analysis of the data, and then choosing the suitable statistical distribution out of a number of candidate models. One of the most important families of such candidate distributions with wide applicability is the Weibull family. The first modeling of the Weibull distribution to engineering data dates back to Weibull (1951). Since the advent and popularity of the Weibull distribution in life data analysis prompted by Professor Weibull himself and the followers, in order to widen the applicability of Weibull analysis, many generalizations of the Weibull distribution, called the Weibull models, have been proposed and studied. These Weibull models can exhibit various shapes of FRF, not only monotonic but also unimodal and bathtub shaped, which are very common FRF shapes of modern mechanic and electronic units, such as computer processors.

The cumulative distribution function (CDF) of the 3-parameter Weibull distribution is as follows

$$F(t) = 1 - \exp\left\{-\left(\frac{t-\tau}{\alpha}\right)^{\beta}\right\},\ t \geq \tau \tag{2.1}$$

where $\alpha > 0$, $\beta > 0$ and $\infty > \tau > -\infty$ are called the scale, shape and location parameters respectively.

When the location parameter τ is equal to 0 or after the Weibull variable undergoes a horizontal shift of -τ, the 3-parameter Weibull distribution reduces to the Weibull distribution

$$F(t) = 1 - \exp\left\{-\left(\frac{t}{\alpha}\right)^{\beta}\right\},\ t \geq 0 \tag{2.2}$$

Estimation of the shape and scale parameters of the Weibull distribution has shown to be relatively easy. In contrast, with the inclusion of the additional location parameter τ, estimation of the parameters of the 3-parameter Weibull becomes much more complicated. Focus is on designing feasible and efficient estimation procedures recently. A survey on the various estimators and their properties will be presented in the final section of this chapter.

From the taxonomy of Murthy *et al*. (2004)[1], frequently used Weibull models can be roughly classified into 3 different types according to the different procedures of generalization, type-1 from direct transformation of the Weibull variable, type-2 from transformation of the Weibull distribution function, sometimes with one or more additional parameters, and type 3 involving more than one Weibull distribution or distribution from type-1.

Type-1 Weibull models are the basic members of the Weibull family, including the Weibull distribution (Weibull, 1939, 1951; Murthy *et al*., 2004[1]; Dodson, 2006), type-1 extreme value distribution (White, 1969; Kotz and Nadarajah, 2000; De Haan and Ferreira, 2006), and inverse Weibull distribution (Drapella, 1993; Khan *et al*., 2008). These distributions have been extensively studied and applied to practical application.

Type 3 Weibull models are composed of one or more Weibull or inverse Weibull distributions, members including Weibull or inverse Weibull mixture (Mendenhall and Hader, 1958; Kao, 1959; Chen *et al*., 1989; Jiang and

Kececioglu, 1992; Jiang and Murthy, 1995; Nagode and Fajdiga, 2000; Sultan *et al*., 2007; Mosler and Scheicher, 2008; Touw, 2009), Weibull or inverse Weibull competing risk (David, 1970; Jiang and Murthy, 1997(1), 2003; Davison and Neto, 2000; Jiang *et al*., 2001; Balasooriya and Low, 2004; Bousquet *et al*., 2006; Pascual, 2007, 2008), Weibull or inverse Weibull multiplicative (Jiang and Murthy, 1995, 1997(2)), and Weibull sectional (Kao, 1959; Mann *et al*., 1974; Jiang *et al*., 1999). These Weibull models are flexible at modeling life data, but due to the difficulty involved in analytic parameter estimation such as MLE, graphic parameter estimation methods resorting to WPP are often employed in practice.

Type-2 Weibull models are mostly newly proposed models. They are derived from the Weibull distribution, with one or more additional parameters, and therefore are able to exhibit a wider range of shapes of FRF. In addition, unlike type 3 Weibull models which contain coefficient parameters weighing the importance of the submodels, type-2 Weibull models do not have the difficulty in ML estimation procedure caused by estimating these parameters, so statistical properties of MLE and then other characteristics of the models can be studied conveniently and systematically. Because of these advantages of type-2 Weibull models, they attract a lot of research attention and application interest. The main part of the thesis will be centered on some of type-2 Weibull models, so in the next section a detailed literature review on the relevant models will be given, but before that we will briefly survey the existing research on the other models. To highlight the relationship between the models and the Weibull distribution, we use $G(t)$ to denote CDF of the models and $F(t)$ to denote CDF of the Weibull distribution.

2.1.1 Exponentiated Weibull

The exponentiated Weibull distribution was proposed by Mudholkar and Srivastava (1993). CDF of the distribution is

$$G(t) = [F(t)]^{\nu} = \left\{1 - \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right]\right\}^{\nu}, \ t \geq 0 \tag{2.3}$$

with $\nu > 0$ the additional parameter. When $\nu = 1$, the exponentiated Weibull reduces to the Weibull distribution.

As stated in Mudholkar *et al*. (1995), FRF of the exponentiated Weibull distribution can exhibit monotonic, unimodal and bathtub shapes. Statistical properties and parametric characterization of the distribution were investigated in Mudholkar and Hutson (1996) and Nassar and Eissa (2003). Nadarajah and Gupta (2005) and Choudhury (2005) considered the derivation of the moments. Ashour and Afify (2007) considered the analysis under type-1 progressive interval censoring and derived the ML estimators and the corresponding asymptotic variances. Jiang and Murthy (1999) presented a graphic study of the distribution and proposed to use WPP to estimate the model parameters. Bayesian parameter estimation was studied in Cancho *et al*. (1999), Cancho and Bolfarine (2001), Nassar and Eissa (2004), and Singh *et al*. (2005). Ortega *et al*. (2006) adapted local influence methods to detect influential observations with exponentiated Weibull regression models for censored data. As a practical application in software reliability study in Ahmed *et al*. (2008), the exponentiated Weibull distribution was incorporated into the modeling process and the researchers found that the proposed software reliability growth model is wider and effective SRGM.

2.1.2 Generalized Weibull

Mudholkar and Kollia (1994) proposed a generalization of the Weibull distribution, which they called the generalized Weibull family. CDF is

$$G(t) = 1 - \left[1 - \lambda(1 + \alpha t)^{1/\alpha}\right]^{1/\lambda} \tag{2.4}$$

where $-\infty < \alpha, \lambda < \infty$.

Another slightly different version with three model parameters was proposed in Mudholkar *et al*. (1996)

$$G(t) = 1 - \left[1 - \lambda(t/\sigma)^{1/\alpha}\right]^{1/\lambda} \tag{2.5}$$

where $\alpha, \sigma > \infty$, $-\infty < \lambda < \infty$. This model reduces to the Weibull distribution as $\lambda \to 0$.

According to Mudholkar and Kollia (1994) and Mudholkar *et al*. (1996), the supports for both CDF are dependent on the sign of the parameters, instead of invariably on the positive real line. Regarding the shape of FRF, Mudholkar *et al*. (1996) indicated that FRF of the latter model can exhibit monotonic, unimodal and bathtub shapes. However, there seems that no discussion on the shape of FRF of the former model (2.4) is available yet.

2.1.3 Additive Weibull

The additive Weibull distribution was proposed by Xie and Lai (1996). CDF is

$$G(t) = 1 - \exp\left\{-(at)^b - (ct)^d\right\},\ t \geq 0 \tag{2.6}$$

where $a, c \geq 0$, $b > 1$, $0 < d < 1$. The model reduces to the Weibull distribution when either a or c equals to 0.

FRF of this model is not only able to be monotonic, but also bathtub shaped.

The additive Weibull distribution is essentially a special case of the 2-component Weibull competing risk distribution of which one shape parameter is larger than 1 and the other smaller than 1, but its good property in describing bathtub shaped failure rate data and the application of the simplified version makes it an important generalization of the Weibull distribution, and so we put it here as a member of the type-2 Weibull models for a more detailed literature survey.

Motivated by the idea of Xie and Lai (1996), Wang (2000) proposed the additive Burr XII distribution, which is also able to describe bathtub shaped failure rate data. Lai *et al*. (2004) recommended adding a constant random failure term to the additive Weibull distribution to achieve a better fit to some data. Bebbington *et al*. (2006) proposed using the curvature of FRF to evaluate the length of the useful period for a bathtub curve of the additive Weibull. Bebbington *et al*. (2007[(2)]) showed that the addition of a constant competing risk to the additive Weibull can lead to complex effects on the mean residual life, which may be of great use in actuarial and reliability studies.

2.1.4 Extended Weibull

The extended Weibull distribution was proposed by Marshall and Olkin (1997). CDF is

$$G(t) = \frac{F(t)}{F(t) + \nu \bar{F}(t)}, \; t \geq 0 \tag{2.7}$$

where $F(t)$ and $\bar{F}(t)$ are a Weibull CDF and its corresponding survival function (SF), and ν is the additional parameter. This model reduces to the Weibull distribution when $\nu = 0$.

Marshall and Olkin (1997) studied FRF of the extended Weibull distribution and showed that except for monotonic shape, FRF of the model can exhibit increasing-decreasing-increasing and decreasing-increasing-decreasing shapes. However, no exactly unimodal or bathtub shaped failure rate curve can be achieved. Marshall and Olkin (1997) proved the geometric-extreme stability property of the model, which could be a favorable feature for practical application. Hirose (2002) derived log-likelihood function and likelihood equations for the extended Weibull distribution and showed the usefulness of the model for fitting breakdown voltage data. Ghitany *et al*. (2005) presented another derivation of the model and discussed the application to censored data. Adamidis *et al*. (2005) proposed to use EM algorithm to estimate the model parameters when $F(t)$ is an exponential CDF. Sankaram and Jayakumar (2007) showed that the extended Weibull distribution satisfies the property of proportional odds function and then gave a physical interpretation of the model. Zhang and Xie (2007) described a graphic parameter estimation method for the model, and discussed application related issues such as burn-in time and replacement time determination.

Motivated by the idea of Marshall and Olkin (1997), Jayakumar and Mattew (2006) extended the Burr type-2 distribution, Ghitany *et al*. (2007) the Lomax distribution, and Ghitany and Kotz (2007) the linear failure rate distribution.

2.1.5 Weibull Extension

The Weibull extension distribution was proposed by Xie *et al*. (2002). CDF is

$$G(t) = 1 - \exp\left\{\lambda\alpha\left[1 - e^{(t/\alpha)^\beta}\right]\right\},\ t \ \geq\ 0 \tag{2.8}$$

where $\lambda, \alpha, \beta > 0$.

When $\alpha = 1$, the model reduces to the 2-parameter model of Chen (2000); when $\alpha \to \infty$, the model reduces to the Weibull distribution.

Xie *et al*. (2002) showed that FRF of the Weibull extension distribution can exhibit monotonic and bathtub shapes. Tang *et al*. (2003) carried out a detailed statistical analysis of the Weibull extension distribution. Nadarajah and Gupta (2005) derived explicit algebraic formulas for the moments of the distribution. Wu *et al*. (2004) proposed an exact statistical test for the shape parameter of the model of Chen (2000). Gupta *et al*. (2008) carried out a Bayesian estimation of the model parameters using Markov chain Monte Carlo simulation, and observed that the in spite the model cannot provide good fit to the higher order observations which are responsible for the increasing part of the hazard rate, it behaves quite well overall.

2.1.6 Flexible Weibull

Bebbington *et al*. (2007[1]) proposed the flexible Weibull distribution. CDF is

$$G(t) = 1 - \exp\left(-e^{\alpha t - \beta/t}\right),\ t > 0 \tag{2.9}$$

where $\alpha, \beta > 0$. When $\beta = 0$, the model reduces to the Type-1 extreme value distribution, and thus may be regarded as a generalization of the Weibull distribution.

Bebbington *et al*. (2007[1]) proved that FRF of the distribution can exhibit increasing, increasing average, and increasing-decreasing-increasing, called modified bathtub, shapes. Bebbington *et al*. (2007[2]) constructed a competing risk model involving a flexible Weibull distribution and an exponential distribution, and showed that the new model performs well for human mortality data.

2.1.7 Model by Dimitrakopoulou et al. (2007)

Dimitrakopoulou *et al*. (2007) proposed a 3-parameter generalization of the Weibull distribution. CDF of the distribution is

$$G(t) = 1 - \exp\left(1 - \left(1 + \lambda t^{\beta}\right)^{\alpha}\right),\ t > 0 \tag{2.10}$$

where $\alpha, \beta > 0$ are shape parameters and $\lambda > 0$ is a scale parameter. The model reduces to the Weibull distribution when $\alpha = 1$.

As stated in the above, the motivation of the distribution comes from evaluating the reliability of a series system. FRF of the distribution was shown to be able to exhibit monotonic, unimodal and bathtub shapes. Likelihood equations were derived.

2.2 Modified Weibull and odd Weibull

2.2.1 Modified Weibull Distribution

As one of the type-2 Weibull models, the modified Weibull distribution (Lai *et al*. (2003)) attracts some interest among researchers and practitioners because of its ability in modeling bathtub shaped failure rate data, simplicity and flexibility of FRF and ease of handling parameter estimation using least squares method. CDF of this distribution is

$$G(t) = 1 - \exp\left\{-at^b e^{\lambda t}\right\},\ t \geq 0 \tag{2.11}$$

where $a > 0$, $b, \lambda \geq 0$.

The distribution function once appeared in an earlier paper Gurvich *et al*. (1997), but with different parameterization. In the paper, the first and second moments of the distribution were derived, but without explicit forms, and a least squares parameter estimation method was formulated. However, the model is not exactly the same as the modified Weibull since Gurvich *et al*. (1997) did not confine the parameter λ to be positive. When $\lambda < 0$, the support of the CDF does not cover the positive half of the real line, but only a portion (from 0 to a finite value). In fact, negative estimates of λ were yielded in Gurvich paper when fitting the model to glass fiber data, which was the motivation of the research. Therefore, it is inaccurate to say the two models are identical (Nadarajah and Kotz, 2005).

According to Lai *et al*. (2003), the distribution is able to model monotonic and bathtub failure rate data. The model has the Weibull and Type-1 extreme value distributions as special cases, and is an approximation of the Beta-Integrated distribution in the limit case. Lai *et al*. (2003) depicted FRF and WPP plotting for this distribution, and suggested a multiple linear regression method to estimate the model parameters based on a sample data. The log-likelihood function and likelihood equations for complete data were derived in the paper, and MLE procedures were briefly stated. In Lai *et al*. (2004), the relationship between FRF and MRL was visibly demonstrated, and the modified Weibull distribution was claimed to be very flexible for modeling different reliability situations.

Ng (2005) carried out an MLE study of the model parameters for progressively type-2 censored data, and suggested transformed confidence intervals for the parameters based on asymptotic lognormality could achieve higher coverage probabilities than traditional confidence intervals based on asymptotic normality, since the parameters are assumed to be positive. Regarding the performances of parameter estimation methods, Ng (2005) showed that MLE performs better than LSE, for both bias and MSE. As to censoring schemes, progressively type-2 censoring schemes are better than conventional type-2 censoring, from the estimation point of view.

Xie *et al*. (2004) proposed the difference between the turning points of FRF and MRL as a measure of the flatness of the constant period of a bathtub curve, and showed that for the modified Weibull, this criterion is in line with another one which measures flatness with the length of the period whose failure rate is within a tolerance interval of the minimum failure rate.

Bebbington *et al*. (2008) investigated the performance of MLE of the turning point of FRF for this distribution and constructed a confidence interval for the estimator based on asymptotic normality.

2.2.2 Odd Weibull Distribution

The odd Weibull distribution was recently proposed in Cooray (2006). CDF is

$$G(t) = \frac{F^{\theta}(t)}{F^{\theta}(t) + \bar{F}^{\theta}(t)}, \ t \geq 0 \tag{2.12}$$

where $-\infty < \theta < \infty$ is the additional parameter. If $\theta > 0$, $F(t)$ and $\bar{F}(t)$ are a Weibull CDF and the corresponding SF, while if $\theta < 0$, $F(t)$ and $\bar{F}(t)$ are inverse Weibull CDF and SF. The model reduces to the Weibull distribution when $\theta = 1$ and inverse Weibull distribution when $\theta = -1$.

As described in Cooray (2006), the model originated from evaluating the randomness of the odds of death using the Weibull distribution. It is capable of modeling monotonic, unimodal and bathtub shaped failure rate data, and the advantage of the model over some other Weibull models is that the second portion of the bathtub curve of FRF could be quite flat and long, which is a favorable feature for real data modeling.

Regarding parameter estimation, the log-likelihood function was set up for right censored data in Cooray (2006). The inverse property of MLE was shown, which could be useful when the odd Weibull densities are non-unimodal. TTT transforms were employed to help determine the shape of FRF and test the goodness-of-fit against exponentiality null hypothesis, and a simulation study was done to tabulate the upper percentage points of the TTT test statistic.

Despite the good properties of the distribution, except Cooray (2006), few works have been done to its statistical characteristics and parameter estimation, partly because the form of the distribution function and failure rate function are complicated. In order to cover this gap, a study of the important statistical characteristics and an alternative graphic parameter estimation method is presented in this thesis.

2.3 Parameter Estimation Methods

The Weibull models are widely applied to life data analysis for all kinds of systems. As a normal procedure for empirical modeling, after a model is chosen to characterize the life mode of the system, model parameters are estimated with a given sample data set.

Sample data could be complete or censored. There are several different classifications of censoring (Murthy *et al*., 2004[2])

- Right, left, or interval censoring
- Type-1 time based censoring, type-2 failure based censoring, or random censoring
- Single or multiple censoring

Censoring is an important and often used technique in practice to save time and cost, yet effective in investigating the life behavior of a system. In the perspective of reliability study, right type-1 and type-2 censoring, random censoring, interval censoring (grouped data) are of particular interest.

To handle data with different failure patterns and censoring schemes, for different models in assumption, a few traditional parameter estimation methods have been applied and studied for the Weibull models, as well as some initiative methods been developed.

Priority of the estimation methods is not fixed. For a particular model with a certain data censoring scheme, a method may behave better than others in consideration of bias, but performs poorer in the criterion of dispersion. A method could produce an estimate that is both accurate and precise, but the difficulty and computing time involved in the search of the estimate may offset the benefit. In a word, none of the estimation methods is universally and overwhelming best. For the Weibull models, a lot of research has been done to study the properties of the estimators, compare the advantages and disadvantages of the estimation methods, and providing solutions that could increase the performance. In the following part of this section, we give a literature survey on this subject.

2.3.1 Method of Moments

The method of moments makes use of sample moments such as mean and variance to estimate the model parameters.

The mean and variance of a Weibull variable are Gamma functions of the model parameters, so it is impossible to give moment estimators for the parameters with closed forms. However, iterative procedures can be applied to numerically obtain the estimates. Situation is similar for the 3-parameter Weibull distribution, except that the third moment skewness needs to be evaluated.

A critical drawback of the method of moments for parameter estimation is its inability for censored data (Murthy *et al*., 2004[(2)]). This disadvantage hinders its application in lifetime study.

For discussions on this topic, see Menon (1963), Cohen *et al*. (1984), Cran (1988), Rekkas and Wong (2005), Cao (2005), Gaeddert and Annamalai (2005), Nadarajah and Gupta (2005), Merganič and Sterba (2006), Nadarajah and Kotz (2007), Carrasco *et al*. (2008) etc.

2.3.2 Method of Percentiles

Another parameter estimation method initiating from the relationship between the distribution functions and the model parameters is the method of percentiles.

To estimate the two parameters of the Weibull distribution, at least two percentiles are needed, while for the 3-parameter Weibull, at least three are needed. Hassanein (1971) obtained the best linear unbiased estimates with 2, 4 and 6 sample percentiles and discussed optimum spacing of the sample percentiles for the Weibull distribution assuming the shape parameter known.

Selection of the percentiles is pivotal to the performance of the estimators. Under different conditions, percentiles estimators with different probabilities were discussed in Menon (1963), Dubey (1967), Zanakis (1979), Wang and Keats (1995) and Marks (2005).

Chen (2004) and Chen and Chen (2009) employed the pivotal property of the parameters embedded in the relationships of percentiles of the 3-parameter Weibull distribution, and proposed a simulation based method of constructing confidence intervals and point estimate of the location parameter.

2.3.3 Maximum Likelihood Estimation

Since the invention of this method by Sir. R. A. Fisher in the early part of last century, ML estimation has been one of the most popular parameter estimation methods for statistical distributions. The idea of MLE is to maximize the likelihood function, which is the joint probability function of the available data under a statistical distribution whose parameters are unknown, by changing the values of the model parameters, and then find the parameter estimates.

Under mild regularity conditions, the MLE has the inverse of the Fisher information matrix as the asymptotic variance-covariance matrix. Statistical inferences can be drawn from the normal distribution assumption of the estimators, and thus confidence intervals, statistical significance tests, etc can be constructed. In addition, from the Cramer-Rao theory, the variance of a MLE is asymptotically the smallest, as compared to other unbiased estimators. Another advantage is that unlike the two methods reviewed above, MLE works well with censored data.

For the Weibull distribution, some earlier works (Kao, 1958; Cohen, 1965; Thoman *et al*., 1970; Engelhardt and Bain, 1974; etc) built the foundation of ML estimation procedures and properties. Among them, Cohen (1965) derived the likelihood function and equations for complete and right censored sample data, recommended iterative procedures to solve the likelihood equations, and used the Observed Fisher Information matrix to approximate the variance-covariance of the estimators. Thoman *et al*. (1969) and Watkins (1996) investigated the bias of MLE of the Weibull parameters and Ross (1996), Montanari *et al*. (1997), Hiross (1999), Yang and Xie (2003), Yang *et al*. (2003, 2007), and Ferrari *et al*. (2007) presented several methods to reduce the bias.

For the 3-parameter Weibull distribution, Harter and Moore (1965) derived the log-likelihood function and likelihood equations, and indicated that when the value of the shape parameter is less than 2, MLE may not exist because of the irregularity of the likelihood function. Also see Blischke (1974) for further details. To deal with the difficulty in convergence of MLE searching, several techniques and modified estimators were proposed (Lemon, 1975; Cohen and Whiten, 1982; Balakrishnan and Kateri, 2008; Cousineau, 2009[1]; etc).

ML estimation is widely used for other models of the Weibull family. For more details, see the literature listed under the specific distributions in Section 2.2.

Another topic regarding MLE with theoretic and application importance is the existence and uniqueness of the estimators. For the Weibull distribution, a simplification of the likelihood equations can easily lead to the proof of this property of MLE (e.g. see Farnum and Booth (1997)). Wang and Fei (2003) proved the MLE of the shape parameter of the Weibull distribution is unique in a multiple step-stress accelerated life test. To overcome the difficulty of ML estimation, Hirose and Lai (1997) reparameterized the 3-parameter Weibull distribution and embedded it in a large family. However, for distributions with more complicated CDF, other techniques are needed.

Makelanen *et al*. (1981) proved that to verify the fact that MLE of the parameters of a distribution exist and are unique, one just needs to show that the Hessian matrix of the log-likelihood is negative definite and the likelihood is constant on the boundary of the parameter space. Following this track, Gupta *et al*. (1999) proved MLE of the parameters of the log-logistic distribution for right censored sample data uniquely exist. Zhou *et al*. (2007) generalized the result to the same distribution for grouped data. A similar result was obtained for the Normal distribution in Balakrishnan and Mi (2003). Mi (2006) even extended the discussion to a much broader distribution class, location-scale families, for complete and partially grouped data.

However, till now there have been no works on the existence and uniqueness properties of the ML estimators of the parameters of Weibull generalization models, which we think poses a theoretical necessity. This consideration leads to the study of the MLE of the parameters of the modified Weibull distribution presented in this thesis. It should be noted that the approach we proposed could easily be extended to other 3 parameter generalization models of the Weibull distribution.

2.3.4 Least Squares Estimation and Weibull Probability Plot

LSE is obtained by minimizing the sum of squared errors between the sample data and the fitted distribution function. This estimation method is very popular for model fitting, especially in linear and non-linear regression. As a graphic realization of LSE, WPP is easy for implementation and so attract a lot of interest among practitioners and researchers. The first application of WPP appeared in Weibull (1951), in which the parameter estimation of the case studies was done by manual curve fitting.

Due to the importance in application, LSE and WPP have been studied extensively (Kao, 1959; Bain and Antle, 1967; Hossain and Howlader, 1996; etc). However, similar to MLE, LSE for Weibull parameters are always biased. To decrease the bias, several approaches have been tried by researchers (White, 1969; Bergman, 1986; Lu *et al*., 2004; Hung and Liu, 2005; Wu *et al*., 2006; Zhang *et al*., 2006(1), 2006(2), 2007; etc).

As a graphic approach, WPP has its advantages over analytical methods such as MLE and moment estimation, due to its visualization capability and even more importantly, the ease of implementation when numerical computations involved in the analytical methods are complicated or unstable. In this thesis, a trial-and-error WPP based method is proposed to estimate the parameters of the odd Weibull distribution, which we believe is of practical importance when the model is used in real data analysis.

2.3.5 Bayesian Estimation and Markov Chain Monte Carlo

From the view of frequentists, probability is interpreted as the degree to which an event is believed to happen. To estimate the model parameters of a statistical distribution, Bayesian approach starts with assigning a prior distribution to the parameters, and then calculates the estimates based on the posterior distribution in which the information contained in the realizations has been added.

The selection of the prior distribution affects the performance of the estimators. At the beginning of the Bayesian procedure, if some historical data of the system or expert opinion exist, one can incorporate these messages into the prior distribution to show some pre-knowledge of the parameters. When no or very little information about the parameters is known, uninformative priors are often used (see Gelman *et al*. (2004) for example).

For the Bayesian estimation of Weibull parameters, the prior distribution was extensively discussed (Canavos and Tsokos, 1973; Sinha and Sloan, 1988; Kaminskiy and Krivtsov, 2005; Zhang and Meeker, 2005; etc).

Usually numerical methods are needed when the posterior distribution of the parameters is hard to derive directly. Dellaportas and Wright (1991) discussed the problems for the Weibull parameters and proposed an approximation for the posterior mean with Gauss-Hermite method. See Singh *et al*. (2005) and Nassa and Eissa (2004) for the numerical evaluation of Bayesian estimates for the parameters of the exponentiated Weibull distribution.

Another approach to deal with this problem is resorting to the MCMC algorithm (Metropolis *et al*., 1953; Hastings, 1970).

MCMC methodology provides a convenient and efficient way to sample from complex, high dimensional statistical distributions. As one of the MCMC methods, Gibbs sampler generates a sequence of samples from the joint distribution of the random variables, for the purpose of approximating the joint distribution or computing expected values relating to the distribution (Gilks *et al*., 1995). This sampling scheme was first introduced by Geman and Geman (1984), but the applicability to statistical modeling for Bayesian computation was demonstrated by Gelfand and Smith (1990).

To see if the observations generated from the sampling does follow the underlying distribution after running long enough, the convergence of the Gibbs sampler has to be checked. Several evaluation methods have been proposed (Gelfand and Smith, 1990; Casella and George, 1992; etc).

The application of Gibbs sampler in parameter estimation of the Weibull models is very popular among recent years. Green *et al*. (1994) discussed parameter estimation for the 3-parameter Weibull distribution, and showed that when handling tree diameter data, ML estimation for the location parameter has a large chance to be negative, which contradicts the fact, while if given a proper prior distribution, the estimate generated from Gibbs sampling is always positive. Pang *et al*. (2007) claimed that MCMC is quite versatile and flexible for estimating the parameters of the 3-parameter Weibull distribution, and showed that these Gibbs estimators perform better than ML estimation, in the consideration of flexibility and the ease of constructing exact probability intervals. Pang *et al*. (2005) studied the interval estimation of the coefficient of variation (CV) for several statistical models, and indicated that the Gibbs estimators behave well even when the distribution is quite skewed, and the convergence of the Markov chain to a stationary process is reasonably fast. When both the Weibull and lognormal distributions are suitable to model a given data set, Upadhyay and Peshwani (2003) recommended using the Gibbs sampler to choose the right model through a simulation based Bayesian study. For the exponentiated Weibull distribution, Cancho and Bolfarine (2001) estimated the parameters and carried out model selection against other distributions via Gibbs sampling. Gupta *et al*. (2008) did a Bayesian analysis of the Weibull extension distribution and deployed a hybrid strategy to carry out the MCMC estimation. See Kottas (2006), Gupta *et al*. (2008), Kundo (2008), Zhao *et al.* (2008) for some other recent discussions.

An important step in Gibbs sampling is to sample from the posterior distribution. Since in most cases the distribution is so complicated that it is difficult or impossible to sample directly, rejection sampling techniques are required. For Gibbs sampling from a distribution which is complicated in form and evaluation of the density function is computationally expensive, Gilks and Wild (1992) introduced a rejection sampling scheme, called adaptive rejection sampling. This sampling technique is suitable for any univariate log-concave PDF. The advantage of adaptive rejection sampling is that it is adaptive: both the envelope function and the squeezing function converge to the target density function as sampling proceeds, and the reconstructions of the envelope function and the squeezing function only need negligible computational cost, thus it is very efficient compared to direct sampling or traditional rejection sampling.

Despite the volume of MCMC estimation of Weibull related model parameters, few systematic simulation works are available regarding the comparison of the efficiency and effectiveness between this method and the others. In this thesis, an application of MCMC estimation is carried for the modified Weibull distribution and a simulation design is made.

2.4 Parameter Estimation for 3-Parameter Weibull

As indicated previously, MLE of the parameters of the 3-parameter Weibull distribution does not satisfy the so-called regularity conditions, in the sense that when the shape parameter $\beta<1$, the likelihood function is not bounded so MLE do not exist, while when $1\leq\beta<2$, MLE of the parameters exist but do not follow an asymptotic normal distribution and are inefficient, and only when $\beta>2$, the distribution function is regular and MLE of the model parameters exist and are consistent (see Rockette *et al*. (1974), Smith (1985), Kantar and Senoglu (2008) and Cousineau (2009)).

There is a lot of literature on the estimation of the 3-parameter Weibull parameters. Earlier study dates back to Harter and Moore (1965), Dubey (1966, 1967), Lemon (1975), Zanakis (1979), Smith and Naylor (1987), etc. More recent works are mainly summarized in the review articles as follows.

Tang (2003) highlighted the practical importance of the failure-free life (FFL), which is essentially the value of the location parameter τ, of the 3-parameter Weibull distribution, and implemented two estimation procedures, D-Method from Drapella (1999) and LSE, using Excel^{TM} Solver in two case studies. The discussion was enlightening and showed that the proper identification of the presence of such factor is beneficial and may lead to in-depth findings of the underlying principle of the lifetime system.

Assuming the shape parameter known, Kantar and Senoglu (2008) treated the 3-parameter Weibull distribution with only two unknown parameters. The authors reviewed nine estimators, namely MLE, method of moments, maximum product of spacing (Cheng and Amin, 1983), modified MLE I, II (Cohen and Whitten, 1982), Tiku's modified MLE (Tiku, 1967), LSE, weighted LSE (Swain *et al.*, 1988) and percentile estimators (Kao, 1958).

Cousineau (2009) reviewed some estimation methods for fitting the 3-parameter Weibull distribution, namely maximum product of spacing (Cheng and Amin, 1983), weighted MLE (Cousineau, 2009[1]), method of moments (Harter and Moore, 1965), and claimed that all these methods are superior to MLE.

In the rest of this section, we will briefly review some of the abovementioned and other methods.

2.4.1 D-Method

In Drapella (1999) and O'Connor (2002), the following equation is obtained by introducing the WPP transformation $y_i = \ln\left(-\ln \hat{R}(t_i)\right)$, where $\hat{R}(t_i)$ is an estimate of the survivability at time t_i,

$$\frac{y_k - y_j}{y_l - y_k} \ln\left(\frac{t_l - \tau}{t_k - \tau}\right) - \ln\left(\frac{t_k - \tau}{t_j - \tau}\right) = 0 \tag{2.13}$$

where $j<k<l$.

It should be noted that in (2.13) there is only one unknown parameter τ. Upon solving for τ, the estimates of α and β can be easily obtained.

As indicated in Tang (2003), a critical drawback of this method is that the result relies on the choice of *j*, *k* and *l*, and it is difficult to determine which sets are optimal.

From the similar idea, Chen and Chen (2009) designed a simulation-based confidence interval construction method for the location parameter τ. The authors defined a function of τ as

$$\omega(\tau)=\frac{\ln(t_l-\tau)-\ln(t_k-\tau)}{\ln(t_k-\tau)-\ln(t_j-\tau)}=\frac{\ln\{(t_l-\tau)/\alpha\}^{\beta}-\ln\{(t_k-\tau)/\alpha\}^{\beta}}{\ln\{(t_k-\tau)/\alpha\}^{\beta}-\ln\{(t_j-\tau)/\alpha\}^{\beta}} \tag{2.14}$$

It can be seen that when α, β and τ are the correct model parameters, $\omega(\tau)$ is parameter free, due to the fact that $\{(t_i-\tau)/\alpha\}^{\beta}$ is simply the order statistic from the standard exponential distribution. In addition, $\omega(\tau)$ increases in τ. Hence, by simulating samples from the standard exponential distribution and calculating $\omega=\dfrac{\ln(z_l)-\ln(z_k)}{\ln(z_k)-\ln(z_j)}$ for all the samples, a confidence interval of τ can be constructed based on $\omega(\tau)$ in that $\Pr(\omega_{1-\alpha/2}<\omega(\tau)<\omega_{\alpha/2})=1-\alpha$, and the corresponding upper and lower limit of τ being identified as

$$\omega(\tau_{1-\alpha/2})=\omega_{1-\alpha/2} \text{ and } \omega(\tau_{\alpha/2})=\omega_{\alpha/2} \tag{2.15}$$

From their simulation study, Chen and Chen (2009) also found the optimal selection of *j*, *k* and *l* as *j*=1, *l*=*n* and *k*=[(*n*+2)/3], where *n* is the sample size and [(*n*+2)/3] denotes the integral part of (*n*+2)/3.

2.4.2 Least Squares Estimation

LSE has always been a popular method for the Weibull family distributions. For the 3-parameter Weibull distribution, LSE aims to minimize the sum of squared error

$$\left(\tilde{\alpha},\tilde{\beta},\tilde{\tau}\right)=\arg\min_{\alpha,\beta,\tau}\sum\left[\beta\ln\left(t_i-\tau\right)-\beta\ln\left(\alpha\right)\right]^2 \tag{2.16}$$

Denoting $x_i(\tau)=\ln(t_i-\tau)$, from the linear regression theory, $\tilde{\alpha}$ and $\tilde{\beta}$ can be expressed as functions of τ, and incorporating them back into (2.16) we can obtain an optimization problem with only a variable τ. Upon solving the problem, estimates of the three parameters can be obtained.

Tang (2003) proposed a variant of the objective function by defining $z_i(\beta)=\left(-\ln\left(\hat{R}_i\right)\right)^{1/\beta}$, where $\hat{R}_i$ is the estimated survivability at time t_i

$$\tilde{\beta}=\arg\min_{\beta}\sum\left[t_i-\tau(\beta)-\alpha(\beta)z_i(\beta)\right]^2 \tag{2.17}$$

where $\alpha(\beta)$ and $\tau(\beta)$ are functions of β.

The advantage of the above estimation is that, in applications such as accelerated testing, "… it is desirable to have a common estimate of β…" (Tang, 2003).

A disadvantage of the ordinary LSE is that the estimation error is often larger than that of MLE. To reduce the error, weighted LSE are designed.

For the Weibull distribution, Swain *et al*. (1988) suggested the weights to be

$$w_i = \frac{(n+1)^2(n+2)}{i(n-i+1)} \tag{2.18}$$

Instead, Hung (2001) suggested the weights to be

$$w_i = \frac{\left[\hat{R}_i \ln\left(\hat{R}_i\right)\right]^2}{\sum\left[\hat{R}_i \ln\left(\hat{R}_i\right)\right]^2} \tag{2.19}$$

With these weight factors incorporated into (2.16), weighted LSE objective is revised as

$$\left(\tilde{\alpha}, \tilde{\beta}, \tilde{\tau}\right) = \arg\min_{\alpha,\beta,\tau} \sum w_i \left[\beta \ln(t_i - \tau) - \beta \ln(\alpha)\right]^2 \tag{2.20}$$

From the simulation comparison study, Hung (2001) showed the squared error loss of the weighted LSE universally smaller than that of the ordinary LSE.

More discussions on ordinary LSE and weighted LSE for the 3-parameter Weibull distribution can be found in Jukić *et al*. (2008), Marković *et al*. (2009), et al.

2.4.3 Maximum Product of Spacing

To avoid inconsistencies which could be encountered when maximizing the log-likelihood function with β<1, Cheng and Amin (1983) proposed to replace the likelihood function with the probability function, namely the spacing function, which is defined as

$$D_i = \int_{t_{i-1}}^{t_i} f(t \mid \alpha, \beta, \tau) dt \tag{2.21}$$

The MPS estimates of the parameters are obtained by maximizing the geometric mean of these spacings

$$G = \left\{ \prod_{i=1}^{n+1} D_i \right\}^{1/(n+1)} \tag{2.22}$$

Cheng and Amin (1983) discussed the sufficiency, consistency and asymptotic efficiency properties of the MPS estimators, and proved that they are better estimators than the MLE in terms of these properties for the 3-parameter Weibull distribution. Through a simulation study, Cousineau (2009) showed MPS estimators perform better than MLE in terms of bias and squared error.

Extensions on this method are discussed in Ghosh and Jammalamadaka (2001), Heathcote *et al*. (2002, 2004) and Cousineau *et al*. (2004).

2.4.4 Bayesian and Other Methods

Smith and Naylor (1987) compared the different behaviors of MLE and Bayesian estimators of the parameters of the 3-parameter Weibull distribution in detail, and concluded that the Bayesian method has a practical advantage, that it does not depend on the asymptotic of the log-likelihood function and has the freedom to choose different priors for the ease of numerical implementations.

As a numerical realization of the Bayesian method, MCMC techniques for fitting the 3-parameter Weibull distribution to tree diameter data was discussed in Greet *et al*. (1994). See Pang *et al*. (2005) and Zhao *et al*. (2008) for other applications of MCMC to the 3-parameter Weibull distribution.

Recent discussions on other estimation methods for the 3-parameter Weibull distribution include Lockhart and Stephens (1994), Hirose (1996, 2002), Offinger (1998), Tiku and Akkaya (2004), Cao and McCarty (2006), Balakrishnan and Kateri (2008), Cousineau (2009[1]).

Chapter 3. Statistical Analysis of the Modified Weibull Distribution

In the field of reliability study, an important topic is to model the life behavior of the system with a suitable statistical distribution. Among the various distributions already studied, the Weibull distribution has been proven to be flexible and versatile at describing monotonic failure rate data. However, for many modern complex systems which exhibit unimodal or bathtub shaped failure rates, the Weibull distribution is inadequate. In order to extend its application, generalizations of the Weibull distribution have been proposed. Among them the 3-parameter Weibull models are of much interest since these models are more flexible than the Weibull distribution at modeling non-monotonic failure rate data, and they have only one additional parameter as compared to the Weibull distribution, which keeps as much simplicity as possible for model analysis and data fitting.

The modified Weibull distribution (Lai *et al*., 2003) is one of such Weibull models. The CDF of this distribution is

$$G(t) = 1 - \exp\left\{-at^b e^{\lambda t}\right\},\ t \geq 0 \tag{3.1}$$

where $a > 0$, $b, \lambda \geq 0$, but b and λ cannot be zero at the same time.

When $\lambda = 0$, the modified Weibull reduces to the Weibull distribution; when $b = 0$, it reduces to the type-1 extreme value distribution and the support extends to the whole x axis. The modified Weibull distribution is also the asymptotic approximation of the Beta-Integrated distribution (Lai *et al*., 2003).

FRF of the modified Weibull distribution is able to exhibit monotonic and bathtub shapes. Several research papers have addressed the parameter estimation and burn-in related issues of the distribution, but little study has been carried out to look into the statistical properties, which is a prerequisite for the distribution to be applied to real lifetime data modeling. In this chapter, a systematic structural statistical analysis of the distribution is presented, discussion including moments, PDF and FRF. After the statistical analysis, iterative steps of obtaining parameter MLE under a progressively type-2 censoring scheme are described, and model selection method using a chi-square test is proposed. To illustrate the application of the distribution, two examples of modeling lifetime data are presented. At the end of this chapter, a tentative study of maximum probability estimation (MPE) of the 3-parameter Weibull parameters is presented.

3.1 Statistical Properties

The SF is

$$\overline{G}(t) = 1 - G(t) = \exp\left\{- a t^{b} e^{\lambda t}\right\} \tag{3.2}$$

PDF and FRF are accordingly

$$g(t) = a(b + \lambda t)t^{b-1}e^{\lambda t}\exp\{-at^b e^{\lambda t}\} \tag{3.3}$$

$$h(t) = \frac{g(t)}{\overline{G}(t)} = a(b + \lambda t)t^{b-1}e^{\lambda t} \tag{3.4}$$

3.1.1 Moments

The *k*th moment of a modified Weibull random variable is

$$\mu_k' = E(T^k) = \int_0^\infty t^k g(t)dt = k\int_0^\infty t^{k-1}\overline{G}(t)dt = k\int_0^\infty t^{k-1}\exp\{-at^b e^{\lambda t}\}dt \tag{3.5}$$

For $t \geq 0$, $\overline{G}(t) \leq \exp(-at^b)$, so

$\mu_k' \leq k\int_0^\infty t^{k-1}e^{-at^b}dt = \frac{k}{a^{k/b}b}\int_0^\infty x^{k/b-1}e^{-x}dx = \frac{k}{a^{k/b}b}\Gamma(k/b)$, and μ_k' is a finite positive value.

Denoting $\mu = E[T] = \mu_1'$ as the mean, the *k*th central moment is

$$\mu_k = E(T-\mu)^k = \sum_{i=0}^{k}\binom{k}{i}(-1)^{k-i}\mu_k'\mu^{k-i} \tag{3.6}$$

$\mu_k < \mu_k'$ is also finite.

From (3.5) and (3.6), the raw moments and central moments are not able to be expressed in closed form and have to be evaluated numerically.

However, the following property of a modified Weibull variable is useful in studying the central moments.

Lemma 3.1. If T follows a modified Weibull distribution with parameters (a, b, λ), then λT is also modified Weibull distributed, but with parameters (a/λ^b, b, 1).

Proof: $\Pr(\lambda T > t) = \Pr\left(T > \frac{t}{\lambda}\right)$

$$= \exp\left\{-a\left(\frac{t}{\lambda}\right)^b e^{\lambda\frac{t}{\lambda}}\right\}$$

$$= \exp\left\{-\frac{a}{\lambda^b} t^b e^t\right\}. \text{ Q.E.D.}$$

While a scaling transformation changes the mean of the variable proportionally, it changes variance quadratically, hence keeps skewness and kurtosis intact.

$$\mu^* = E(\lambda T) = \lambda E(T) = \lambda\mu$$

$$\sigma^{*2} = E\{\lambda T - E(\lambda T)\}^2 = \lambda^2 E\{T - E(T)\}^2 = \lambda^2\sigma^2$$

$$\gamma_1^* = \frac{\mu_3^*}{(\mu_2^*)^{\frac{3}{2}}} = \frac{E\{\lambda T - E(\lambda T)\}^3}{\sigma^{*3}} = \frac{\lambda^3 E\{T - E(T)\}^3}{\lambda^3\sigma^3} = \frac{\mu_3}{(\mu_2)^{\frac{3}{2}}} = \gamma_1$$

$$\gamma_2^* = \frac{\mu_4^*}{(\mu_2^*)^2} - 3 = \frac{E\{\lambda T - E(\lambda T)\}^4}{\sigma^{*4}} - 3 = \frac{\lambda^4 E\{T - E(T)\}^4}{\lambda^4\sigma^4} - 3 = \frac{\mu_4}{\mu_2^2} - 3 = \gamma_2$$

Ignoring the scaling effects, we fix λ at 1, change a and b simutaneously and tabulate these basic statistics of the modified Weibull variable in the following table 3.1.

Table 3.1 Mean, variance, skewness and kurtosis of modified Weibull for different parameters

a	*b*	0.1	0.2	0.5	1.0	2.0	5.0
0.1	Mean	1.7769	1.7340	1.6223	1.4837	1.3195	1.1464
	Var	0.9848	0.8836	0.6499	0.4111	0.1950	0.0479
	Skew	-0.1874	-0.1725	-0.1444	-0.1402	-0.2123	-0.4789
	Kurt	-0.8244	-0.7941	-0.7074	-0.5770	-0.3593	0.1642
0.2	Mean	1.2089	1.1958	1.1609	1.1177	1.0687	1.0236
	Var	0.7318	0.6518	0.4727	0.2981	0.1456	0.0396
	Skew	0.1537	0.1463	0.1147	0.0428	-0.1191	-0.4775
	Kurt	-0.9766	-0.9287	-0.8046	-0.6506	-0.4236	0.3206
0.5	Mean	0.5828	0.6039	0.6561	0.7168	0.7887	0.8782
	Var	0.3644	0.3301	0.2507	0.1695	0.0930	0.0304
	Skew	0.8117	0.7404	0.5565	0.3204	0.0051	-0.4354
	Kurt	-0.3344	-0.3901	-0.4952	-0.5501	-0.4581	0.0963
1.0	Mean	0.2526	0.2853	0.3729	0.4819	0.6149	0.7802
	Var	0.1438	0.1391	0.1227	0.0971	0.0630	0.0247
	Skew	1.6371	1.4393	0.9997	0.5564	0.0972	-0.4179
	Kurt	2.0603	1.4545	0.3869	-0.2692	-0.4479	0.0481
2.0	Mean	0.0712	0.0974	0.1821	0.3057	0.4718	0.6918
	Var	0.0312	0.0358	0.0452	0.0487	0.0409	0.0200
	Skew	3.2686	2.6579	1.5899	0.8103	0.1850	-0.4019
	Kurt	11.8823	7.7631	2.4579	0.2360	-0.4082	0.0431
5.0	Mean	0.0039	0.0106	0.0533	0.1533	0.3251	0.5884
	Var	0.0008	0.0017	0.0068	0.0157	0.0216	0.0149
	Skew	12.0332	7.2333	2.7278	1.1568	0.2910	-0.3832
	Kurt	179.739	66.3643	9.5261	1.2717	-0.2256	0.0244

The general patterns of these statistics can be exhibited graphically. For example, for $a = 0.1$, the following figure shows the trends as b increases

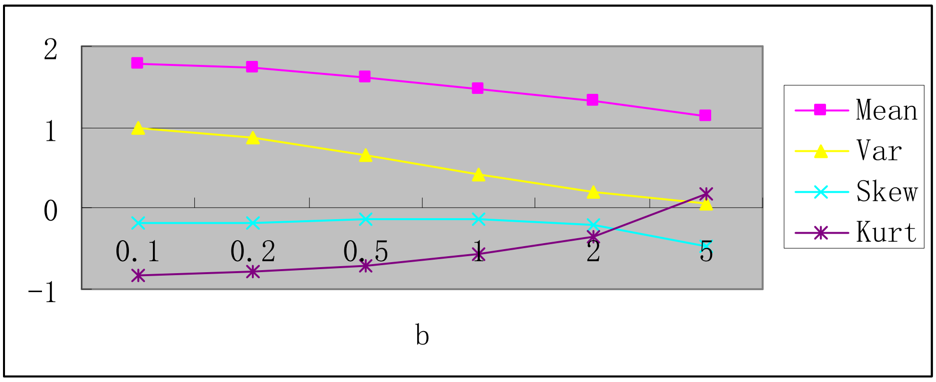


Figure 3.1 Mean, Variance, Skewness and Kurtosis of a MW Variable as *b* Increases, *a* = 0.1

From another perspective, fixing $b = 5.0$, as a increases, changes of these statistics are exhibited in the following figure

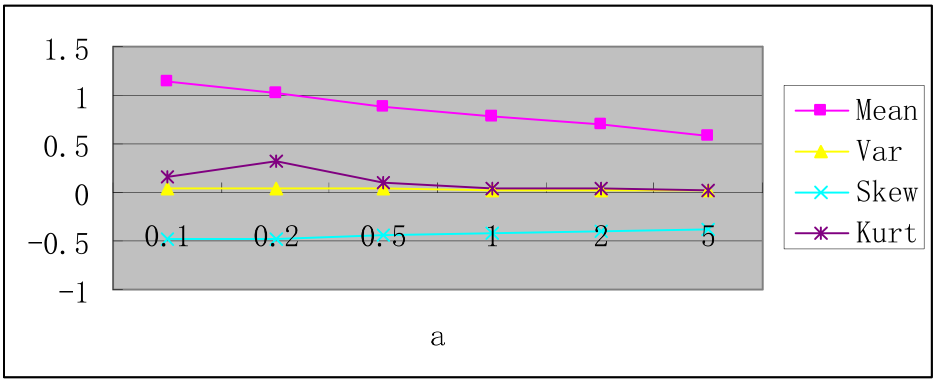


Figure 3.2 Mean, Variance, Skewness and Kurtosis of a MW Variable as *a* Increases, *b* = 5.0

Overall, the trends of these statistics are summarized as follows

1. When a<0.5, mean decreases as b increases; when a≥0.5, mean increases as b increases; mean is decreasing in a.
2. When a<2, variance decreases as b increases; when a≥0.5, variance increases as b increase; variance is decreasing in a.
3. Skewness decreases in b but increases in a; and the modified Weibull distribution could be either positive skew (skewed to the left) or negative skew (skewed to the right).
4. When a<1, kurtosis increases as b increases; when a≥1, kurtosis decreases as b increases; kurtosis basically increases in a.

3.1.2 Probability Density Function

Differentiating PDF $g(t)$ in (3.3) with respect to t yields

$$g'(t) = at^{b-2}e^{\lambda t}\exp\{-at^b e^{\lambda t}\}\left[(b+\lambda t)^2\left(1-at^b e^{\lambda t}\right)-b\right] \tag{3.7}$$

When $t > 0$, the sign of $g'(t)$ is determined by the sign of

$$K(t) = (b+\lambda t)^2\left(1-at^b e^{\lambda t}\right)-b \tag{3.8}$$

If we let $x = \lambda t$ and $c = a/\lambda^b$, then

$$K(t) = (b+x)^2\left(1-cx^b e^x\right)-b \tag{3.9}$$

We discuss the following different cases:

1. $\lambda = 0$: In this case the modified Weibull reduces to the Weibull distribution. Then $g(t)$ is monotonically decreasing if $b \le 1$ and is unimodal if $b > 1$, with mode

$$T^* = a\left(\frac{b-1}{b}\right)^{1/b} \tag{3.10}$$

2. $b = 0$: In this case it reduces to the type-1 extreme value distribution. Then $g(t)$ is monotonically decreasing if $a \ge 1$ and is unimodal if $a < 1$, with mode

$$T^* = \frac{1}{\lambda}\log\left(\frac{1}{a}\right) \tag{3.11}$$

Note in this case the support of CDF is the whole real line, but T^* is always positive.

3. $0 < b < 1$: From (3.8) $K(0) = b^2 - b < 0$; as t increases, $K(t)$ increases, but whether $K(t)$ can be larger than 0 depends on the value of c (if c is small, there may be a T^* that $K(T^*) = 0$; if c is large, such T^* does not exist); $\lim_{t\to\infty} K(t) = -\infty$. In other words, if c is small, $K(t)$ is initially negative, then positive, and finally negative; if c is large, $K(t)$ is invariably negative. However, there seems to be no simple explicit relationship between the value of b and threshold value of c.

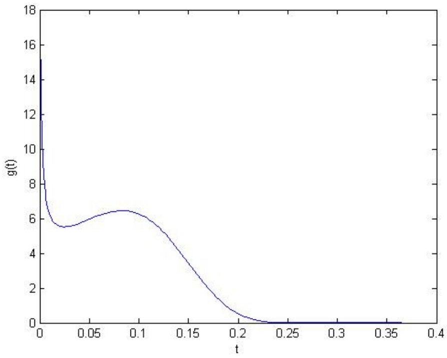
$\lim_{t \to 0} g(t) = \infty$ and $\lim_{t \to \infty} g(t) = 0$. In both cases of c large or small, $g(t)$ is S-shaped. If c is small, $g(t)$ is initially decreasing and finally decreasing, but with a transitional period where it is increasing; if c is large, $g(t)$ is monotonically decreasing.

4. $b = 1$: From (3.8) $K(0) = 0$; as t increases, $K(t)$ increases and reaches its maximum at T^*, and then decreases; $\lim_{t \to \infty} K(t) = -\infty$. In a word, $K(t)$ is unimodal, initially positive and then negative.

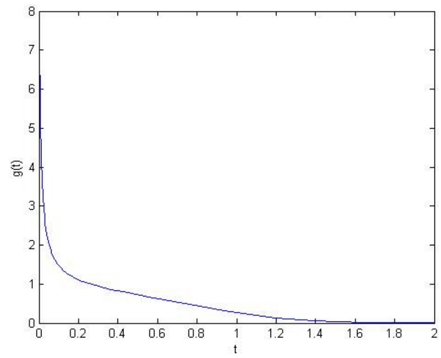
$g(0) = a$ and $\lim_{t \to \infty} g(t) = 0$. $g(t)$ is unimodal in such case.

5. $b > 1$: From (3.8) $K(0) = b^2 - b > 0$; $\lim_{t \to \infty} K(t) = -\infty$. $K(t)$ is initially positive and then negative, so $g(t)$ is unimodal.

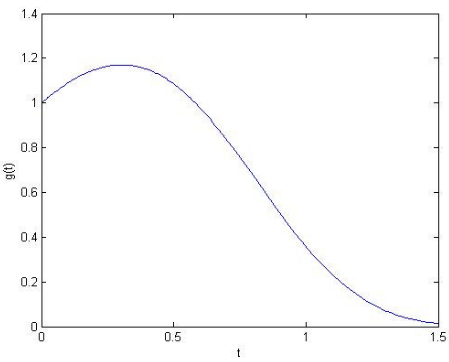
Figure 3.1 shows some typical shapes of PDF with different parameters



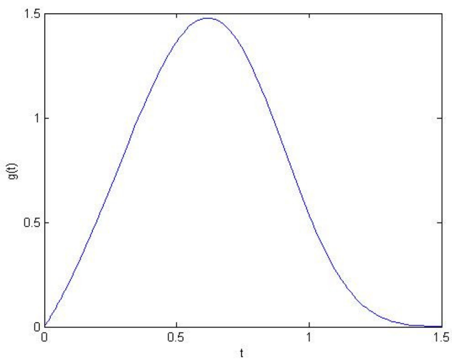
a=1, b=0.5, λ=12 (c=0.2887)



a=1, b=0.5, λ=1 (c=1)



a=1, b=1, λ=1



a=1, b=2, λ=1

Figure 3.3 Plots of PDF of the modified Weibull distribution for different parameters

3.1.3 Failure Rate Function

Differentiating FRF $h(t)$ with respect to t, equation (3.4) yields

$$h'(t) = at^{b-2}e^{\lambda t}\left\{(b+\lambda t)^2 - b\right\} \tag{3.12}$$

The different shapes of FRF have been studied in Lai *et al*. (2003). We summarize the results here.

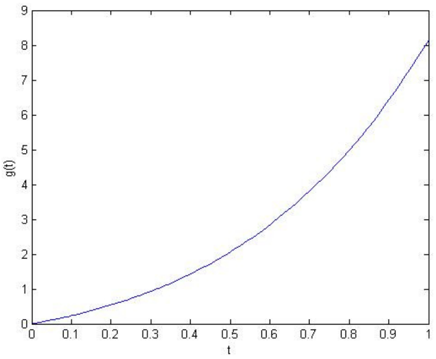
1. $b \geq 1$, $h'(t) > 0$ for all $t > 0$, so $h(t)$ is monotonically increasing. $h(0) = 0$ if $b > 1$, $h(0) = ab$ if $b = 1$; and $\lim_{t \to \infty} h(t) = \infty$.

2. $0 < b < 1$, $\lambda > 0$, $h'(t) < 0$ for $0 \leq t < t^* = \dfrac{\sqrt{b} - b}{\lambda}$, $h'(t^*) = 0$, and $h'(t) > 0$ for $t > t^*$. $\lim_{t \to 0} h(t) = \infty$, $\lim_{t \to \infty} h(t) = \infty$. $h(t)$ is bathtub shaped, with $t^* = \dfrac{\sqrt{b} - b}{\lambda}$ being the change point.

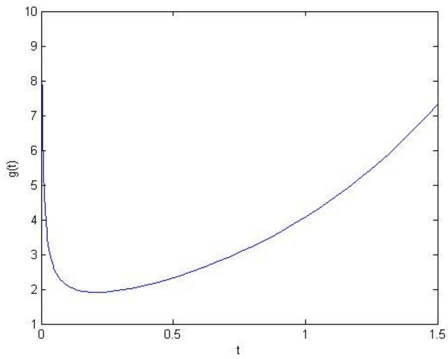
3. $0 < b < 1$, $\lambda = 0$, $h'(t) < 0$ for all $t > 0$, so $h(t)$ is monotonically decreasing. $\lim_{t \to 0} h(t) = \infty$ and $\lim_{t \to \infty} h(t) = 0$.

4. $b = 0$, $h'(t) > 0$ for all $-\infty < t < \infty$, so $h(t)$ is monotonically increasing. $\lim_{t \to -\infty} h(t) = 0$ and $\lim_{t \to \infty} h(t) = \infty$.

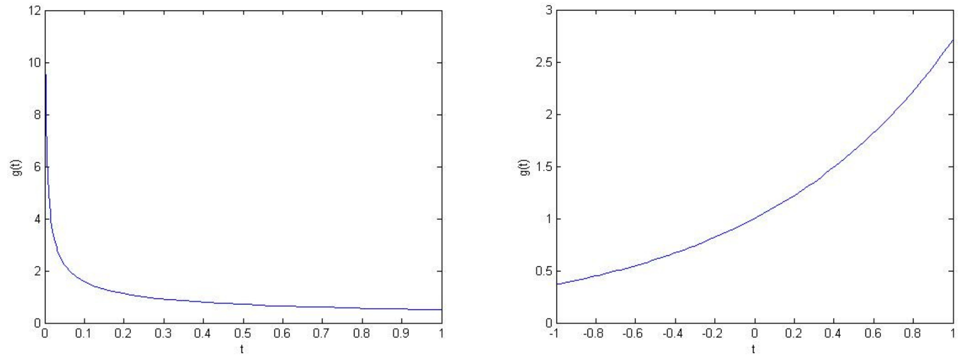
Figure 3.2 shows some typical shapes of FRF with different parameters



a=1, *b*=1, λ=1



a=1, *b*=0.5, λ=1



a=1, b=0.5, λ=0 a=1, b=0, λ=1

Figure 3.4 Plots of FRF of the modified Weibull distribution for different parameters

3.2 Statistical Inferences

For the modified Weibull distribution, the WPP transformation $x = \log(t)$, $y = \log(-\log(1-F(t)))$ yields

$$y = \log(a) + bt + \lambda \exp(t)$$

which is a nonlinear relationship, but with relatively simpler form as compared to that of other Weibull related distributions, such as exponentiated Weibull (Jiang and Murthy, 1999), which cannot be expressed in a multi-linear expression.

Lai *et al*. (2003) discussed the graphic method of implementing WPP and an alternative plot and the analytic method applying multiple linear regression analysis to estimate the model parameters.

3.2.1 Maximum Likelihood Estimation

A more rigorous method is maximum likelihood (ML) estimation. Ng (2005) derived the likelihood equations and second derivatives of the log-likelihood function with regard to the parameters in the case of progressively type-2 censored data, and recommended an iterative procedure to obtain the MLE of the model parameters. We write the log-likelihood function, likelihood equations and second derivatives of the log-likelihood function here as reference.

Let $t_1 < t_2 < \cdots < t_m$ denote the failure times of the items and $r_1 < r_2 < \cdots < r_m$ be the numbers of censored items at the corresponding failure times. The underlying log-likelihood function is

$$\begin{aligned} L(a,b,\lambda) &= \log\left(C\prod_{i=1}^{m} f(t_i \mid a,b,\lambda)\left(\overline{F}(t_i \mid a,b,\lambda)\right)^{r_i} \right) \\ &= \log(C) + m\log(a) + (b-1)\sum_{i=1}^{m}\log t_i + \lambda\sum_{i=1}^{m} t_i + \sum_{i=1}^{m}\log(b+\lambda t_i) - a\sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i} \end{aligned} \tag{3.13}$$

where $C = n(n-1-r_i)(n-2-r_1-r_2)\ldots(n-m-r_1-\ldots-r_m)$ and is usually neglected in analysis.

Under mild regularity conditions (Gong and Samaniego, 1981; Godambe, 1960), the maximum of $L(a,b,\lambda)$ in (3.13) occurs when its derivatives are equal to 0. The likelihood equations are derived by taking differentiates of $L(a,b,\lambda)$ with respect to a, b and λ

$$\frac{\partial L}{\partial a} = \frac{m}{a} - \sum_{i=1}^{m} (1 + r_i) t_i^b e^{\lambda t_i} = 0 \tag{3.14}$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{m} \log(t_i) + \sum_{i=1}^{m} \frac{1}{b + \lambda t_i} - a \sum_{i=1}^{m} (1 + r_i) t_i^b e^{\lambda t_i} \log t_i = 0 \tag{3.15}$$

$$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^{m} t_i + \sum_{i=1}^{m} \frac{t_i}{b + \lambda t_i} - a \sum_{i=1}^{m} (1 + r_i) t_i^{b+1} e^{\lambda t_i} = 0 \tag{3.16}$$

The second derivatives of $L(a,b,\lambda)$ with regard to a, b and λ are

$$\frac{\partial^2 L}{\partial a^2} = -\frac{m}{a^2} \tag{3.17}$$

$$\frac{\partial^2 L}{\partial b^2} = -\sum_{i=1}^{m} \frac{1}{(b + \lambda t_i)^2} - \sum_{i=1}^{m} (1 + r_i) a t_i^b e^{\lambda t_i} (\log t_i)^2 \tag{3.18}$$

$$\frac{\partial^2 L}{\partial \lambda^2} = -\sum_{i=1}^{m} \frac{t_i^2}{(b + \lambda t_i)^2} - \sum_{i=1}^{m} (1 + r_i) a t_i^{b+2} e^{\lambda t_i} \tag{3.19}$$

$$\frac{\partial^2 L}{\partial a \partial b} = -\sum_{i=1}^{m} (1 + r_i) t_i^b e^{\lambda t_i} \log t_i \tag{3.20}$$

$$\frac{\partial^2 L}{\partial a \partial \lambda} = -\sum_{i=1}^{m} (1 + r_i) t_i^{b+1} e^{\lambda t_i} \tag{3.21}$$

$$\frac{\partial^2 L}{\partial b \partial \lambda} = -\sum_{i=1}^{m} \frac{t_i}{(b + \lambda t_i)^2} - a \sum_{i=1}^{m} (1 + r_i) t_i^{b+1} e^{\lambda t_i} \log t_i \tag{3.22}$$

The Hessian matrix of $L(a,b,\lambda)$ is

$$H(a,b,\lambda) = \frac{\partial^2 L(a,b,\lambda)}{\partial(a,b,\lambda)\partial(a,b,\lambda)'} \tag{3.23}$$

To get the MLE of the model parameters, an iterative procedure is needed here since equations (3.14) – (3.16) do not have closed form solutions. Methods of iteration include Newton-Raphson (Jennrich and Sampson, 1976), and Expectation Maximization (Dempster *et al*, 1977).

In our current research, the Newton-Raphson iteration method is used to obtain the parameter estimates. The algorithm is comprised of the following steps,

0. Use LSE or some other proper estimates of the parameters as starting point of iteration, denote the estimates as (a_0, b_0, λ_0), and set k=0;

1. Calculate $(L_a, L_b, L_\lambda)_{(a_k, b_k, \lambda_k)}$, which is the derivative vector of $L(a, b, \lambda)$ with regard to a, b and λ, at point (a_k, b_k, λ_k);

2. Calculate Hessian matrix $H(a_k, b_k, \lambda_k)$, and the inverse $H^{-1}(a_k, b_k, \lambda_k)$;

3. Update (a, b, λ) as

   $$(a_{k+1}, b_{k+1}, \lambda_{k+1}) = (a_k, b_k, \lambda_k) - (L_a, L_b, L_\lambda)_{(a_k, b_k, \pi_k)} H^{-1}(a_k, b_k, \lambda_k);$$

4. Set k=k+1, and then go back to step 1. Continue the iterative steps until $\left|(a_{k+1}, b_{k+1}, \lambda_{k+1}) - (a_k, b_k, \lambda_k)\right|$ is smaller than a threshold value.

The final estimates of (a, b, λ) are the MLE of the parameters, denoted as $(\hat{a}, \hat{b}, \hat{\lambda})$.

At the MLE point $(\hat{a}, \hat{b}, \hat{\lambda})$, the negative Hessian matrix $-H(a, b, \lambda)\big|_{(\hat{a}, \hat{b}, \hat{\lambda})}$ is called the Observed Fisher Information matrix and its inverse is the asymptotic approximate of the variance-covariance matrix of $(\hat{a}, \hat{b}, \hat{\lambda})$ under normality approximation.

3.2.2 Statistical Decision

To test if the modified Weibull distribution is a more appropriate model, than the Weibull or type-1 extreme value distribution, to model a set of data, one can use the likelihood ratio test.

When the null hypothesis is the Weibull distribution, the test statistic is given as

$$\Lambda_1 = 2\left\{L(\hat{a}, \hat{b}, \hat{\lambda}) - L_1(\hat{a}_1, \hat{b}_1)\right\} \tag{3.24}$$

where $L_1(a, b)$ is the log-likelihood function under the null hypothesis (2-parameter Weibull) and $(\hat{a}_1, \hat{b}_1)$ are the MLE of $L_1(a, b)$, and $L(\hat{a}, \hat{b}, \hat{\lambda})$ is the log-likelihood of the modified Weibull, which is the alternative hypothesis.

According to Akaike (1974), under the null hypothesis, Λ_1 is asymptotically distributed as chi-square χ^2_{p-q} distribution, where p is the number of unknown parameters of the distribution in the alternative hypothesis, and q is the number of unknown parameters in the null hypothesis. Therefore, in this case, Λ_1 is asymptotically distributed as χ^2. With a given significance level α, one should reject the null assumption if Λ_1 is larger than $\chi^2_{(\alpha)}$, the upper $100 \times (1-\alpha)$ percentile point of a chi-square variable, or otherwise do not reject the assumption.

However, it should be noted that when MLE of the parameters are used to compute the χ^2 statistic for goodness of fit test, the test statistic stochastically dominates that would be expected under the chi-square theory (Chernoff and Lehmann, 1954), and the result is the probability of rejection, when the null hypothesis if true, is greater than the desired significance level.

Similarly, when the null hypothesis is type-1 extreme value distribution, the test statistic is

$$\Lambda_2 = 2\left\{L\left(\hat{a}, \hat{b}, \hat{\lambda}\right) - L_2\left(\hat{a}_2, \hat{\lambda}_2\right)\right\} \tag{3.25}$$

where $L_2(a, \lambda)$ is the log-likelihood function under the null hypothesis (type-1 extreme value) and $\left(\hat{a}_2, \hat{\lambda}_2\right)$ are the MLE of $L_2(a, \lambda)$.

Under the null hypothesis, Λ_2 is also asymptotically distributed as chi-square χ^2 distribution. Therefore we can make statistical decisions based on the asymptotic distribution.

3.3 Illustrative Examples

3.3.1 Aarset Data

The light bulb lifetime data from Aarset (1987) is often cited by researchers as a good example with bathtub shaped failure rate (e.g. Xie *et al*. (2002), Lai *et al*. (2003)). Since this data set will be used several times in the thesis, we put it here for reference

Table 3.2 Lifetimes of 50 devices

0.1	0.2	1	1	1	1	1	2	3	6
7	11	12	18	18	18	18	18	21	32
36	40	45	46	47	50	55	60	63	63
67	67	67	67	72	75	79	82	82	83
84	84	84	85	85	85	85	85	86	86

Due to the S-shape of the TTT plot, the data exhibit a bathtub shaped failure rate (Aarset, 1987).

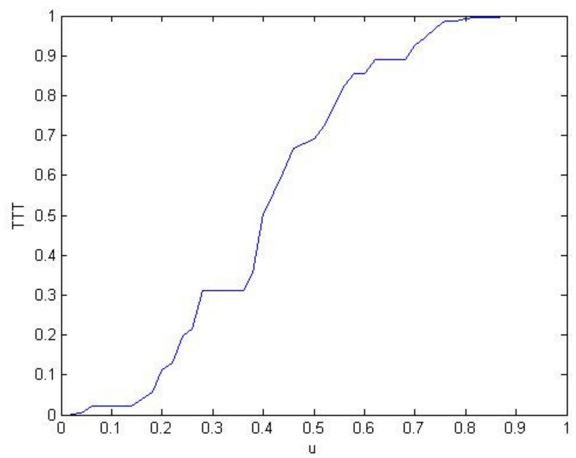


Figure 3.5 TTT transformation of the Aarset data

Using the modified Weibull distribution to model the data, Lai *et al*. (2003) obtained the parameter estimates using a regression procedure $(\tilde{a},\tilde{b},\tilde{\lambda}) = (0.0876, 0.389, 0.01512)$, and Ng (2005) obtained the MLE $(\hat{a},\hat{b},\hat{\lambda}) = (0.0624, 0.355, 0.02332)$. Both estimates of the shape parameter *b* support the assumption of bathtub shaped FRF. The following figure presents the fit of the modified Weibull distribution in a WPP plot

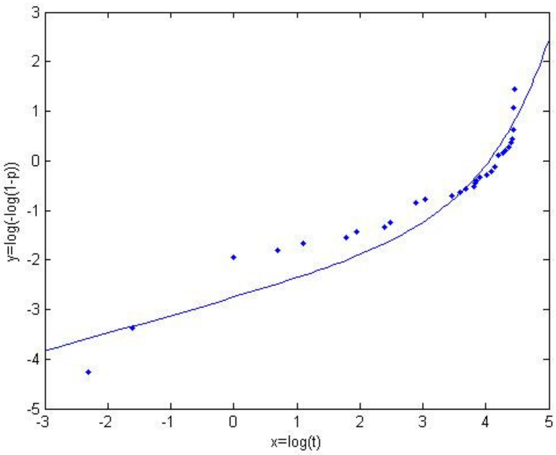


Figure 3.6 Goodness of fit of the modified Weibull to Aarset data

The model fits the data well, though it does not capture the pattern perfectly. The following figure exhibit the fits of the other two popular models, the exponentiated Weibull and the Weibull extension, both with MLE as model parameters.

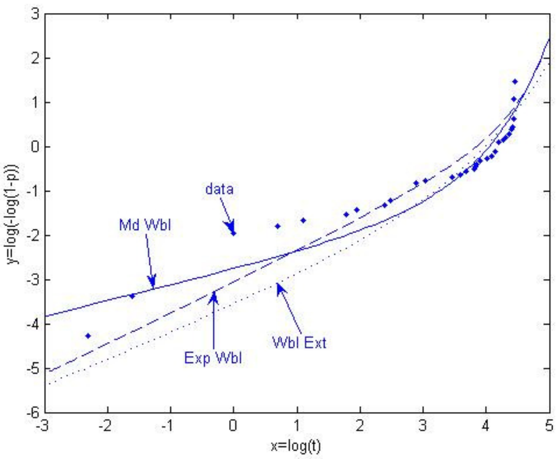


Figure 3.7 Goodness of fit of the modified Weibull (solid), exponentiated Weibull (dashed) and Weibull extension (dotted)

However, it is quite difficult to say which model fits the data better than the others. In chapter 6 we will have more details on the comparison and there we will find that another newly proposed model, the odd Weibull, fit the data far better.

Assuming a progressively type-2 censoring scheme for this data set, where only 35 failures are observed, $r_3 = 4$, $r_{10} = 4$, $r_{23} = 3$, $r_{33} = 4$ and $r_i = 0$, $1 \le i \ne 3, 10, 23, 33 \le 35$. A progressively type-2 censored sample is obtained by censoring a predetermined number of units, once a failure or a number of consecutive failures happen. This scheme differs from a traditional type-2 censoring scheme in that it provides more flexibility, and retains more information from the experiment with the same cost (e.g. Bairamov, 2006).

Table 3.3 Simulated progressively type-2 censored sample

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
t_i	0.1	0.2	1	2	3	6	7	11	12	18	21	32	36	40	45	46	47	50
r_i	0	0	4	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0
i	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	
t_i	55	60	63	63	67	72	75	79	82	82	83	84	84	84	85	86	86	
r_i	0	0	0	0	3	0	0	0	0	0	0	0	0	0	4	0	0	

By maximizing $L(a,b,\lambda)$ iteratively, the MLE are obtained as $\left(\hat{a}_0, \hat{b}_0, \hat{\lambda}_0\right) = (0.0376, 0.3398, 0.0263)$ and the likelihood is $L\left(\hat{a}, \hat{b}, \hat{\lambda}\right) = -170.65$, and the approximates of variance are $\operatorname{var}(\hat{a}_0) = 4.1718 * 10^{-4}$, $\operatorname{var}\left(\hat{b}_0\right) = 0.0206$ and $\operatorname{var}\left(\hat{\lambda}_0\right) = 3.7684 * 10^{-5}$.

To see if it is necessary to use the modified Weibull distribution instead of Weibull, assuming the Weibull distribution as the underlying model, then the parameter estimates are $\left(\hat{a}_1, \hat{b}_1\right) = (0.0128, 1.044)$ and the maximized likelihood is $L_1\left(\hat{a}_1, \hat{b}_1\right) = -181.21$. Hence the test statistic $\Lambda_1 = 2\left[L\left(\hat{a}, \hat{b}, \hat{\lambda}\right) - L_1\left(\hat{a}_1, \hat{b}_1\right)\right] = 21.12 > \chi^2_{0.05} = 3.8415$ at 0.05 significance level, so the null hypothesis is rejected. Therefore, we claim that the modified Weibull distribution provides a better fit for the data than the Weibull distribution.

Alternatively, assuming the type-1 extreme value distribution as the underlying model, then the parameter estimates are $\left(\hat{a}_2, \hat{b}_2\right) = (0.0639, 0.0385)$ and the maximized likelihood is $L_2(\hat{a}_2, \hat{c}_2) = 181.02$. The test statistic $\Lambda_1 = 20.94 > \chi^2_{0.05} = 3.8415$. Hence, at the 0.05 significance level, we can also claim that the modified Weibull provides a better fit for the data than the type-1 extreme value distribution.

3.3.2 Kumar Data

The time between successive failures data (in hours) of load-haul-dump machines for loading rock in underground mines were gathered and studied in Kumar *et al.* (1989). The following table contains this data set.

Table 3.4 Time between successive failures of LHD machines

16	39	71	95	98	110	114	226	294	344
555	599	757	822	963	1077	1167	1202	1257	1317
1345	1372	1402	1536	1625	1643	1675	1726	1736	1772
1796	1799	1814	1868	1894	1970	2042	2044	2094	2127
2291	2295	2299	2317						

The TTT plot also has an S-shape, so the failure rate function has a bathtub shape, and then we can use the modified Weibull distribution to fit the data.

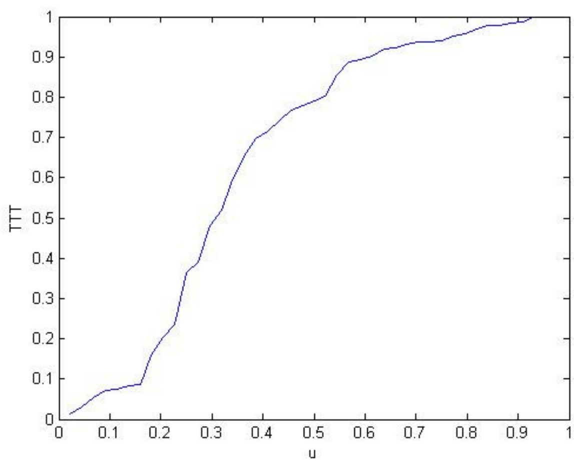


Figure 3.8 TTT transformation of the Kumar data

MLE of the model parameters of a modified Weibull distribution fitting are $\left(\hat{a},\hat{b},\hat{\lambda}\right)=\left(0.007,0.4322,0.0011\right)$, which support the assumption that FRF is bathtub shaped. The following figure shows the fit of the modified Weibull model.

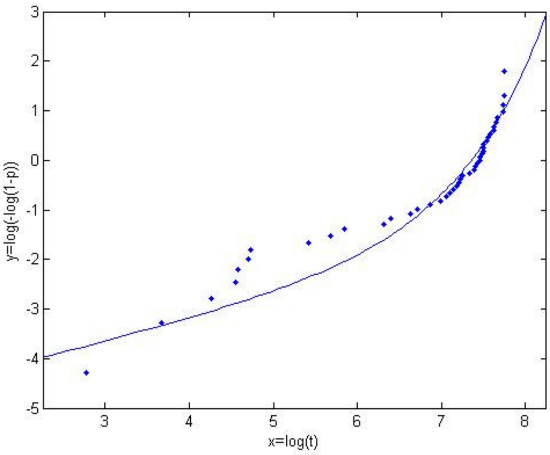


Figure 3.9 Goodness of fit of the modified Weibull to Kumar data

3.4 Maximum Probability Estimation for 3-Parameter Weibull

As a generalization of MLE, the maximum probability estimation (MPE) method, which was introduced by Weiss and Wolfowitz (1967), does not have some of the intrinsic inadequacies of the general MLE method, most importantly the prerequisite of "regularities class". MPE raised a lot of interest during the 60s' and 70s' and was discussed in detail by Weiss and Wolfowitz (1970), Kuβ (1972), Dudewicz (1973), Wegner (1976), Roussas (1977), Akahira (1991), etc. However, during the recent years, there seems to be no updated research and application around this method.

It is well known that for the 3-parameter Weibull distribution (2.1), when the shape parameter $\beta<1$, the likelihood function is not bounded so MLE do not exist, when $1\leq\beta<2$, MLE of the parameters exist but do not satisfy the usual regularity conditions and hence are inefficient, and only when $\beta>2$, the distribution function is regular and MLE of the model parameters exist and are consistent (see Rockette *et al*., 1974; Smith, 1985; and Kantar and Senoglu, 2008). The CDF of this distribution is put down below for reference

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

The difficulties in ML estimation of the parameters of the 3-parameter Weibull distribution and the ways to circumvent them have been reviewed in Chapter 2.

Since researchers claim that MPE is able to derive parameter estimates even when the underlying distribution does not meet the regularity conditions, in this section we will formulate the estimation procedure for the parameters of the 3-parameter Weibull distribution and study the characteristics of the estimators.

Given an underlying PDF $f(t \mid \theta)$, where θ is the unknown or partially unknown parameter vector and Θ is the corresponding parameter space, and a sample of data $\{t_1, t_2, \cdots, t_n\}$ from the distribution, MLE of the parameters are obtained by maximizing the likelihood function $l(\theta \mid t_1, t_2, \cdots, t_n) = \prod_{i=1}^{n} f(t_i \mid \theta)$ over the parameter space Θ, that is

$$\hat{\theta} = \arg\max_{\theta \in \Theta} l(\theta \mid t_1, t_2, \cdots, t_n) \tag{3.26}$$

Instead of directly maximizing $l(\theta \mid t_1, t_2, \cdots, t_n)$, MPE method estimates the parameters via maximizing the integral of $l(\theta \mid t_1, t_2, \cdots, t_n)$ over the neighborhood of each $\theta \in \Theta$, that is

$$\tilde{\theta} = \arg\max_{\theta \in \Theta} \int_{R(\theta)} l(\theta \mid x_1, x_2, \cdots, x_n) d\theta \tag{3.27}$$

where $R(\theta) = \{\theta' \mid \|\theta' - \theta\| < d\}$ and d is a proper constant.

For the 3-parameter Weibull distribution, since the "non-regularity" occurs on the location parameter τ, to make things simpler, the first step to study this problem would be to estimate τ, assuming the other two parameters α and β are known.

If we estimate τ using by (3.27), the computation would be very difficult. This is because as sample size n increases, the likelihood function $l(\tau)$ would become quite steep and hence make the integration computationally intractable. Therefore, as what is always done in ML estimation, we can modify the formula to maximize the integral of the log-likelihood instead of likelihood itself

$$\tilde{\tau} = \underset{\tau < \min\{t_1, \cdots, t_n\}}{\arg\max} \int_{R(\tau)} L(\tau \mid t_1, t_2, \cdots, t_n) d\tau \tag{3.28}$$

where $L(\tau \mid t_1, t_2, \cdots, t_n) = n \log\beta - n\beta\log\alpha + \sum_{i=1}^{n}(\beta - 1)\log(t_i - \tau) - \sum_{i=1}^{n}\left(\frac{t_i - \tau}{\alpha}\right)^{\beta}$ is the log-likelihood function.

Let $R(\tau) = \{\tau' \mid -r/\sqrt{n} < \tau' - \tau < r/\sqrt{n}\}$ where r is a small positive number. For $\tau < \min\{t_1, \cdots, t_n\} - r/\sqrt{n}$, omitting irrelevant terms, the integral of the log-likelihood function $L(\tau)$ can be expressed as

$$\begin{aligned} H(\tau) = {} & (\beta - 1)\sum\left(t_i - \tau + r/\sqrt{n}\right)\log\left(t_i - \tau + r/\sqrt{n}\right) + \frac{1}{\alpha^{\beta}(\beta + 1)}\sum\left(t_i - \tau - r/\sqrt{n}\right)^{\beta+1} \\ & - (\beta - 1)\sum\left(t_i - \tau - r/\sqrt{n}\right)\log\sum\left(t_i - \tau - r/\sqrt{n}\right) - \frac{1}{\alpha^{\beta}(\beta + 1)}\sum\left(t_i - \tau + r/\sqrt{n}\right)^{\beta+1} \end{aligned} \tag{3.29}$$

To maximize (3.29), we differentiate $H(\tau)$ with respect to τ

$$\begin{aligned} H'(\tau) = {} & (\beta - 1)\sum\log\left(t_i - \tau - r/\sqrt{n}\right) + \frac{1}{\alpha^{\beta}}\sum\left(t_i - \tau + r/\sqrt{n}\right)^{\beta} \\ & - (\beta - 1)\sum\log\left(t_i - \tau + r/\sqrt{n}\right) - \frac{1}{\alpha^{\beta}}\sum\left(t_i - \tau - r/\sqrt{n}\right)^{\beta} \end{aligned} \tag{3.30}$$

When β>1, $H'(\tau)$ it monotonically decreasing and has a unique zero point within $\left(-\infty, \min\{t_i, i=1,\cdots,n\} - r/\sqrt{n}\right)$. Hence $H(\tau)$ has a unique local maximum, which is the MPE.

In contrast, when β<1, $H'(\tau) > 0$ for all $\tau < \min\{t_1,\cdots,t_n\} - r/\sqrt{n}$, so $H(\tau)$ is monotonically increasing and is unbounded as r approaches 0.

To conclude, when β≥1, the MPE of τ exists and is unique. When β<1, the MPE of τ does not exist.

It turns out that our procedure for pursuing the MPE of the location parameter τ does not always guarantee the existence of the estimator, and hence is not usable unless β≥1. Nevertheless, our main purpose of this section is to bring attention from researchers to MPE, and hopefully in the future applicable estimation procedures can be generated based on this method.

3.5 Summary

In this chapter, the moments, PDF and FRF of the modified Weibull distribution are discussed. The log-likelihood function, likelihood equations and second derivatives of the log-likelihood are derived, and the iterative procedures for approaching the MLE of the model parameters are described. The likelihood-ratio test is employed to make statistical decisions between the modified Weibull distribution and its special cases, the Weibull and type-1 extreme value distributions. A practical example is presented to illustrate the use of the distribution to model real lifetime data and the adaptability of the distribution to bathtub shaped failure rate data. At the final part, MPE method is studied and applied to the estimation of the parameters of the 3-parameter Weibull distribution.

Chapter 4. On the Existence and Uniqueness of the MLE of the Modified Weibull Distribution

Given a set of sample data and data-fit statistical distribution, the MLE of the parameters of the distribution are obtained by maximizing the log-likelihood function. For distributions with more than one model parameter, iterative procedures are often needed to carry out the maximization. In the iterative steps, the property of the matrix composing of the second derivatives of the log-likelihood function with respect to the parameters, called the Hessian matrix, is very important. Besides, if the MLE of the parameters exist, the minus of the Hessian matrix at the MLE point is the Observed Fisher Information matrix and its inverse is the approximate variance-covariance matrix of the MLE.

In this chapter, we discuss the problem involved in the computation of the Observed Fisher Information matrix and the MLE for a broad class of the Weibull models, and then we apply the techniques developed to study the ML estimation procedures and properties of the model parameters of the modified Weibull distribution.

Part of the content is published in Jiang *et al*. (2010).

Consider the class of distributions introduced by Gurvich *et al*. (1997),

$$G(t) = 1 - \exp\{-aM(t)\} \tag{4.1}$$

where $a > 0$ is a model parameter and $M(t)$ is an increasing function of *t* with or without parameters.

When $M(t) = t$, $G(t)$ is the exponential CDF; when $M(t) = t^b$ with a parameter *b*, $G(t)$ is the Weibull CDF; when $M(t) = \exp(bt)$, $G(t)$ is the type-1 extreme value CDF. If $M(t)$ is a function with two or more parameters, $G(t)$ could represent a broad class of distributions, including the Weibull extension (Xie *et al*., 2002) and the modified Weibull (Lai *et al*., 2003).

4.1 Simplification of Observed Fisher Information Matrix

In this section we consider $M(t \mid b, c)$ (which is abbreviated as $M(t)$ in the following) with two parameters *b* and *c*, so CDF $G(t)$ of (4.1) has three parameters in all. Taking the derivative of $G(t)$ with respect to *t* we obtain PDF of the distribution

$$g(t) = am(t)\exp\{-aM(t)\} \tag{4.2}$$

where $m(t \mid b, c) = \frac{\partial}{\partial t} M(t)$, and we write it as $m(t)$ in the following text.

Then the log-likelihood function with a complete sample $t_1, t_2, \cdots, t_n$ (though right censoring is essentially the same) is

$$L(a,b,c) = n\log a + \sum \log(m(t_i)) - a\sum M(t_i) \tag{4.3}$$

Differentiating $L(a,b,c)$ with respect to a, b and c and equating the derivatives to zero, we obtain the likelihood equations

$$\frac{\partial L}{\partial a} = \frac{n}{a} - \sum M(t_i) = 0 \tag{4.4}$$

$$\frac{\partial L}{\partial b} = -a\sum M_b(t_i) + \sum \frac{m_b(t_i)}{m(t_i)} = 0 \tag{4.5}$$

$$\frac{\partial L}{\partial c} = -a\sum M_c(t_i) + \sum \frac{m_c(t_i)}{m(t_i)} = 0 \tag{4.6}$$

where $M_b(t) = \frac{\partial}{\partial b}M(t)$, $M_c(t) = \frac{\partial}{\partial c}M(t)$, $m_b(t) = \frac{\partial}{\partial b}m(t)$ and $m_c(t) = \frac{\partial}{\partial c}m(t)$.

Take second derivatives

$$\frac{\partial^2 L}{\partial a^2} = -\frac{n}{a^2} \tag{4.7}$$

$$\frac{\partial^2 L}{\partial b^2} = -a\sum M_{bb}(t_i) + \sum \frac{m_{bb}(t_i)}{m(t_i)} - \sum \left[\frac{m_b(t_i)}{m(t_i)}\right]^2 \tag{4.8}$$

$$\frac{\partial^2 L}{\partial c^2} = -a\sum M_{cc}(t_i) + \sum \frac{m_{cc}(t_i)}{m(t_i)} - \sum \left[\frac{m_c(t_i)}{m(t_i)}\right]^2 \tag{4.9}$$

$$\frac{\partial^2 L}{\partial a \partial b} = -\sum M_b(t_i) \tag{4.10}$$

$$\frac{\partial^2 L}{\partial a \partial c} = -\sum M_c(t_i) \tag{4.11}$$

$$\frac{\partial^2 L}{\partial b \partial c} = -a\sum M_{bc}(t_i) + \sum \frac{m_{bc}(t_i)}{m(t_i)} - \sum \frac{m_b(t_i)m_c(x_i)}{[m(t_i)]^2} \tag{4.12}$$

where M_{bb}, M_{bc}, M_{cc}, m_{bb}, m_{bc} and m_{cc} are the second and mixed partial derivatives of $M(t)$ and $m(t)$ with regard to *b* and *c*, respectively.

The Observed Fisher Information matrix is the negative Hessian matrix at the MLE point $(\hat{a}, \hat{b}, \hat{c})$

$$I = -H(\hat{a}, \hat{b}, \hat{c}) = -\left.\frac{\partial^2 L(a,b,c)}{\partial(a,b,c)\partial(a,b,c)'}\right|_{(\hat{a},\hat{b},\hat{c})} \tag{4.13}$$

At MLE $(\hat{a}, \hat{b}, \hat{c})$, likelihood equations (4.4) – (4.6) hold. From (4.5) and (4.6)

$$\sum M_b(t_i) = \frac{1}{a}\sum \frac{m_b(t_i)}{m(t_i)} \tag{4.14}$$

$$\sum M_c(t_i) = \frac{1}{a}\sum \frac{m_c(t_i)}{m(t_i)} \tag{4.15}$$

When $M(t)$ is a multinomial, exponential or some other functions of *t*, as in the modified Weibull and Weibull extension cases, $\frac{m_b(t)}{m(t)}$ and $\frac{m_c(t)}{m(t)}$ in equations (4.14) and (4.15) are usually simpler than $M_b(t)$ and $M_c(t)$ because high order or exponential terms of *t* approach 0 after the dividing operation. Therefore, we can simplify the computation of (4.10) and (4.11) and thus the Observed Fisher Information matrix I by introducing (4.14) and (4.15)

$$\frac{\partial^2 L}{\partial a \partial b} = -\frac{1}{a}\sum \frac{m_b(t_i)}{m(t_i)} \tag{4.16}$$

$$\frac{\partial^2 L}{\partial a \partial c} = -\frac{1}{a}\sum \frac{m_c(t_i)}{m(t_i)} \tag{4.17}$$

However, this simplification procedure only makes sense at the MLE point. When $(a_0, b_0, c_0) \neq (\hat{a}, \hat{b}, \hat{c})$, equations (4.14) and (4.15) generally do not hold. In such cases, (4.16) and (4.17) also do not hold. Consequently, the simplification can only be used for the computation of the approximated variance-covariance matrix of the MLE. In the next section 4.2, another technique is proposed to handle the problem encountered in the MLE searching procedure.

Given a complete sample $t_1, t_2, \cdots, t_n$ from the modified Weibull, the log-likelihood function (3.12) as shown in chapter 3 is

$$L(a, b, \lambda) = n \log a + (b-1)\sum \log t_i + \lambda \sum t_i + \sum \log(b + \lambda t_i) - a\sum t_i^b e^{\lambda t_i} \tag{4.18}$$

At the MLE point, equation (4.14) and (4.15) are

$$\sum t_i^b e^{\lambda t_i} \log t_i = \frac{1}{a}\sum \frac{1}{b + \lambda t_i} + \frac{1}{a}\sum \log t_i \tag{4.19}$$

$$\sum t_i^{b+1} e^{\lambda t_i} = \frac{1}{a}\sum \frac{t_i}{b+\lambda t_i} + \frac{1}{a}\sum t_i \tag{4.20}$$

It is obvious that the right sides of equations (4.19) and (4.20) are easier to compute than the left sides. Substituting them into (4.16) and (4.17), we obtain

$$\frac{\partial^2 L}{\partial a \partial b} = -\frac{1}{a}\sum \frac{1}{b+ct_i} - \frac{1}{a}\sum \log t_i \tag{4.21}$$

$$\frac{\partial^2 L}{\partial a \partial c} = -\frac{1}{a}\sum \frac{t_i}{b+ct_i} - \frac{1}{a}\sum t_i \tag{4.22}$$

4.2 Simplification of the Log-likelihood Function

The evaluation of the MLE of the parameters usually involves iterative steps when there are more than one model parameters. In such case, the maximization process continues until all the likelihood equations are satisfied. However, this process is often not easy and needs a lot of numerical computation, and even worse sometimes the process does not converge. In this section, a technique is proposed to help deal with this problem and is shown to be useful for a broad class of distributions.

We extend the class (4.1) a bit to accommodate more distributions

$$G(t) = 1 - \exp\{-aM(t)\} \text{ or } G(t) = \{1 - \exp[-M(t)]\}^a \tag{4.23}$$

where $M(t)$ is an increasing function of *t* with parameter vector $\boldsymbol{\theta}$ but without *a*. (4.23) implies a log-linear relationship between CDF or SF of the aiming distribution and CDF or SF of the distribution without parameter *a*. The exponentiated Weibull belongs to this class of distributions.

Now the log-likelihood function for the former case of (4.23) is (for the latter case the discussion is similar, so we omit it here for brevity)

$$L(a,\boldsymbol{\theta}) = n\log a + \sum \log(m(t_i)) - a\sum M(t_i) \tag{4.24}$$

Similarly, likelihood equations are derived

$$\frac{\partial L}{\partial a} = \frac{n}{a} - \sum M(t_i) = 0 \tag{4.25}$$

$$\frac{\partial L}{\partial \boldsymbol{\theta}} = -a\sum M_{\boldsymbol{\theta}}^{/}(t_i) + \sum \frac{m_{\boldsymbol{\theta}}^{/}(t_i)}{m(t_i)} = 0 \tag{4.26}$$

Note that (4.26) may be composed of several equations, depending on the dimensionality of $\boldsymbol{\theta}$.

At the MLE point $(\hat{a},\hat{\boldsymbol{\theta}})$, equation (4.25) holds (since $\frac{\partial^2 L}{\partial a^2} = -\frac{n}{a^2} < 0$), so

$$a = \frac{n}{\sum M(t_i)} \tag{4.27}$$

Substituting (4.27) into (4.26), we obtain the following equation(s)

$$\frac{\partial L}{\partial \boldsymbol{\theta}} = -\frac{n\sum M_{\boldsymbol{\theta}}'(t_i)}{\sum M(t_i)} + \sum \frac{m_{\boldsymbol{\theta}}'(t_i)}{m(t_i)} = 0 \tag{4.28}$$

It automatically follows that equations (4.25) and (4.26) are equivalent to (4.27) and (4.28).

Based on (4.28), omitting the constant term $n\log n - n$, we write the concentrated log-likelihood function

$$L^*(\boldsymbol{\theta}) = -n\log\left[\sum M(t_i)\right] + \sum \log\left[m(t_i)\right] \tag{4.29}$$

We say a log-likelihood function can achieve its regular maximum when at the MLE point the corresponding derivatives of the log-likelihood function with respect to the parameters are equal to zero, i.e. likelihood equations hold.

Lemma 4.1 Maximizing $L(a,\boldsymbol{\theta})$ is equivalent to maximizing $L^*(\boldsymbol{\theta})$. Hence, the MLE of $\boldsymbol{\theta}$ in $L(a,\boldsymbol{\theta})$ are the same as the MLE of $L^*(\boldsymbol{\theta})$.

Proof: Suppose $(a_1,\boldsymbol{\theta}_1)$ maximize $L(a,\boldsymbol{\theta})$. Let Θ denote the parameter space of $\boldsymbol{\theta}$. If $\boldsymbol{\theta}_2 \in \Theta$ maximizes $L^*(\boldsymbol{\theta})$ and $L^*(\boldsymbol{\theta}_2) > L^*(\boldsymbol{\theta}_1)$, we let $a_2 = \frac{n}{\sum M(t_i)\big|_{\boldsymbol{\theta}_2}}$,

then

$$
\begin{aligned}
L(a_2,\boldsymbol{\theta}_2) &= n\log n - n\log\left(\sum M(t_i)\big|_{\boldsymbol{\theta}_2}\right) + \sum \log\left(m(t_i)\big|_{\boldsymbol{\theta}_2}\right) - n \\
&= n\log n - n + L^*(\boldsymbol{\theta}_2) \\
&> n\log n - n + L^*(\boldsymbol{\theta}_1) \\
&= L(a_1,\boldsymbol{\theta}_1)
\end{aligned}
$$

The last equality holds because of (4.27).

This contradicts the assumption that $L(a_1,\boldsymbol{\theta}_1)$ is the maximum, so $\boldsymbol{\theta}_1$ maximizes $L^*(\boldsymbol{\theta})$.

On the contrary, suppose $\boldsymbol{\theta}_1^*$ maximizes $L^*(\boldsymbol{\theta})$, then $(a_1^*,\boldsymbol{\theta}_1^*)$ maximize $L(a,\boldsymbol{\theta})$, where $a_1^* = \dfrac{n}{\sum M(t_i)\big|_{\boldsymbol{\theta}_1^*}}$. Otherwise, if $(a_2^*,\boldsymbol{\theta}_2^*)$ maximize $L(a,\boldsymbol{\theta})$ and $L(a_2^*,\boldsymbol{\theta}_2^*) > L(a_1^*,\boldsymbol{\theta}_1^*)$, from (4.27) we have $a_2^* = \dfrac{n}{\sum M(t_i)\big|_{\boldsymbol{\theta}_2^*}}$, and then

$$
\begin{aligned}
L^*(\boldsymbol{\theta}_2^*) &= -n\log\left(\sum M(t_i)\big|_{\boldsymbol{\theta}_2^*}\right) + \sum \log\left(m(t_i)\big|_{\boldsymbol{\theta}_2^*}\right) \\
&= n - n\log n + n\log a_2^* + \sum \log\left(m(t_i)\big|_{\boldsymbol{\theta}_2^*}\right) - a\sum M(t_i)\big|_{\boldsymbol{\theta}_2^*} \\
&= n - n\log n + L(a_2^*,\boldsymbol{\theta}_2^*) \\
&> n - n\log n + L(a_1^*,\boldsymbol{\theta}_1^*) \\
&= L^*(\boldsymbol{\theta}_1^*)
\end{aligned}
$$

This contradicts the assumption that $L^*(\boldsymbol{\theta}_1^*)$ is the maximum, so $(a_1^*,\boldsymbol{\theta}_1^*)$ maximize $L(a,\boldsymbol{\theta})$. Q.E.D.

4.3 Log-Likelihood Function of the Modified Weibull Distribution

A log-likelihood function $L(\theta)$ is said to have a local maximum at point $\hat{\theta}$, if there exists some $\varepsilon > 0$, such that $L(\hat{\theta}) \geq L(\theta)$ for $\theta \in \Theta$ when $|\theta - \hat{\theta}| < \varepsilon$, where Θ is the parameter space of the distribution function. If $L(\theta)$ has only one local maximum, then $L(\hat{\theta})$ is the unique maximum. If for all $\theta \in \Theta$, $L(\hat{\theta}) \geq L(\theta)$, then $\hat{\theta}$ is called the global maximum point of $L(\theta)$. Obviously, a unique maximum is also a global maximum, but a global maximum is not necessarily the unique local maximum.

As pointed out by Makelainen *et al.* (1981), the occurrence of several local maxima of the likelihood would result in the unwanted situation that "summarization of the data by means of a maximum likelihood estimate and its asymptotic variance could be very misleading". Examples of this case include the 2-parameter Cauchy distribution (Barnett, 1966). In addition, conventional asymptotic inferential procedures require that the global maximum point be located as an interior solution to the likelihood equations. Therefore, for the sake of simple numerical optimization of the likelihood and statistical inference of the parameters, analytical results regarding the uniqueness property of the MLE is of practical importance.

For the modified Weibull distribution, Ng (2005) discussed ML estimation of the model parameters for progressively type-2 censored data. The paper recommended iterative steps for the MLE searching, but did neither verify that the procedure can reach the estimate, nor prove that the estimates obtained by solving the likelihood equations maximize the log-likelihood. Bebbington *et al*. (2008) showed that the determinant of the Fisher information matrix is not everywhere positive, so claimed that the MLE of the parameters of the modified Weibull distribution do not always exist. However, this assertion is not accurate. In the rest of this chapter, the techniques developed above are applied to the modified Weibull distribution to prove the existence and uniqueness of the MLE of the model parameters.

Progressively type-2 censoring is a natural generalization of the complete and single right censoring schemes, but has a lot of practical applications. Given any progressively type-2 censored sample data $\{t_1 < \cdots < t_m\}$ $\{r_1, \cdots, r_m\}$, the log-likelihood function under modified Weibull assumption is

$$L(a,b,\lambda) = m\log(a) + (b-1)\sum_{i=1}^{m}\log(t_i) + \lambda\sum_{i=1}^{m} t_i + \sum_{i=1}^{m}\log(b+\lambda t_i) - \sum_{i=1}^{m}(1+r_i)at_i^b e^{\lambda t_i} \tag{4.30}$$

We make a natural assumption that not all t_i are identical, so $m \geq 2$. In fact m has to be at least 3 in order to be large enough to validate the parameter estimation endeavor.

Substituting $a = \dfrac{m}{\sum_{i=1}^{m}(1 + r_i)t_i^b e^{\lambda t_i}}$ into (4.30), the log-likelihood function $L(a, b, \lambda)$ is transformed to with only two unknown parameters

$$L^*(b, \lambda) = \sum_{i=1}^{m} \log(b + \lambda t_i) + (b-1)\sum_{i=1}^{m} \log t_i + \lambda \sum_{i=1}^{m} t_i - m \log\left(\sum_{i=1}^{m}(1 + r_i)t_i^b e^{\lambda t_i}\right) \quad (4.31)$$

From Lemma 4.1, maximizing $L(a, b, \lambda)$ is equivalent to maximizing $L^*(b, \lambda)$.

Before going into the details of $L^*(b, \lambda)$, we present the results of Makelainen *et al*. (1981) as preliminaries.

4.4 Preliminaries

According to the result in Makelainen *et al*. (1981), in order to prove the MLE of the log-likelihood function $L(\boldsymbol{\theta})$ exist and are unique, one needs to show $L(\boldsymbol{\theta})$ is constant on the boundary of the parameter space and its Hessian matrix is negative-definite everywhere. That is,

$$H(\boldsymbol{\theta}) = \frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \quad (4.32)$$

- The determinant of the upper left 1-by-1 corner of $H(\boldsymbol{\theta})$ is negative;
- The determinant of the upper left 2-by-2 corner of $H(\boldsymbol{\theta})$ is positive ;
- The determinant of the upper left 3-by-3 corner of $H(\boldsymbol{\theta})$ is negative;
- …

In the single dimensional case, the above condition says that $L(\boldsymbol{\theta})$ has the same values or limits in its both tails (constant on the boundary) and it is concave everywhere (Hessian matrix negative-definite).

In the multiple dimensional case, concavity is replaced by negative-definiteness.

We found that the Hessian matrix $H(a,b,\lambda)$ as shown in (3.22) is not everywhere negative-definite, which is also observed by Bebbington *et al*. (2008), though what they considered was the Expected Fisher Information matrix, instead of the Observed Fisher Information matrix, which is negative Hessian.

However, as will be seen in the next section, the transformed Hessian matrix $H^*(b,\lambda)$ is indeed negative-definite everywhere.

4.5 Existence and Uniqueness of MLE

Deriving (4.31) with respect to b and λ, we have the new likelihood equations

$$\frac{\partial L^*}{\partial b} = \sum_{i=1}^{m} \log(t_i) + \sum_{i=1}^{m} \frac{1}{b+\lambda t_i} - \frac{m\sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i} \log t_i}{\sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}} = 0 \tag{4.33}$$

$$\frac{\partial L^*}{\partial \lambda} = \sum_{i=1}^{m} t_i + \sum_{i=1}^{m} \frac{t_i}{b+\lambda t_i} - \frac{m\sum_{i=1}^{m}(1+r_i)t_i^{b+1} e^{\lambda t_i}}{\sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}} = 0 \tag{4.34}$$

It should be noted here that a λ which makes (4.34) true could be negative. Hence, we extend the parameter space Θ_1 of the log-likelihood function $L^*(b,\lambda)$

$$\Theta_1 = \{(b,\lambda): b,\lambda \geq 0\} \tag{4.35}$$

to data dependent

$$\Theta^* = \{(b,\lambda): b+\lambda t_i > 0, i = 1,\cdots,m\} \tag{4.36}$$

Hence, the boundary of the new parameter space Θ^* is

$$\partial\Theta^* = \{b=\infty\} \cup \{\lambda = \infty\} \cup \{(b,\lambda): b>0, \lambda = -b/t_m\} \cup \{(b,\lambda): \lambda > 0, b = -\lambda t_1\} \tag{4.37}$$

4.5.1 Constancy on the Boundary

In this section, we will show that $L^*(b,\lambda)$ approaches $-\infty$ on each part of the boundary.

Lemma 4.2 If $m \ge 2$, $\lim_{b\to\infty} \sup_{\lambda > -b/t_m} L^*(b,\lambda) = -\infty$.

Proof: From (4.31), given $b > 0$, we have

$$\begin{aligned} L^*(b,\lambda) &= \sum_{i=1}^{m} \log(b+\lambda t_i) + (b-1)\sum_{i=1}^{m} \log t_i + \lambda \sum_{i=1}^{m} t_i - m\log\left(\sum_{i=1}^{m} (1+r_i) t_i^b e^{\lambda t_i}\right) \\ &< m\log(b+\lambda t_m) + (b-1)\sum_{i=1}^{m} \log t_i + \lambda \sum_{i=1}^{m} t_i - m\log\left((1+r_m) t_m^b e^{\lambda t_m}\right) \end{aligned} \tag{4.38}$$

Note t_m is the largest observed failure time.

Let $g_b(\lambda) = m\log(b+\lambda t_m) + (b-1)\sum_{i=1}^{m} \log t_i + \lambda \sum_{i=1}^{m} t_i - mb\log t_m - m\lambda t_m$ (4.39)

Then the maximum of $g_b(\lambda)$ achieves at the point λ^* where $g_b'(\lambda^*) = 0$, or

$$\lambda^* = \frac{m}{mt_m - \sum_{i=1}^{m} t_i} - \frac{b}{t_m} \tag{4.40}$$

We have

$$g_b\left(\lambda^*\right) = m\log\left(\frac{m}{mt_m - \sum_{i=1}^{m} t_i}\right) - m - \sum_{i=1}^{m}\log t_i + \left(m - \sum_{i=1}^{m}\frac{t_i}{t_m} + \sum_{i=1}^{m}\log t_i - m\log t_m\right)b \tag{4.41}$$

Let $s(x) = 1 - x + \log x$, $s'(x) = \frac{1}{x} - 1 > 0$ for $0 < x < 1$, so $s(x) < s(1) = 0$ for $0 < x < 1$.

Hence, because $t_m > \cdots > t_1$, we have

$$m - \sum_{i=1}^{m}\frac{t_i}{t_m} + \sum_{i=1}^{m}\log t_i - m\log t_m = \sum_{i=1}^{m} s\left(\frac{t_i}{t_m}\right) < 0 \tag{4.42}$$

Therefore, $\lim_{b\to\infty}\sup_{\lambda > -b/t_m} g_b(\lambda) = -\infty$, and since $L^*(b,\lambda) < g_b(\lambda) - m\log(1 + r_m)$, $\lim_{b\to\infty}\sup_{\lambda > -b/t_m} L^*(b,\lambda) = -\infty$. Q.E.D.

Lemma 4.3 If $m \ge 2$, $\lim_{\lambda\to\infty}\sup_{b > -\lambda t_1} L^*(b,\lambda) = -\infty$.

The proof is similar to lemma 4.2, so omitted here. Q.E.D.

Lemma 4.4 $L^*(b,\lambda)\Big|_{\{(b,\lambda): b>0, \lambda = -b/t_m\}} = -\infty$.

Proof: This is true because $\log(b + \lambda t_m)\big|_{\{(b,\lambda):b>0,\lambda=-b/t_m\}} = -\infty$ and other terms are infinite. Q.E.D.

Lemma 4.5 $L^*(b,\lambda)\big|_{\{(b,\lambda):\lambda>0,b=-\lambda t_1\}} = -\infty$.

Proof: This is true because $\log(b + \lambda t_1)\big|_{\{(b,\lambda):\lambda>0,b=-\lambda t_1\}} = -\infty$ and other terms are infinite. Q.E.D.

Theorem 4.6 If $m \geq 2$, $\lim_{(b,\lambda)\to\partial\Theta} L^*(b,\lambda) \to -\infty$, then the log-likelihood function is constant on the boundary $\partial\Theta^*$ of the parameter space Θ^*.

Proof: This is the direct result of lemmas 4.2 – 4.5. Q.E.D.

4.5.2 Negative-Definiteness of Hessian Matrix *H**(*b*,*λ*)

Differentiating (4.33) and (4.34) with respect to b and λ, we get the second and mixed derivatives

$$\frac{\partial^2 L^*}{\partial b^2} = -\sum_{i=1}^{m}\left(\frac{1}{b+\lambda t_i}\right)^2 - \frac{m\sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}(\log t_i)^2}{\sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}} + \frac{m\left(\sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}\log t_i\right)^2}{\left(\sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}\right)^2}$$

(4.43)

$$\frac{\partial^2 L^*}{\partial \lambda^2} = -\sum_{i=1}^{m}\left(\frac{t_i}{b+\lambda t_i}\right)^2 - \frac{m\sum_{i=1}^{m}(1+r_i)t_i^{b+2}e^{\lambda t_i}}{\sum_{i=1}^{m}(1+r_i)t_i^{b}e^{\lambda t_i}} + \frac{m\left(\sum_{i=1}^{m}(1+r_i)t_i^{b+1}e^{\lambda t_i}\right)^2}{\left(\sum_{i=1}^{m}(1+r_i)t_i^{b}e^{\lambda t_i}\right)^2} \tag{4.44}$$

$$\begin{aligned}\frac{\partial^2 L^*}{\partial b \partial \lambda} = &-\sum_{i=1}^{m}\frac{t_i}{(b+\lambda t_i)^2} - \frac{m\sum_{i=1}^{m}(1+r_i)t_i^{b+1}e^{\lambda t_i}\log t_i}{\sum_{i=1}^{m}(1+r_i)t_i^{b}e^{\lambda t_i}} \\ &+ \frac{m\left(\sum_{i=1}^{m}(1+r_i)t_i^{b+1}e^{\lambda t_i}\right)\sum_{i=1}^{m}(1+r_i)t_i^{b+1}e^{\lambda t_i}\log t_i}{\left(\sum_{i=1}^{m}(1+r_i)t_i^{b}e^{\lambda t_i}\right)^2}\end{aligned} \tag{4.45}$$

Then the Hessian matrix is

$$H^*(b,\lambda) = \left(\frac{\partial^2 L^*(b,\lambda)}{\partial(b,\lambda)\partial(b,\lambda)'}\right) = \begin{pmatrix} \frac{\partial^2 L^*}{\partial b^2} & \frac{\partial^2 L^*}{\partial b \partial \lambda} \\ \frac{\partial^2 L^*}{\partial b \partial \lambda} & \frac{\partial^2 L^*}{\partial \lambda^2} \end{pmatrix} \tag{4.46}$$

Lemma 4.7 The upper left 1-by-1 corner $\left(\frac{\partial^2 L^*}{\partial b^2}\right)$ is negative.

Proof: We write (4.43) here

$$\frac{\partial^2 L^*}{\partial b^2} = -\sum_{i=1}^{m}\left(\frac{1}{b+\lambda t_i}\right)^2 - \frac{m\sum_{i=1}^{m}(1+r_i)t_i^{b}e^{\lambda t_i}(\log t_i)^2}{\sum_{i=1}^{m}(1+r_i)t_i^{b}e^{\lambda t_i}} + \frac{m\left(\sum_{i=1}^{m}(1+r_i)t_i^{b}e^{\lambda t_i}\log t_i\right)^2}{\left(\sum_{i=1}^{m}(1+r_i)t_i^{b}e^{\lambda t_i}\right)^2}$$

Applying the Cauchy-Schwarz inequality, that is

$$\left(\sum a_i b_i\right)^2 \le \left(\sum a_i^2\right)\left(\sum b_i^2\right)$$

with

$$a_i = \sqrt{(1+r_i)t_i^b e^{\lambda t_i}}\ \log t_i ,\ b_i = \sqrt{(1+r_i)t_i^b e^{\lambda t_i}}\ ,$$

We have

$$\left(\sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}\ \log t_i\right)^2 \le \sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}(\log t_i)^2 \cdot \sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}\ ,$$

so we obtain

$$-\frac{m\sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}(\log t_i)^2}{\sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}} + \frac{m\left(\sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}\ \log t_i\right)^2}{\left(\sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}\right)^2} \le 0 .$$

Moreover, since $-\sum_{i=1}^{m}\left(\frac{1}{b+\lambda t_i}\right)^2 < 0$, so $\left(\frac{\partial^2 L^*}{\partial b^2}\right) < 0$. Q.E.D.

Lemma 4.8 If $m \ge 2$, $\det\left(H^*(b,\lambda)\right) > 0$ for $(b,\lambda) \in \Theta^*$.

Proof: To simplify notation, the following symbols are used:

$$T = \sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i} \qquad T_1 = \sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}\ \log t_i \qquad T_2 = \sum_{i=1}^{m}(1+r_i)t_i^b e^{\lambda t_i}(\log t_i)^2$$

$$T_3 = \sum_{i=1}^{m}(1+r_i)t_i^{b+1}e^{\lambda t_i} \qquad T_4 = \sum_{i=1}^{m}(1+r_i)t_i^{b+2}e^{\lambda t_i} \qquad T_5 = \sum_{i=1}^{m}(1+r_i)t_i^{b+1}e^{\lambda t_i}\log t_i$$

$$S_1 = \sum_{i=1}^{m}\left(\frac{1}{b+\lambda t_i}\right)^2 \qquad S_2 = \sum_{i=1}^{m}\left(\frac{t_i}{b+\lambda t_i}\right)^2 \qquad S_3 = \sum_{i=1}^{m}\frac{t_i}{(b+\lambda t_i)^2}$$

$$R_1 = \frac{TT_2 - T_1^2}{T^2} \qquad R_2 = \frac{TT_4 - T_3^2}{T^2} \qquad R_3 = \frac{TT_5 - T_1T_3}{T^2}$$

Hence,

$$\det\left(H^*(b,\lambda)\right) = \frac{\partial^2 L^*}{\partial b^2}\frac{\partial^2 L^*}{\partial \lambda^2} - \left(\frac{\partial^2 L^*}{\partial b \partial \lambda}\right)^2$$
$$= \left(S_1 + \frac{m\left(TT_2 - T_1^2\right)}{T^2}\right)\left(S_2 + \frac{m\left(TT_4 - T_3^2\right)}{T^2}\right) - \left(S_3 + \frac{m\left(TT_5 - T_1T_3\right)}{T^2}\right)^2$$

Expanding it, we have

$$L_1 = S_1S_2 - S_3^2 = \sum_{i=2}^{m}\sum_{j=1}^{i-1}\frac{(t_i - t_j)^2}{(b+\lambda t_i)^2(b+\lambda t_j)^2} > 0 \tag{4.47}$$

$$L_2 = R_1R_2 - R_3^2 = \frac{1}{T^3}\left(TT_2T_4 + 2T_1T_3T_5 - T_1^2T_4 - T_2T_3^2 - TT_5^2\right)$$
$$= \frac{1}{T^3}\sum_{i=3}^{m}\sum_{j=2}^{i}\sum_{k=1}^{j}(1+r_i)(1+r_j)(1+r_k)t_i^b e^{\lambda t_i} t_j^b e^{\lambda t_j} t_k^b e^{\lambda t_k}$$
$$\left(t_i\log t_j + t_j\log t_k + t_k\log t_i - t_i\log t_k - t_j\log t_i - t_k\log t_j\right)^2$$
$$> 0 \tag{4.48}$$

Finally, since $TT_4 - T_2^2 = \sum_{i=2}^{m}\sum_{j=1}^{i-1}(1+r_i)(1+r_j)t_i^b e^{\lambda t_i} t_j^b e^{\lambda t_j}(t_i - t_j)^2 > 0$, together with equations (4.47) and (4.48), we have

$$
\begin{aligned}
L_3 = S_1 R_2 + S_2 R_1 - 2S_3 R_3 &> \frac{S_3^2}{S_2} R_2 + S_2 \frac{R_3^2}{R_2} - 2S_3 R_3 \\
&= \left(S_3 \sqrt{\frac{R_2}{S_2}} - R_3 \sqrt{\frac{S_2}{R_2}} \right)^2 \geq 0
\end{aligned}
\tag{4.49}
$$

In sum, $\frac{\partial^2 L^*}{\partial b^2}\frac{\partial^2 L^*}{\partial \lambda^2} - \left(\frac{\partial^2 L^*}{\partial b \partial \lambda}\right)^2 = L_1 + m^2 L_2 + mL_3 > 0$.

This completes the proof of lemma 4.8. Q.E.D.

Lemmas 4.7 and 4.8 lead to the following theorem 4.9.

Theorem 4.9 The Hessian matrix $H^*(b,\lambda)$ is negative-definite at every point of $\Theta^* = \{(b,\lambda): b + \lambda t_i > 0, i = 1,\cdots,m\}$.

4.5.3 Existence and Uniqueness of MLE

From theorem 4.6 and 4.9, and the theorem of Makelainen *et al.* (1981), the existence and uniqueness of the MLE of parameters (b,λ) for the log-likelihood function $L^*(b,\lambda)$ is guaranteed. We have the following main result of this chapter.

Theorem 4.10 Given a progressively type-2 censored sample $\{t_1 < \cdots < t_m\}$ $\{r_1, \cdots, r_m\}$, where $m \geq 3$, the MLE of the parameters of the modified Weibull distribution exist in the parameter space $\Theta = \{(a, b, \lambda): a > 0, b + \lambda t_i > 0, i = 1, \cdots, m\}$ and are unique.

Proof: This is simply because $\hat{a}$ is determined by $\hat{b}$ and $\hat{\lambda}$. Q.E.D.

Theorem 4.10 shows that given any progressively type-2 censored sample, for the modified Weibull distribution we can define a new parameter space Θ which includes the original one $\Theta_0 = \{(a, b, \lambda): a > 0, b \geq 0, \lambda \geq 0\}$ as subspace. With such definition, the MLE of the parameters exist and are unique.

As can be seen, for some sample data, the obtained MLE may not necessarily reside in Θ_0. In such case, the likelihood equations (3.13), (3.14) and (3.15) do not have common non-negative solutions and the fitted modified Weibull model with these MLE is not suitable to model lifetime data. However, we can treat it as a constraint optimization problem subject to inequality constraints $a > 0, b \geq 0, \lambda \geq 0$ and get the MLE satisfying these regularity conditions which maximize the log-likelihood.

4.6 Illustrative Examples

In this section we present several examples to show that given progressively type-2 censored samples the MLE of the modified Weibull parameters exist and are unique.

4.6.1 Data from Aarset (1987)

It was shown in Aarset (1987) that the TTT plot of the lifetimes of the 50 devices indicates a bathtub-shaped failure rate, thus it is appropriate to use the modified Weibull distribution to model the data. The dataset is given as follows.

From table 3.2, we can see that many data coincide, which means that the data might be treated as progressively type-2 censored sample in table 3.3. Hence, 35 failure times can be withdrawn from the table, and the numbers of censored units are $r_3 = 4$, $r_{10} = 4$, $r_{23} = 3$, $r_{33} = 4$, and $r_i = 0, i \neq 3, 10, 23, 33$.

Maximizing the log-likelihood function subject to the non-negative constraints, the MLE of the parameters are $\left(\hat{a}, \hat{b}, \hat{\lambda}\right) = (0.0376, 0.3398, 0.0263)$.

To show the pattern of the likelihood, we plot the log-likelihood function $L^*(b, \lambda)$ in the neighborhood of the MLE $\left(\hat{b}, \hat{\lambda}\right) = (0.3398, 0.0263)$. Since the MLE of b

and λ are 0.3398 and 0.0263 respectively, the plotting area of the two parameters is confined to a subspace $[0,1]\times[0,0.1]$ of the first quadrant. In order to describe the likelihood, we draw two plots, a surface plot and a contour plot.

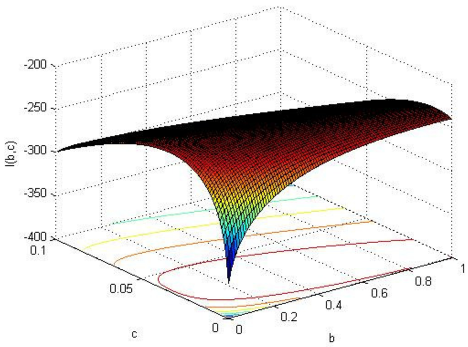


Figure 4.1 Surface plot of $L_1(b, \lambda)$

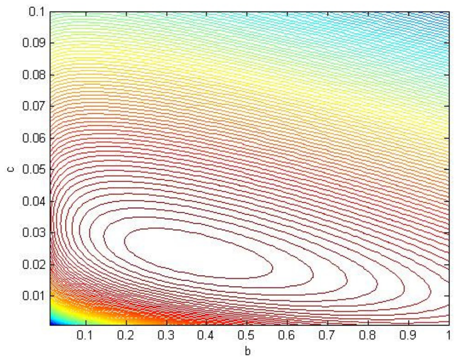


Figure 4.2 Contour plot of $L_1(b, \lambda)$

From the two plots, it is easy to see that $L^*(b,\lambda)$ has only one local maximum in the area $[0,1]\times[0,0.1]$, and it spreads out to the whole $[0,\infty)\times[0,\infty)$ space like climbing down a hill.

4.6.2 A Simulated Example

As another illustrative example, we generated a progressively type-2 censored sample of size 30, where $t=(1.2719,1.8784,3.2781,4.0952,6.2069,9.3385)$ and $r=(4,4,4,4,4,4)$ with model parameters $(a,b,\lambda)=(0.1,1,0.1)$.

Constraint optimization yields MLE $\left(\hat{a},\hat{b},\hat{\lambda}\right)=\left(0.014,1.6729,0\right)$, and the log-likelihood at this point is -23.4387. However, we note that the likelihood equations (4.33) and (4.34) do not hold when $\left(\hat{b},\hat{\lambda}\right)=\left(1.6729,0\right)$. Therefore, we extend the parameter space to Θ^* and maximize $L^*\left(b,\lambda\right)$ in this region. $\left(b,\lambda\right)$ that maximize $L^*\left(b,\lambda\right)$ are located at $\left(\hat{b}_1,\hat{\lambda}_1\right)=\left(1.7621,-0.0189\right)$. At this point the likelihood equations hold and the log-likelihood is -23.4353.

It is interesting to see that the "regular" MLE $\left(\hat{b}_1,\hat{\lambda}_1\right)$ that maximize the log-likelihood and also maintain the likelihood equations are not in the first quarter of the $b-\lambda$ space, so these parameters $\left(\hat{a}_1\hat{b}_1,\hat{\lambda}_1\right)=\left(0.0136,1.7621,-0.0189\right)$ are not suitable for the modified Weibull distribution to model lifetime data. The appropriate estimates of the parameters are $\left(\hat{a},\hat{b},\hat{\lambda}\right)=\left(0.014,1.6729,0\right)$, which reduce the modified Weibull to the Weibull distribution.

4.7 Negative MLE

As we have shown in the previous section, it is possible that the estimates obtained by directly maximizing the log-likelihood function (4.31) over the parameter space Θ do not conform to the non-negativity conditions of the model parameters, i.e. either $\hat{b}<0$ or $\hat{\lambda}<0$. Therefore constraint optimization techniques are required to ensure the parameter estimates are non-negative. In this section, we will do a simulation study to look into this phenomenon and discuss the relationship between the percentage of 0 estimated values, which refers to the negative estimates if they were obtained by directly maximizing the log-likelihood, and the magnitudes of the parameter settings and sample sizes.

We note that if T follows a modified Weibull distribution with parameters (a, b, λ), then $a^{1/b}T$ is also modified Weibull distributed, but with model parameters (1, b, $\lambda a^{-1/b}$). This is because

$$\begin{aligned}\Pr\left(a^{1/b}T > t\right) &= \Pr\left(T > a^{-1/b}t\right)\\ &= \exp\left(-t^b e^{\lambda a^{-1/b} t}\right)\end{aligned}$$

Since rescaling a random variable T and a random sample t does not alter the estimation of b and changes the estimate of λ proportionately, we can simplify the simulation by generating data from parameters (1, b, λ), only changing the values of b and λ, and estimating the parameters based on these generated samples.

The values of the parameters b and λ are picked from (0.1, 0.2, 0.5, 1, 2) × (0.1, 0.2, 0.5, 1, 2), and for each pair of the parameters, we generate 5000 samples with sample size of 10, 20, 50 and 100. We calculate the maximum likelihood estimates of the parameters for each of the samples and count the number of estimates which are zero. The simulation results are as follows (since $\hat{a}$ must be positive, it is not included in the following tables). The figure in the upper right corner of each cell is the number of instances where $\hat{b} = 0$, while the figure in the lower left corner is the number of instances where $\hat{\lambda} = 0$, out of 5000.

Table 4.1 Number of Zero Estimates for Sample Size 10

λ \ b	0.1	0.2	0.5	1	2
0.1	1 1	0 28	0 533	13 1458	24 2049
0.2	0 4	0 32	3 491	10 1263	35 1934
0.5	1 3	0 25	2 416	8 1107	44 1777
1	0 4	1 30	5 321	8 1031	90 1743
2	0 2	0 31	1 365	8 985	158 1669

Table 4.2 Number of Zero Estimates for Sample Size 20

λ \ b	0.1	0.2	0.5	1	2
0.1	0 0	0 2	0 314	0 1303	6 1981
0.2	0 0	0 2	0 237	0 1042	2 1851
0.5	0 0	0 0	0 139	1 807	3 1648
1	0 0	0 1	0 108	5 625	7 1526
2	0 0	0 0	0 97	4 557	21 1389

Table 4.3 Number of Zero Estimates for Sample Size 50

λ \ b	0.1	0.2	0.5	1	2
0.1	0 / 0	0 / 0	0 / 80	0 / 1075	0 / 1925
0.2	0 / 0	0 / 0	0 / 28	0 / 702	0 / 1689
0.5	0 / 0	0 / 0	0 / 6	0 / 353	0 / 1344
1	0 / 0	0 / 0	0 / 2	0 / 209	0 / 1025
2	0 / 0	0 / 0	0 / 1	0 / 125	1 / 841

Table 4.4 Number of Zero Estimates for Sample Size 100

λ \ b	0.1	0.2	0.5	1	2
0.1	0 / 0	0 / 0	0 / 7	0 / 836	0 / 1859
0.2	0 / 0	0 / 0	0 / 0	0 / 419	0 / 1563
0.5	0 / 0	0 / 0	0 / 0	0 / 101	0 / 976
1	0 / 0	0 / 0	0 / 0	0 / 34	0 / 608
2	0 / 0	0 / 0	0 / 0	0 / 12	0 / 455

From the tables above, we can observe that as the value of λ increases, the number of zero estimates of b increases but that of λ decreases. While as the value of b increases, the numbers of zero estimates of b and λ both increases.

The following figures are the illustrations of the parameter estimates of the samples generated from a parameter setting a=1, b=1, λ=0.1 and sample size n=50

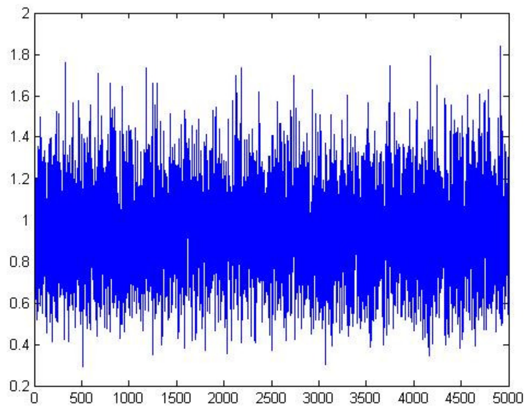


Figure 4.3 Parameter Estimates of *a*

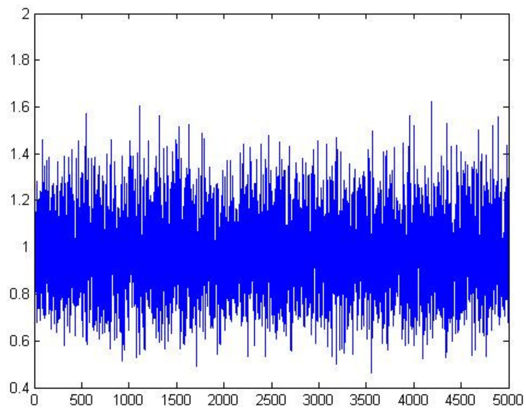


Figure 4.4 Parameter Estimates of *b*

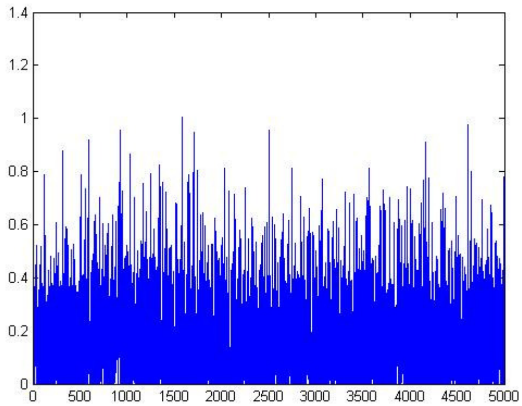


Figure 4.5 Parameter Estimates of λ

4.8 Summary

In this chapter we examine the log-likelihood function of a class of distributions, which includes many useful models for lifetime data analysis, such as the Weibull distribution and the modified Weibull. A simple technique is proposed to simplify the computation of the elements of the Observed Fisher Information matrix. In addition, the form of the class of distributions makes it possible to decrease the number of variables in the log-likelihood function.

Using the techniques developed, maximum likelihood estimation of the model parameters of the modified Weibull distribution with progressively type-2 censored samples is studied. The property of the log-likelihood function is investigated by introducing the simple transformation to decrease the dimensionality of the parameter vector while keeping the analysis tenable.

Existence and uniqueness of the MLE of the model parameters are proved. However, we found that the unique MLE that maximize the log-likelihood function may not be appropriate estimates of the parameters of the modified Weibull distribution to model lifetime data, and non-negative constraints have to be imposed on the parameters. Several examples are presented to illustrate the uniqueness property of the MLE.

Chapter 5. MCMC Estimation of Modified Weibull Parameters

ML estimation of the model parameters of statistical distributions is straightforward, and under mild regularity conditions the estimates are asymptotically unbiased. However, when the sample size is small, the MLE for the Weibull parameters are usually quite biased (Watkins, 1996; Montanari *et al*., 1997). Based on the pivotal property of the Weibull parameters discovered by Thoman *et al*. (1969), several methods have been proposed to reduce the bias (Ross, 1996; Montanari *et al*., 1997; Hiross, 1999; Yang and Lin, 2007).

For the modified Weibull distribution with 3 model parameters, such techniques are not readily available and even if they exist it would involve too many factors to make the implementation complicated. As an alternative to ML estimation, we consider the Bayesian method employing the Markov chain Monte Carlo (MCMC) techniques and compare its estimation accuracy and dispersion against MLE.

As a family of the powerful tools for sampling from multivariate statistical distributions, Bayesian methods implemented through MCMC have been developed and applied to estimate the model parameters based on a complete or censored sample. Merits of MCMC estimation for the Weibull parameters have been discussed in Green *et al*. (1994) and Pang *et al*. (2005, 2007). Advantages of

MCMC estimation over other estimation methods include small bias when the sample size is not large, ease of constructing exact probability intervals, and convenience of incorporating prior/expert information into consideration, etc.

Another advantage of MCMC estimation is discussed in Green *et al*. (1994), who showed that the MLE of the location parameter of the 3-parameter Weibull distribution has a large chance to be negative, hence fails to meet the condition the true underlying distribution, but MCMC estimate can always be positive with a proper choice of the prior distribution. For the modified Weibull distribution, in chapter 4 we have shown that direct maximization of the log-likelihood function could yield negative estimates of the parameter λ, which are not suitable for modeling life time data.

In this chapter, we study the Gibbs sampler for the parameters of the modified Weibull distribution based on a Bayesian framework and make a comparison between these estimates and the MLE regarding their bias and variability.

The content of the current chapter is published in Jiang *et al*. (2008)[(1)].

5.1 Bayesian Model

For the modified Weibull distribution, the Bayesian model is constructed by specifying a prior distribution for a, b and λ, and then multiplying with the likelihood function to obtain the posterior distribution function. Given a complete sample of data $t = (t_1, t_2, \cdots, t_n)$, the likelihood function is

$$l(a,b,\lambda;t) = a^n \prod (b + \lambda t_i)\left(\prod t_i\right)^{b-1} \exp\left\{\lambda \sum t_i - a \sum t_i^b e^{\lambda t_i}\right\} \tag{5.1}$$

Denote the prior of a, b and λ as $p(a,b,\lambda)$. The joint posterior distribution is

$$p(a,b,\lambda \mid t) \propto l(a,b,\lambda;t)p(a,b,\lambda) \tag{5.2}$$

Here the prior distribution is given in advance, usually based on prior information of the parameters, which is from historical data, previous experiences and expert suggestions, but sometimes the choice of prior is just for mathematical convenience. For the current model, since there are no constraints for the parameters except for non-negativity, and we have no reason to prefer one value over another for each of the parameters, it is convenient to assume independent generalized uniform distributions on the positive supports for the three parameters, i.e. $p(a,b,\lambda) = p(a)p(b)p(\lambda)$, $p(a) = p(b) = p(\lambda) \propto 1$, $a > 0, b, \lambda \geq 0$. In such case, the joint posterior PDF is proportional to the likelihood function

$$p(a,b,\lambda \mid t) \propto a^n \prod (b + \lambda t_i)\left(\prod t_i\right)^{b-1} \exp\left\{\lambda \sum t_i - a \sum t_i^b e^{\lambda t_i}\right\} \tag{5.3}$$

5.2 Gibbs Sampler Parameter Estimation

5.2.1 Steps of Gibbs Sampling

As introduced in Gelfand and Smith (1990) and Ibrahim *et al*. (2001), the steps of using Gibbs sampler to draw samples of the parameters from the Bayesian posterior distribution are sequentially sampling from the full conditional distribution of each parameter based on the given samples of other parameters. For the modified Weibull Bayesian model (5.3), letting $p_a(a \mid b, \lambda, t)$, $p_b(b \mid a, \lambda, t)$ and $p_\lambda(\lambda \mid a, b, t)$ denote the full conditional CDF of a, b and λ, the steps can be described as follows:

(0) Arbitrarily choose an starting point (a_0, b_0, λ_0), and set k=0;

(1) Generate $(a_{k+1}, b_{k+1}, \lambda_{k+1})$ as follows:

   a. Sample a_{k+1} from $p_a(a \mid b_k, \lambda_k, t)$;

   b. Sample b_{k+1} from $p_b(b \mid a_{k+1}, \lambda_k, t)$;

   c. Sample λ_{k+1} from $p_\lambda(\lambda \mid a_{k+1}, b_{k+1}, t)$.

(2) Set k=k+1, and then go to step (1). Continue the iterative steps until a predetermined number of runs is reached.

Under mild regularity conditions, Geman and Geman (1984) showed that

(a_k, b_k, λ_k) converge to the true values of (a, b, λ) in distribution as *k* approaches infinity. Therefore, we can make inferences about the parameters with the Markov chain obtained, such as estimation by taking the average of the corresponding values in the chain. However, in normal conditions, the successive observations are not independent in a Markov chain. If an independent identically distributed (iid) sample is needed, suitably spaced observations may be required, say every 40th (Green *et al*., 1994). In addition, a suitable burn-in is needed to diminish the influence of the starting values of the parameters. To check the convergence of the Markov chain, in most cases where the computational cost is not too high to afford, it is preferred to run the Gibbs sampler several times with different starting points and check whether after a sufficiently long run the different Markov chains will converge to the same stationary distribution (Gelfand and Smith, 1990).

5.2.2 Adaptive Rejection Sampling

An important step in Gibbs sampling is to sample from the full conditional distributions. Since in most cases the distributions are so complicated that it is difficult or impossible for direct sampling, rejection sampling techniques are required. Gilks and Wild (1992) introduced an adaptive rejection sampling scheme for Gibbs sampling when the target distribution is complicated and evaluation of the full conditional PDF is computationally expensive. This sampling method is an extension of the rejection sampling and is suitable for any log-concave PDF, i.e. $\frac{\partial^2 \log f(\theta)}{\partial \theta^2} < 0$. The advantage of the adaptive rejection sampling scheme is that it is adaptive: the envelope function converges to the target conditional PDF as sampling proceeds, and the reconstructions of the envelope function and the squeezing function only need negligible computational cost, thus it is very efficient compared to direct sampling or traditional rejection sampling. Denote $r(\theta) = \log(f(\theta))$. The steps of adaptive rejection sampling can be described as

0. Determine $T_k = (\theta_1, \theta_2, \cdots, \theta_k)$ as the set of three or four initial abscissae for $r(\theta)$, where $r'(\theta_1) > 0$ and $r'(\theta_k) < 0$, k=3 or 4. Define the envelope function of $r(\theta)$ as $u_k(\theta)$, which is a piecewise linear function, with each linear part being the tangent of $r(\theta)$ at the abscissa;

1. Define the envelope function of $r(\theta)$ as $u_k(\theta)$, which is a piecewise linear function, with each linear part being the tangent of $h(\theta)$ at the abscissa in T_k;

2. Define $s_k(\theta)$ as PDF being proportional to $\exp\{u_k(\theta)\}$

$$s_k(\theta) = \exp\{u_k(\theta)\} \Big/ \int \exp\{u_k(\theta')d\theta'\} \tag{5.4}$$

3. Sample a value θ^* from $s_k(\theta)$ and a value w independently from uniform(0,1) distribution. Perform the following rejection test:

   If $w \le \exp\{h(\theta^*) - u_k(\theta^*)\}$, then accept θ^*;

   otherwise reject θ^*.

4. If θ^* is accepted in step 3, θ^* is the observation wanted and then the sampling process is stopped. Otherwise, include θ^* in T_k to form T_{k+1}, rearrange the elements of T_{k+1} in ascending order, then let $k = k + 1$, go back to step 1 and run through all the steps left.

5.2.3 Convergence Diagnostics

When using the Gibbs sampler method to estimate the model parameters, an important practical issue has to be considered, i.e. convergence diagnostics of the Markov chain. Convergence of the Markov chain ensures that the distribution estimated is a proper approximate of the target distribution.

Research papers on this topic are of vast volume, e.g. Baftery and Banfield (1991), Gelman and Rubin (1992), Casella and George (1992), Roberts and Smith (1994), Zellner and Min (1995), Cowles and Carlin (1996), Belisle (1998), etc.

Among the various methods, the one proposed by Gelman and Rubin (1992) is the most popular. The method involves two steps. The first step is to generate *m* sets of starting points of the Gibbs sampler. The second step is to simulate Markov chains for each of the starting points, for the desired number of iterations, say 2*n*. Convergence is monitored by estimating the following shrink factor,

$$\sqrt{\hat{R}} = \sqrt{\left(\frac{n-1}{n} + \frac{m+1}{mn}\frac{B}{W}\right)\frac{df}{df-2}}$$

where *B* is the variance between the means of the *m* Markov chains, *W* is the average of the *m* within-chain variances, and *df* is the degree of freedom of the Student's *t* distribution, which is approximated by the last *n* observations of the first Markov chain. The Gibbs sampler is deemed to converge when this shrink factor is close to 1.

5.2.4 Gibbs Estimation of Parameters of the Modified Weibull

Given a complete sample of lifetime data $t = \{t_1, t_2, \cdots, t_n\}$, the full conditional PDF of *a* is

$$\begin{aligned} p_a(a \mid b, \lambda, t) &= \frac{p(a, b, \lambda \mid t)}{p(b, \lambda \mid t)} \\ &\propto p(a, b, \lambda \mid t) \\ &\propto L(a, b, \lambda; t) \end{aligned} \tag{5.5}$$

The first proportionality follows because $p(b, \lambda \mid t)$ is the joint posterior of *b* and λ, so no term involving *a* exists in $p(b, \lambda \mid t)$, and the second proportionality is based on the fact that the prior is generalized uniform distributed. In other words, (5.5) means

$$p_a(a \mid b, \lambda, t) \propto a^n \exp\left\{-a \sum t_i^b e^{\lambda t_i}\right\} \tag{5.6}$$

Similarly, the full conditional PDF of b and λ are respectively

$$p_b(b \mid a, \lambda, t) \propto \prod (b + \lambda t_i)\left(\prod t_i\right)^{b-1} \exp\left\{-a \sum t_i^b e^{\lambda t_i}\right\} \tag{5.7}$$

$$p_\lambda(\lambda \mid a, b, t) \propto \prod (b + \lambda t_i) \exp\left\{\lambda \sum t_i - a \sum t_i^b e^{\lambda t_i}\right\} \tag{5.8}$$

It is easy to see that the full conditional PDF of a is a Gamma distribution with the scale parameter $\sum t_i^b e^{\lambda t_i}$ and shape parameter $n+1$. Therefore it is convenient to generate an observation of a from (5.6).

No similar simple distributions are available for b and λ. Though, we can apply the adaptive rejection sampling technique to draw observations from $p_b(b \mid a, \lambda, t)$ and $p_\lambda(\lambda \mid a, b, t)$. In the beginning, the usage of the technique has to be validated, i.e. the two PDF are log-concave.

Theorem 5.1 The full conditional PDF $p_b(b \mid a, \lambda, t)$ and $p_\lambda(\lambda \mid a, b, t)$ are both log-concave.

Proof: $$\frac{\partial \log\left(p_b\left(b \mid a, \lambda, t\right)\right)}{\partial b} = \sum \frac{1}{b + \lambda t_i} + \sum \log\left(t_i\right) - a \sum t_i^b e^{\lambda t_i} \log t_i$$

$$\frac{\partial^2 \log\left(p_b\left(b \mid a, \lambda, t\right)\right)}{\partial b^2} = -\sum \left(\frac{1}{b + \lambda t_i} \right)^2 - a \sum t_i^b e^{\lambda t_i} \left(\log t_i\right)^2 < 0$$

Similar log-concavity property holds for $p_\lambda\left(\lambda \mid a, b, t\right)$

$$\frac{\partial\left(\log\left(p_\lambda\left(\lambda \mid a, b, t\right)\right)\right)}{\partial \lambda} = \sum \frac{t_i}{b + \lambda t_i} + \sum t_i - a \sum t_i^{b+1} e^{\lambda t_i}$$

$$\frac{\partial^2 \left(\log\left(p_\lambda\left(\lambda \mid a, b, t\right)\right)\right)}{\partial \lambda^2} = -\sum \left(\frac{t_i}{b + \lambda t_i} \right)^2 - a \sum t_i^{b+2} e^{\lambda t_i} < 0 .$$ Q.E.D.

Hence we can use the highly efficient adaptive rejection technique presented above to generate random observations from $p_b\left(b \mid a, \lambda, t\right)$ and $p_\lambda\left(\lambda \mid a, b, t\right)$.

We run the Gibbs sampler in the procedure presented in section 5.2.1 and with the adaptive rejection sampling technique presented in section 5.2.2. With any arbitrary starting values of the parameters, we find that very quickly the Markov chain converges to a steady state. Therefore, discarding the first few observations as burn-in, we can use the remaining observations in the Markov chain to calculate the estimates and probability intervals of the parameters.

5.3 Illustrative Example

In this section, we present an example to illustrate the estimation procedures discussed in this chapter. The lifetime data are from Aarset (1987).

In our study, based on the assumption that these data are from the modified Weibull distribution, we run the Gibbs sampler to generate 3 Markov chains at the length of 30,000 with different starting points of parameters. Doing convergence diagnostics following the Gelman and Rubin (1992) procedure, we find that the Markov chains converge together to a stationary process after approximately 2000 observations. Therefore, burn-in of 5000 observations is more than enough to erase the effect of starting point. For one of the Markov chains, discarding the first 5000 and taking every 10th as iid observations, this step serving for the purpose of diminishing the autocorrelation, we can plot the empirical distributions of the model parameters and thus give their estimates and probability intervals.

The Gibbs estimates of the parameters are: $(\hat{a}, \hat{b}, \hat{\lambda}) = (0.0604, 0.3493, 0.0229)$.

A $100(1-\alpha)\%$ probability interval for any parameter may be estimated by taking the $100\alpha/2$ th and $100(1-\alpha/2)$ th percentiles of the generated sample. Table 5.1 lists the 90% and 95% probability intervals for the three parameters.

Table 5.1 Gibbs Estimates and Two-Sided 90% & 95% Probability Intervals for *a*, *b*, and *λ*

Parameter	Estimate	90% P.I.	95% P.I.
a	0.0604	[0.0332, 0.1235]	[0.0287, 0.1413]
b	0.3493	[0.1925, 0.5210]	[0.1643, 0.5560]
λ	0.0229	[0.0154, 0.0307]	[0.0139, 0.0321]

As a reference, the maximum likelihood estimates and confidence intervals based on a progressively Type-2 right censoring scheme for the same data are in table 5.2 (see Ng (2005)).

Table 5.2 MLE and Two-Sided 90% & 95% Confidence Intervals for *a*, *b*, and *λ*

Parameter	Estimate	90% C.I.	95% C.I.
a	0.0714	[0.0354, 0.1444]	[0.0309, 0.1652]
b	0.398	[0.2419, 0.6564]	[0.2198, 0.7222]
λ	0.01702	[0.0084, 0.0256]	[0.0068, 0.0273]

From table 5.1 and table 5.2, we can see that point estimates of the parameters obtained in both methods are close to each other. Regarding the interval estimates, contrary to our intuition, the length of the Gibbs sampler probability intervals is smaller than that of the MLE confidence intervals, for each of the three parameters on both 90% and 95% significance levels.

In the following section, a simulation study is done to look into the biasness and dispersion, and hence the probability interval properties, of the estimators.

5.4 Simulation Study

A Monte Carlo simulation study is conducted to compare the performance of Gibbs estimators (MCMCE) and MLE of the model parameters of the modified Weibull distribution. For each of the following sets of parameters, we simulated 1000 sets of data with sample sizes n=20, 50, 100 and 200, respectively, and based on each data set we computed MLE and MCMCE for the model parameters. The priors of the parameters are generalized uniform distributions. In order to obtain MCMCE, we run the Gibbs sampler to construct Markov chains at the length of 500.

1) a=0.5, b=1.0, λ=0.1;
2) a=1.0, b=0.5, λ=0.1;
3) a=0.5, b=1.0, λ=0.2;
4) a=1.0, b=0.5, λ=0.2.

As stated in section 5.3, we find that the starting values of the parameters do not affect the convergence of the Markov chain. Therefore, to minimize the influence of the choice of starting points and reduce the running time of the simulation routine, we choose the true values of the parameters as the starting values and do not consider burn-in.

We take mean as the estimates of the parameters a, b, and λ, since the empirical posterior distributions of them are all fairly symmetric.

The following tables list the results of the simulation study. Denote $\left(\hat{a}, \hat{b}, \hat{\lambda}\right)$ as MLE and $\left(\widehat{a}, \widehat{b}, \widehat{\lambda}\right)$ as MCMCE. Bias and MSE are calculated for each of the parameter sets and sample sizes.

Table 5.3 Comparison of MLE and MCMCE for (a, b, λ)=(1, 0.5, 0.1)

n		Bias a	MSE a	Bias b	MSE b	Bias λ	MSE λ
20	MLE $\hat{\theta}$	-0.0641	0.1754	-0.0978	0.2735	0.1262	0.2063
	MCMCE $\widehat{\theta}$	0.0276	0.1298	0.0484	0.2108	0.0039	0.0123
50	MLE $\hat{\theta}$	-0.0331	0.1129	-0.0478	0.1839	0.0567	0.1216
	MCMCE $\widehat{\theta}$	0.0117	0.0888	0.0215	0.13550	0.0031	0.0182
100	MLE $\hat{\theta}$	-0.0138	0.0799	-0.0142	0.1402	0.0235	0.0785
	MCMCE $\widehat{\theta}$	0.0072	0.0651	0.0149	0.1008	0.0011	0.0252
200	MLE $\hat{\theta}$	-0.0089	0.0582	-0.0035	0.1077	0.0122	0.0582
	MCMCE $\widehat{\theta}$	0.0013	0.0482	0.0117	0.0816	0.0014	0.0306

Table 5.4 Comparison of MLE and MCMCE for (a, b, λ)=(0.5, 1, 0.1)

n		Bias a	MSE a	Bias b	MSE b	Bias λ	MSE λ
20	MLE $\hat{\theta}$	-0.0739	0.3060	-0.0088	0.1272	0.0829	0.1732
	MCMCE $\widehat{\theta}$	0.0592	0.2405	0.0458	0.1166	0.0028	0.0201
50	MLE $\hat{\theta}$	-0.0109	0.2000	0.0047	0.0824	0.0222	0.0763
	MCMCE $\widehat{\theta}$	0.0407	0.1734	0.0255	0.0758	-0.0008	0.0280
100	MLE $\hat{\theta}$	-0.0098	0.1413	0.0023	0.0614	0.0130	0.0520
	MCMCE $\widehat{\theta}$	0.0174	0.1265	0.0134	0.0569	0.0021	0.0320
200	MLE $\hat{\theta}$	-0.0046	0.0996	0.0008	0.0421	0.0063	0.0330
	MCMCE $\widehat{\theta}$	0.0069	0.0962	0.0056	0.0413	0.0027	0.0283

Table 5.5 Comparison of MLE and MCMCE for (a, b, λ)=(0.5, 1, 0.2)

n		Bias a	MSE a	Bias b	MSE b	Bias λ	MSE λ
20	MLE $\hat{\theta}$	-0.0543	0.1881	-0.0503	0.3159	0.1371	0.2751
	MCMCE $\widehat{\theta}$	0.0321	0.1263	0.0787	0.2360	0.0018	0.0331
50	MLE $\hat{\theta}$	-0.0191	0.1268	-0.0151	0.2094	0.0458	0.1528
	MCMCE $\widehat{\theta}$	0.0203	0.0951	0.0401	0.1541	-0.0043	0.0489
100	MLE $\hat{\theta}$	-0.0025	0.0944	-0.0016	0.1659	0.0161	0.1099
	MCMCE $\widehat{\theta}$	0.0178	0.0766	0.0248	0.1285	-0.0063	0.0605
200	MLE $\hat{\theta}$	-0.0026	0.0687	0.0001	0.1250	0.0070	0.0815
	MCMCE $\widehat{\theta}$	0.0094	0.0614	0.0161	0.1097	-0.0054	0.0632

Table 5.6 Comparison of MLE and MCMCE for (a, b, λ)=(1, 0.5, 0.2)

n		Bias a	MSE a	Bias b	MSE b	Bias λ	MSE λ
20	MLE $\hat{\theta}$	-0.0284	0.3379	0.0105	0.1496	0.0880	0.2188
	MCMCE $\widehat{\theta}$	0.0912	0.2514	0.0563	0.1308	0.0005	0.0483
50	MLE $\hat{\theta}$	-0.0115	0.2026	0.0047	0.0995	0.0316	0.1120
	MCMCE $\widehat{\theta}$	0.0516	0.1767	0.0288	0.0931	-0.0004	0.0602
100	MLE $\hat{\theta}$	-0.0011	0.1614	0.0032	0.0675	0.0156	0.0805
	MCMCE $\widehat{\theta}$	0.0303	0.1517	0.0153	0.0654	0.0018	0.0614
200	MLE $\hat{\theta}$	-0.0065	0.1126	0.0013	0.0481	0.0085	0.0533
	MCMCE $\widehat{\theta}$	0.0085	0.1121	0.0071	0.0485	0.0033	0.0501

For the comparison of the estimates, we observe the following:

- For estimation of a, though for some cases the bias of MLE is smaller than MCMCE, MCMCE has overwhelming advantage over MLE in the index of MSE. Therefore, MCMCE is more stable than MLE, despite when the sample size is large (say, larger than 100), MLE is less biased than MCMCE. In general, MCMCE is better than MLE in estimating a. Another interesting observation is that MLE consistently underestimates a (bias is negative), while MCMC overestimates (bias is positive).

- For estimation of b, similar to the cases of estimating a, MCMCE is better than MLE in general.

- For estimation of λ, MCMCE is always better than MLE.

- When the sample size is small (say, less than 100), MCMCE behaves better than MLE in both indexes, bias and MSE. The advantage of MCMCE over MLE is especially remarkable in the estimation of parameter λ.

- We can easily obtain the probability intervals for the parameters through MCMCE from the empirical distributions of the parameters. Contrastingly, the construction of confidence intervals involved in MLE needs the local estimate of the asymptotic variance-covariance matrix of MLE, or the Observed Fisher Information matrix. Therefore, the calculation of probability intervals for MCMCE is easier and may be more accurate.

Based on the simulation results, we suggest the use of MCMCE instead of MLE for parameter estimation when the sample size is not very large (say, less than 100). When the sample size is large, MCMCE is still more stable (smaller variability) than MLE, but the bias is larger.

When considering computational cost, MCMCE has no advantage over MLE, since the generalization of the Markov chain usually takes far more time than the optimization procedure needed in maximizing the log-likelihood function. To make a good choice between MCMCE and MLE, decision makers are suggested to take into account the pros and cons of the both methods.

5.5 Summary

Gibbs sampler, as one of the MCMC algorithms, is introduced to estimate the parameters of the modified Weibull distribution based on a Bayesian framework. The adaptive rejection sampling technique is used to sample from the full conditional distributions of the parameters. Gibbs estimation is compared with ML estimation for several different parameter sets and sample sizes, and it is found that the former outperforms the latter for small samples and has the advantage of being easy to construct probability intervals.

Chapter 6. Statistical Characterization and Parameter Estimation of Odd Weibull

Based on the idea of evaluating the distribution of the "odds of death" of a lifetime variable, the odd Weibull distribution proposed by Cooray (2006) has recently been shown to be useful for testing goodness-of-fit of the Weibull and inverse Weibull. The model is also very versatile in modelling lifetime data because its failure rate function can be increasing, decreasing, constant, bathtub shaped and unimodal. In this chapter, a detailed parametric characterization of the statistical properties of this distribution is carried out. Shapes of WPP with different model parameters are presented and the graphic parameter estimation steps are iterated. Burn-in and useful period related issues of the bathtub shaped failure rate curve are discussed. An application example is shown to illustrate the parameter estimation procedure and the superior fit of the model for some real data to the other 3-parameter generalizations of the Weibull distribution.

CDF of this distribution is

$$F(t) = 1 - \left(1 + \left(e^{(t/\theta)^{\alpha}} - 1\right)^{\beta}\right)^{-1}, \; 0 < t < \infty \tag{6.1}$$

with $\theta > 0$ the scale parameter and $\alpha\beta > 0$ the shape parameters. Note that when $\beta = 1$, $F(t)$ is CDF of Weibull, and when $\beta = -1$, it is CDF of inverse Weibull.

The quantile function can be shown to be

$$Q(u) = \theta\left(\ln\left(1 + \left(u/(1-u)\right)^{1/\beta}\right)\right)^{1/\alpha} \tag{6.2}$$

Starting from a Weibull distribution $F_W(t)$ (when the shape parameters α and β are positive), the odd Weibull CDF can be expressed as

$$F(t) = \frac{F_W^{\beta}(t)}{F_W^{\beta}(t) + (1 - F_W(t))^{\beta}} \tag{6.3}$$

while if starting from an Inverse Weibull $F_I(t)$ (when the shape parameters α and β are negative), the odd Weibull CDF is

$$F(t) = \frac{F_I^{-\beta}(t)}{F_I^{-\beta}(t) + (1 - F_I(t))^{-\beta}} \tag{6.4}$$

Taking derivative of the distribution function with respect to t, PDF and then FRF can be obtained respectively

$$f(t) = \left(\frac{\alpha\beta}{t}\right)\left(\frac{t}{\theta}\right)^{\alpha} e^{(t/\theta)^{\alpha}}\left(e^{(t/\theta)^{\alpha}} - 1\right)^{\beta-1}\left(1 + \left(e^{(t/\theta)^{\alpha}} - 1\right)^{\beta}\right)^{-2} \tag{6.5}$$

$$h(t) = \left(\frac{\alpha\beta}{t}\right)\left(\frac{t}{\theta}\right)^{\alpha} e^{(t/\theta)^{\alpha}}\left(e^{(t/\theta)^{\alpha}} - 1\right)^{\beta-1}\left(1 + \left(e^{(t/\theta)^{\alpha}} - 1\right)^{\beta}\right)^{-1} \tag{6.6}$$

As the name stands, the odd Weibull distribution originates from the idea of evaluating the “odds of death” of a Weibull or Inverse Weibull random variable

(Cooray, 2006). The logit function, i.e. the logarithm of the odds, of CDF of the odd Weibull distribution can be written as the product of the logit function of the corresponding Weibull or Inverse CDF and the shape parameter β or $-\beta$

$$\text{logit}(F(t)) = \log\frac{F(t)}{1-F(t)} = \beta\log\frac{F_W(t)}{1-F_W(t)} = \beta\text{logit}(F_W(t))$$

or

$$\text{logit}(F(t)) = \log\frac{F(t)}{1-F(t)} = -\beta\log\frac{F_I(t)}{1-F_I(t)} = -\beta\text{logit}(F_I(t))$$

This relationship between the odd Weibull and the Weibull/Inverse Weibull distribution may be useful in logistic regression analysis of some lifetime data.

However, up to now little research has been done to investigate the behaviors of this new model. It is often helpful to study the statistical properties and parameter estimation of a distribution before it is widely used to model real data. Therefore, the purpose of this chapter is to provide a systematic characterization of the basics of the odd Weibull model.

The content of the current chapter is published in Jiang *et al*. (2008)[(2)].

6.1 Statistical Characteristics

6.1.1 Shape of Failure Rate Function

The shape of FRF is important for modeling lifetime data. Compared to the Weibull distribution which has monotonic failure rate, the odd Weibull distribution is able to exhibit monotonic, bathtub-shaped, unimodal and some other failure rate shapes.

As from (6.6), the form of FRF is complicated, so analytic methods such as studying the derivative can be useful to get information of the shape of FRF.

Taking logarithm on $h(t)$ and differentiating the function with respect to $z = (t/\theta)^{\alpha}$, we obtain

$$\log\left(h\left(t_{(z)}\right)\right) = \log\left(\frac{\alpha\beta}{\theta}\right) + \frac{\alpha - 1}{\alpha}\log(z) + z + (\beta - 1)\log\left(e^{z} - 1\right) - \log\left(1 + \left(e^{z} - 1\right)^{\beta}\right)$$

We can learn the monotonicity property of $h(t)$ via examining the sign of the derivative

$$\frac{\partial \log h}{\partial z} = \frac{\alpha - 1}{\alpha}\frac{1}{z} + \frac{\beta e^{z} - 1 - \left(e^{z} - 1\right)^{\beta}}{\left(e^{z} - 1\right)\left(1 + \left(e^{z} - 1\right)^{\beta}\right)} \qquad (6.7)$$

As z increase from 0 to ∞ (infinity), we have

1. $1/z$ decreases from ∞ to 0. Therefore, if $\alpha > 1$, $\frac{\alpha - 1}{\alpha}\frac{1}{z}$ decreases from ∞ to 0; if $0 < \alpha < 1$, $\frac{\alpha - 1}{\alpha}\frac{1}{z}$ increases from $-\infty$ to 0; if $\alpha < 0$, $\frac{\alpha - 1}{\alpha}\frac{1}{z}$ decreases from ∞ to 0.
2. Denote $g(z) = \frac{\beta e^z - 1 - (e^z - 1)^\beta}{(e^z - 1)(1 + (e^z - 1)^\beta)}$. If $\beta > 1$, $g(z)$ decreases from ∞ to 0; if $0 < \beta < 1$, $g(z)$ increases from $-\infty$ to 0; if $\beta < 0$, $g(z)$ increases from $-\infty$ to 0.

From 1 and 2, it is obvious that when $\alpha > 1$ and $\beta > 1$, it follows $\frac{\partial \log h}{\partial z} > 0$ and then $h(t)$ is monotonically increasing; while when $0 < \alpha < 1$ and $0 < \beta < 1$, it follows $\frac{\partial \log h}{\partial z} < 0$ and then $h(t)$ is monotonically decreasing.

Regarding the shapes of $h(t)$ when α and β take different values, according to Cooray (2006), it is very difficult to do the classification analytically, and boundary lines have to be obtained numerically. The author showed that when $(\alpha < 0, \beta < 0)$, $(\alpha > 1, \alpha\beta > 1)$, $(\alpha < 1, \alpha\beta < 1)$, $(\alpha > 1, \alpha\beta \le 1)$ and $(\alpha < 1, \alpha\beta \ge 1)$, the shapes are unimodal, increasing, decreasing, bathtub and unimodal, respectively. Typical shapes of FRF are exhibited in the following figure.

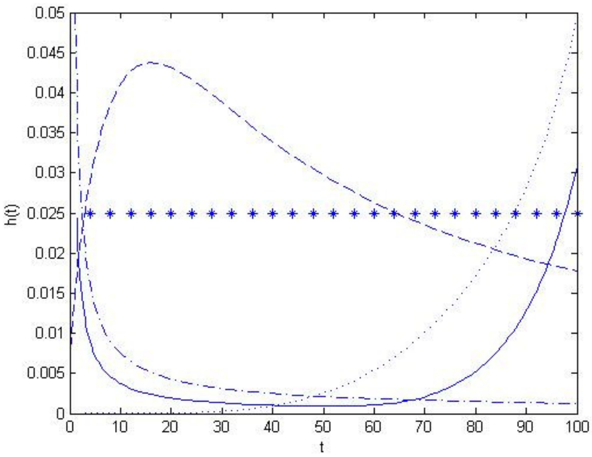
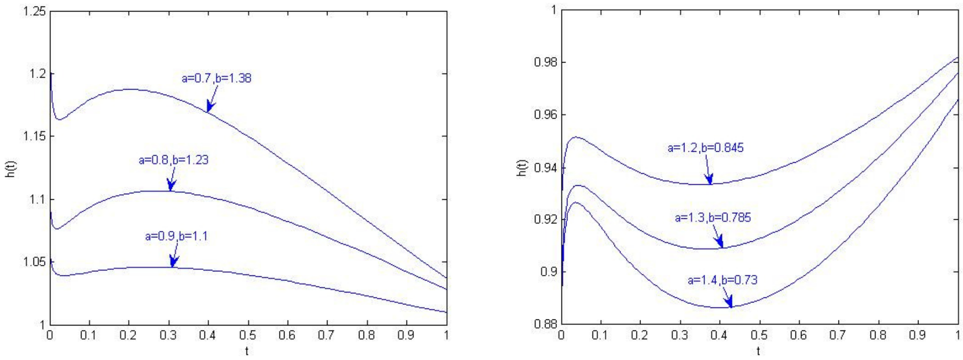


Figure 6.1 Shapes of failure rate function. Unimodal (dashed line), increasing (dotted line), decreasing (dot dashed line), and bathtub shaped (dark line)

It was also indicated that in the regions of $(\alpha > 1, \alpha\beta > 1)$ and $(\alpha < 1, \alpha\beta < 1)$, $h(t)$ may have some other shapes. Numeric analysis shows that the "other shapes" are S and inverse-S, which appear in the two regions when the model shape parameters α and β are near the boundary line $\alpha\beta = 1$. These are the only shapes that have been observed.

The following figures exhibit the shapes of FRF near the boundary line $\alpha\beta = 1$.



(a) α<1, $\alpha\beta$<1 (b) α>1, $\alpha\beta$>1

Figure 6.2 FRF for (α<1, $\alpha\beta$<1) and (α>1, $\alpha\beta$>1) when $\alpha\beta$ is close to 1

6.1.2 Tails of Failure Rate Function

The left and right tails of FRF determine the aging behavior of the model during the infant and elderly periods, so are important for the application of the distribution to lifetime data.

When $\alpha > 0, \beta > 0$, as $t \to 0$, $h(t) \approx \alpha\beta t^{\alpha\beta-1} / \theta^{\alpha\beta}$; as $t \to \infty$, $h(t) \approx \alpha\beta t^{\alpha-1} / \theta^{\alpha}$.

When $\alpha < 0, \beta < 0$, as $t \to 0$, $h(t) \approx \alpha\beta t^{\alpha-1} e^{\beta(t/\theta)^{\alpha}} / \theta^{\alpha}$; as $t \to \infty$, $h(t) \approx \alpha\beta / t$.

Therefore, the pattern of the left tail ($t \to 0$) is classified as follows:

1) $\alpha > 0, \beta > 0, \alpha\beta < 1$, $h(t) \to \infty$, the left tail is unbounded.

2) $\alpha > 0, \beta > 0, \alpha\beta = 1$, $h(t) \to 1/\theta$.

3) $\alpha > 0, \beta > 0, \alpha\beta > 1$, $h(t) \to 0$.

4) $\alpha < 0, \beta < 0$, $h(t) \to 0$, since $\ln\left(t^{\alpha-1}e^{\beta(t/\theta)^\alpha}\right) = (\alpha-1)\ln(t) + \beta(t/\theta)^\alpha \to -\infty$.

The pattern of the right tail ($t \to \infty$) is similarly classified:

1) $0 < \alpha < 1, \beta > 0$, $h(t) \to 0$.

2) $\alpha = 1, \beta > 0$, $h(t) \to \beta/\theta^\alpha$

3) $\alpha > 1, \beta > 0$, $h(t) \to \infty$, the right tail is unbounded.

4) $\alpha < 0, \beta < 0$, $h(t) \to 0$.

The interesting case is $\alpha = 1, \beta > 0$. With such parameters, as $t \to \infty$, FRF $h(t)$ approaches a finite horizontal line $y = \beta/\theta^\alpha$. As pointed out in Bain (1978), in many cases the lifetime of units in a regular maintenance program has a FRF that reaches a stable condition after sufficient long time because of proper maintenance. To model such a life behavior, the Weibull distribution is not a good choice, but the odd Weibull distribution could be.

6.1.3 Moments

The *k*-th moment of the random variable from the odd Weibull distribution is given by

$$E\left(T^k\right) = k\int_0^\infty t^{k-1}\overline{F}(t)dt$$

$$= \frac{k\theta^k}{|\alpha|} \int_0^\infty \frac{(\log(1+y))^{k/\alpha-1}}{(1+y)(1+y^\beta)} dy \tag{6.8}$$

The moments cannot be obtained in closed form, so have to be computed numerically.

As shown in (6.8), if α is a positive integer and $\beta = 1$, the α-th moment is

$$E(T^\alpha) = \theta^\alpha \int_0^\infty \frac{1}{(1+y)^2} dy = \theta^\alpha \tag{6.9}$$

6.1.4 Extreme Value Property

Let $T_1, T_2, \cdots, T_n$ be a random sample from the odd Weibull distribution, and let $T_{1:n} \le T_{2:n} \le \cdots \le T_{n:n}$ denote the sample order statistics. Also, let $U_{1:n} \le U_{2:n} \le \cdots \le U_{n:n}$ denote the order statistics from a uniform $[0,1]$ distribution. From the quantile function (6.2), the order statistics $T_{i:n}$ have the form

$$T_{i:n} = \theta\left(\ln\left(1 + (U_{i:n}/(1-U_{i:n}))^{1/\beta}\right)\right)^{1/\alpha} \tag{6.10}$$

It is well known that both $nU_{1:n}$ and $n(1-U_{n:n})$ converge in distribution to the standard exponential random variable Z. We have the following:

Theorem 6.1. If $\alpha > 0, \beta > 0$, as $n \to \infty$

$$n^{1/\alpha\beta} T_{1:n} \xrightarrow{D} \theta Z^{1/\alpha\beta} \tag{6.11}$$

$$(\ln n)^{1-1/\alpha} T_{n:n} - \frac{\theta \ln n}{\beta^{1/\alpha}} \xrightarrow{D} -\frac{\theta \ln Z}{\alpha\beta^{1/\alpha}} \tag{6.12}$$

Proof: As from equation (6.10)

$$T_{1:n} = \theta\left(\ln\left(1 + \left(U_{1:n}/(1-U_{1:n})\right)^{1/\beta}\right)\right)^{1/\alpha}$$

Expanding $u/(1-u)$ at $u = 0$, we see that

$$u/(1-u) = u + O(u^2)$$

which implies that

$$n^{1/\alpha\beta}\theta\left(\ln\left(1 + \left(U_{1:n}/(1-U_{1:n})\right)^{1/\beta}\right)\right)^{1/\alpha} = \theta(nU_{1:n})^{1/\alpha\beta} + O\left(\left(nU_{1:n}^2\right)^{1/\alpha\beta}\right)$$

Therefore, as $n \to \infty$, $n^{1/\alpha\beta} T_{1:n} \approx \theta(nU_{1:n})^{1/\alpha\beta} \xrightarrow{D} \theta Z^{1/\alpha\beta}$, and the asymptotic distribution of $n^{1/\alpha\beta} T_{1:n}$ is Weibull.

Again, from equation (6.10)

$$T_{n:n} = \theta\left(\ln\left(1 + \left(U_{n:n}/(1-U_{n:n})\right)^{1/\beta}\right)\right)^{1/\alpha}$$

Expanding $\ln\left(1 + (u/(1-u))^{1/\beta}\right)$ at $u = 1$, we see that

$$\ln\left(1 + (u/(1-u))^{1/\beta}\right) = -1/\beta \ln(1-u) + O(1-u)$$

Let $y = \ln(n(1-U_{n:n}))$, we have

$$\frac{(\ln n)^{1/\alpha} - (\ln n - y)^{1/\alpha}}{(\ln n)^{1/\alpha - 1} / \alpha} \to y \text{ as } n \to \infty$$

Therefore, as $n \to \infty$, $(\ln n)^{1-1/\alpha} T_{n:n} - \frac{\theta \ln n}{\beta^{1/\alpha}} \xrightarrow{D} -\frac{\theta \ln Z}{\alpha \beta^{1/\alpha}}$. Because $\ln Z$ follows an extreme value distribution, $(\ln n)^{1-1/\alpha} T_{n:n} - \frac{\theta \ln n}{\beta^{1/\alpha}}$ is asymptotically extreme value distributed.

Theorem 6.2. If $\alpha < 0, \beta < 0$, as $n \to \infty$

$$(\ln n)^{1-1/\alpha} T_{1:n} - \frac{\theta \ln n}{(-\beta)^{1/\alpha}} \xrightarrow{D} -\frac{\theta \ln Z}{\alpha(-\beta)^{1/\alpha}} \tag{6.13}$$

$$n^{-1/\alpha\beta} T_{n:n} \xrightarrow{D} \theta Z^{-1/\alpha\beta} \tag{6.14}$$

The proof is similar to that of theorem 6.1. Contrary to the case with positive shape parameters, the asymptotic distribution of $(\ln n)^{1-1/\alpha} T_{1:n} - \frac{\theta \ln n}{(-\beta)^{1/\alpha}}$ is extreme value distribution, and $n^{-1/\alpha\beta} T_{n:n}$ is asymptotically inverse Weibull distributed.

6.2 WPP Plotting

In Weibull analysis, WPP is a very convenient and useful tool in model selection and parameter estimation. When a Weibull distribution is fitted to a sample data set, WPP can show whether Weibull fitting is suitable or not, as well as provide estimates of the parameters.

As to Weibull related distributions, WPP parameter estimation is quite crude, because eyeball observation and nonlinear regression based on asymptotic approximates are required. Nevertheless, WPP can also serve as a good tool for model selection and a starting point of more refined analytic parameter estimation methods such as MLE or Bayesian estimation. See Jiang and Murthy (1999) and Zhang and Xie (2007) for example.

For the odd Weibull distribution, WPP transformations yield:

$$x = \ln t,\ y = \ln(-\ln(1 - F(t))) \tag{6.15}$$

Put CDF (6.1) in and we obtain

$$y = \ln\left(\ln\left(1 + \left(e^{\left(e^{x}/\theta\right)^{\alpha}} - 1\right)^{\beta}\right)\right) \tag{6.16}$$

This is a smooth curve and denote it as C.

6.2.1 Weibull Case α>0, β>0

1) For $t \to 0$ ($x \to -\infty$)

In this case, $1+\left(e^{(t/\theta)^{\alpha}}-1\right)^{\beta} \approx 1+(t/\theta)^{\alpha\beta}$. Hence, as $t \to 0$

$$y = \ln\left(\ln\left(1+\left(e^{(t/\theta)^{\alpha}}-1\right)^{\beta}\right)\right) \approx \ln\left((t/\theta)^{\alpha\beta}\right) = \alpha\beta(x-\ln\theta) \qquad (6.17)$$

This is a straight line and let L_1 denote it. L_1 intercepts x-axis at $x_0 = \ln\theta$, and its slope is $\alpha\beta$.

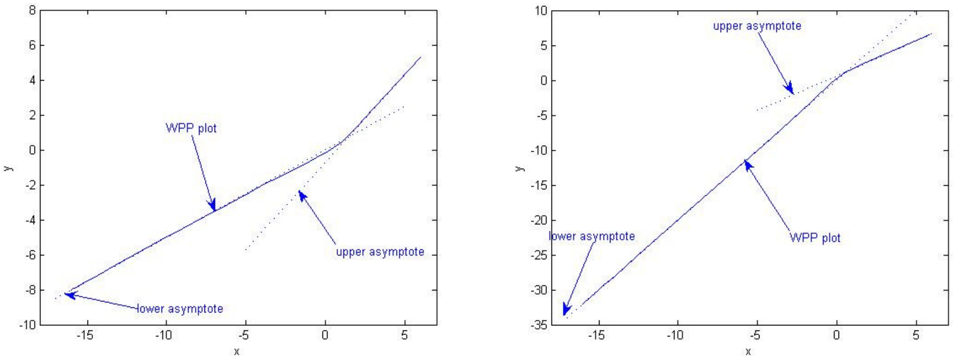
2) For $t \to \infty$ ($x \to \infty$)

In this case, $1+\left(e^{(t/\theta)^{\alpha}}-1\right)^{\beta} \approx e^{\beta(t/\theta)^{\alpha}}$. Hence, as $t \to \infty$

$$y = \ln\left(\ln\left(1+\left(e^{(t/\theta)^{\alpha}}-1\right)^{\beta}\right)\right) \approx \ln\left(\beta(t/\theta)^{\alpha}\right) = \ln\beta + \alpha(x-\ln(\theta)) \qquad (6.18)$$

This is also a straight line and let L_2 denote it. The y-coordinate of the intersection of L_2 with vertical line $L_3: x = \ln\theta$ is $y_0 = \ln\beta$, and the slope of L_2 is α.

The following figure shows the typical WPP plot for the odd Weibull distribution with positive shape parameters. When $0 < \beta < 1$, C is convex; when $\beta > 1$, C is concave.



(a) Typical WPP plot for $0<\beta<1$ (b) Typical WPP plot for $\beta>1$

Figure 6.3 WPP plot of odd Weibull with positive shape parameters

6.2.2 Inverse Weibull Case *α*<0, *β*<0

1) For $t \to 0$ ($x \to -\infty$)

In this case, we have $1+\left(e^{(t/\theta)^\alpha}-1\right)^\beta \approx 1+e^{\beta(t/\theta)^\alpha}$. Therefore, as $t \to 0$

$$y = \ln\left(\ln\left(1+\left(e^{(t/\alpha)^\beta}-1\right)^\theta\right)\right) \approx \ln\left(e^{\beta(t/\theta)^\alpha}\right) = \beta e^{\alpha x}/\theta^\alpha \tag{6.19}$$

2) For $t \to \infty$ ($x \to \infty$)

In this case, $1+\left(e^{(t/\theta)^\alpha}-1\right)^\beta \approx (t/\theta)^{\alpha\beta}$. Therefore, as $t \to \infty$

$$y = \ln\left(\ln\left(1+\left(e^{(t/\theta)^\alpha}-1\right)^\beta\right)\right) \approx \ln(\alpha\beta\ln(t/\theta)) = \ln(\alpha\beta)+\ln(x-\ln(\theta)) \tag{6.20}$$

WPP plotting for the odd Weibull distribution with negative shape parameters does not yield linear asymptotes in either tail. Therefore, it is not able to use WPP to fit an odd Weibull distribution with negative shape parameters to a sample data. Figure 6.4 shows a typical WPP plot in such case. Note here WPP is only able to exhibit a concave shape.

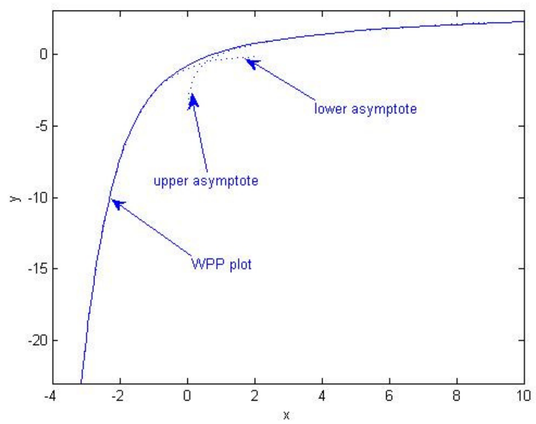


Figure 6.4 WPP plot of odd Weibull with negative shape parameters

6.3 Modeling a Sample Data Set

Normally fitting a Weibull-related distribution to a sample data set and using WPP to estimate the model parameters is composed of two stages. The first stage is plotting the data on a WPP paper; the second stage is estimating the parameters by checking the slope and intersection of the asymptotes.

When using WPP to estimate the parameters of the odd Weibull distribution for a sample data set, because of the different shapes of the plot (see section 6.2), one first needs to determine whether the shape parameters are positive or negative. As mentioned in Cooray (2006), one can draw a TTT plot (see Aarset (1987), Barlow and Campo (1975), Bergman and Klefsjo (1984)) to determine the shape of the failure rate. Then the shape parameters α and β of the distribution can be located into the corresponding region from the shape of FRF according to the classification discussed in section 6.1.1. Moreover, specifically for the odd Weibull distribution, it is easy to check the sign of the shape parameters on a WPP plot. If the WPP curve C is concave, its right tail is asymptotically horizontal, and left tail is asymptotically vertical, then the shape parameters α and β could be negative. Otherwise, α and β must be positive.

6.3.1 Weibull Case *α*>0, *β*>0

Assuming the shape parameters are positive, we can perform the parameter estimation procedure in the following steps:

Stage 1: Plotting WPP for the data

1. Rearrange the data so that t_i, $i = 1, \cdots, n$ is in increasing order;

2. Compute x_i and y_i, $i = 1, \cdots, n$, as follows

$$x = \ln t_i,\ y = \ln(-\ln(R(t_i))) \tag{6.21}$$

where $R(t_i)$ is the empirical survival function at t_i. The computation of $R(t_i)$ depends on the type of data (complete, censored). For details, see Nelson (1982). If the data is complete, a good estimator is $R(t_i) = (n - i + 0.7)/(n + 0.4)$.

3. Plot y_i vs. x_i to generate WPP.

If the fitting plot to the sample data has a shape similar to either one in Figure 6.2, the data can be properly modeled by an odd Weibull distribution with positive shape parameters. An obvious property is that the WPP plot is convex or concave, with linear asymptotes in both tails. Otherwise, the odd Weibull distribution is not an appropriate model, or the shape parameters of the fitted odd Weibull distribution are negative. The latter case will be discussed in the next section.

Stage 2: Parameter estimation

If stage 1 shows that an odd Weibull distribution with positive shape parameters is suitable for modelling the data, then the model parameters can be estimated using the following steps:

4. Fit a straight line L_1 to the left side of the WPP plot. From equation (6.17), the slope of L_1 yields $\tilde{\alpha}\tilde{\beta}$.

5. Fit a straight line L_2 to the right side of the WPP plot. From equation (6.18), the slope of L_2 yields $\tilde{\alpha}$. Using this $\tilde{\alpha}$ and $\tilde{\alpha}\tilde{\beta}$ obtained in step 4, we get $\tilde{\beta}$.

6. Vertically move L_2 by $-\ln\tilde{\beta}$ to generate another line L_3. L_3 is parallel to L_2, and its functional form is $y = \beta(x - \ln\theta)$. The x-axis of the intersection of L_1 and L_3 yields $\ln\tilde{\theta}$, then accordingly $\tilde{\theta}$.

When estimating θ, it is important to ensure that the intersection point lies on the x-axis. To satisfy this condition, some adjustment of L_1 and L_3 (L_2) may be needed. For example, if the intersection is under x-axis, then we do the following

7. Move the line with the larger slope to the right and/or the line with the smaller slope to the left until the intersection point is on the x-axis, whilst ensuring both or either of L_1 and L_2 still fit the tails of the WPP curve C well.

8. If the revised L_1 and/or L_2 do not fit C well, then adjust the slope(s) of the unfitted line(s) and then go back to step 4 or 5 and through 6 and 7 to estimate the parameters.

The above steps may be required to repeat several times until good estimates can be reached.

The graphic parameter estimation approach is able to give accurate estimates to the parameters of the Weibull distribution (e.g. Weibull (1951)). However, to Weibull-related distributions, the graphic approach is generally crude, since it is based on observation and some measures of approximating asymptotes to the real lines are inevitable. Nevertheless, the plotting and estimation are helpful for identifying the intrinsic life mode in the data and doing model selection (e.g. Murthy *et al*. (2004)[(2)]). In addition, the estimates obtained in graphic approach can be used as starting point to obtain more refined estimates using statistical methods such as maximum likelihood.

6.3.2 Inverse Weibull Case $\alpha<0$, $\beta<0$

As discussed above, when failure rate of the data has a unimodal shape or WPP of the data is concave, the shape parameters of the fitted odd Weibull distribution are probable to be negative, but could still be positive. In such case, one can still use the WPP method introduced in section 6.3.1 to estimate the model parameters, and see whether the model is well fitted or not. If not, one has sufficient reason to doubt the assumption of positive shape parameters, and then can use the inverse property introduced by Cooray (2006) to transform the data so that WPP parameter estimation steps are still useful.

Cooray (2006) shows that if a random variable X follows odd Weibull distribution with parameters $(\alpha_0, \beta_0, \theta_0)$, $1/X$ is still odd Weibull distributed, and the distribution parameters $(\alpha_1, \beta_1, \theta_1)$ have the following relationship with $(\alpha_0, \beta_0, \theta_0)$

$$\alpha_1 = -\alpha_0,\ \beta_1 = -\beta_0,\ \theta_1 = 1/\theta_0 \tag{6.22}$$

From this property, if the failure rate or WPP of a sample data set indicates the shape parameters of the odd Weibull distribution are negative, one can invert the data x_1, x_2,…, x_n to $1/x_1$, $1/x_2$,…, $1/x_n$, and then plot WPP for $\{1/x_1, \cdots, 1/x_n\}$ and estimate the model parameters following the steps as iterated in section 6.3.1. If appropriate estimates of $(\alpha_1, \beta_1, \theta_1)$ are obtained, then the original parameters $(\alpha_0, \beta_0, \theta_0)$ can be easily estimated from (6.22). Otherwise, the odd Weibull distribution is not an appropriate model for the data and one should try other models.

6.4 Optimal Burn-In Time and Useful Period

As discussed in the above, when $\alpha > 1$, $\alpha\beta \le 1$, the odd Weibull family is flexible at describing bathtub shaped failure rate data. In this section, some important characteristics of the bathtub curve are discussed.

For a product lifetime exhibiting a bathtub shaped failure rate, an important issue is to determine the optimal burn-in time. A common method is to find the time where the corresponding MRL achieves its maximum (Lai *et al*. (2004), Bebbington *et al*. (2006)). MRL $\mu(t)$ is defined as

$$\mu(t) = \frac{\int_t^\infty (1 - F(t))dt}{1 - F(t)} \tag{6.23}$$

By differentiating $\mu(t)$ with respect to t, the point t^* which maximizes $\mu(t)$ can be found and it is defined in Mi (1995) as a good choice of optimal burn-in time.

Gupta and Akman (1995) proved that for a lifetime distribution, if FRF $h(t)$ is bathtub shaped and $h(0) > 1/\mu$, where μ is the mean time to failure, the corresponding MRL $\mu(t)$ is unimodal with a unique maximum point. For the odd Weibull distribution, when $h(t)$ exhibits a bathtub shape, $h(0) = \infty$, so $\mu(t)$ is unimodal shaped with a unique change point. At the maximum point t^*, $\mu'(t) = 0$, from Muth (1977), there exists a relationship between $\mu(t)$ and $h(t)$

$$\mu'(t)|_{t^*} = \mu(t)h(t)|_{t^*} - 1 = 0 \tag{6.24}$$

For the odd Weibull distribution, when the shape parameters $\alpha > 1$, $\alpha\beta \le 1$, FRF (6.6) is bathtub-shaped, so with the transformation $z = (t/\theta)^{\alpha}$

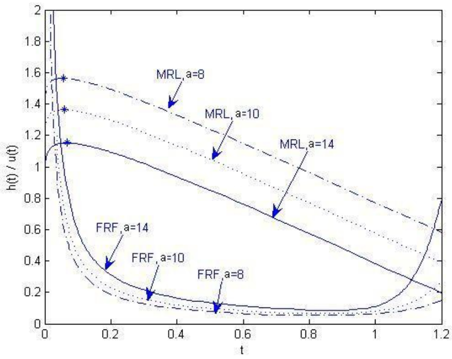
$$\mu'(t) = \alpha\beta z^{\frac{\alpha-1}{\alpha}} e^{z}\left(e^{z}-1\right)^{\beta-1}\int_{z^{1/\alpha}}^{\infty}\left[1+\left(e^{x^{\alpha}}-1\right)^{\beta}\right]^{-1}dx - 1 \tag{6.25}$$

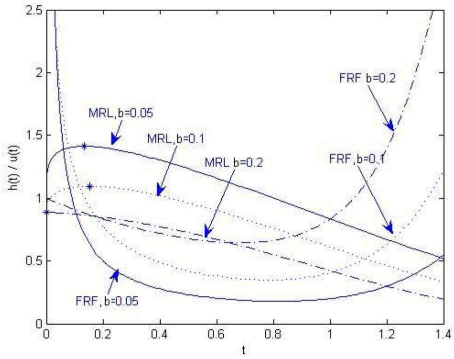
Denote z_1^* as the zero point of (6.25), and then the change point t^* is

$$t^* = \theta\left(z_1^*\right)^{1/\alpha} \tag{6.26}$$

The change point or optimal burn-in time t^* does not have a closed form, but it is unique and can be obtained via numeric methods, such as Newton method.

The following figure shows how $\mu(t)$ and $h(t)$ behave for different parameters, with the maximum point t^* of $\mu(t)$ indicated.





(a) β=0.01, θ=1 (b) α=5, θ=1

Figure 6.5 Typical FRF and MRL curves of odd Weibull

Besides the optimal burn-in time, the length of the useful period or the random risk period of the bathtub curve is important for application. This is because a product with bathtub shaped failure rate is only stable during the useful period, thus it is usually considered the longer this period is the better. Bebbington *et al*. (2006) defines the terms useful period and conservative useful period based on the curvature of $h(t)$ and studies the useful periods of the Additive Weibull distribution. These definitions are intuitively acceptable, but for most bathtub shaped FRF, the curvature is too complicated to deal with. Another definition was given in Xie *et al*. (2004). The authors propose to use the relative difference between the unique change points of FRF $h(t)$ and MRL $\mu(t)$ as an indicator of the length of the useful period. When $h(t)$ exhibits a bathtub shape, Mi (1995) proves the corresponding $\mu(t)$ has a unique change point before $h(t)$, and hence the measure proposed in Xie *et al*. (2004) is well defined. In the current paper, we consider Xie's evaluation of the length of useful period.

The relative difference is defined as

$$d = \frac{b^* - t^*}{b^*} \tag{6.27}$$

where b^* is the unique change point or minimum of FRF, and t^* is the unique change point of MRL.

To find b^*, differentiate $\log(h(t))$ with respect to z, we have

$$\frac{\partial \log(h(t))}{\partial z} = \frac{\alpha - 1}{\alpha}\frac{1}{z} + 1 + (\beta - 1)\frac{e^z}{e^z - 1} - \beta\frac{\left(e^z - 1\right)^{\beta-1} e^z}{1 + \left(e^z - 1\right)^{\beta}} \tag{6.28}$$

Denote z_2^* as the zero point of (6.28), then the change point b^* is

$$b^* = \theta\left(z_2^*\right)^{1/\alpha} \tag{6.29}$$

Similar to t^*, the change point b^* of the FRF does not have closed form, but numerically reachable.

According to Xie *et al*. (2004), the useful period is defined as

$$\left\{t \mid h(t) \le (1 + k)h\left(b^*\right)\right\} \tag{6.30}$$

where k is a tolerant index. Denote the length of the useful period as l.

It is straightforward that the useful period is proportional to the scale parameter θ. Hence the relative difference d is independent of θ, and the length of useful period l is proportional to θ.

A numerical study is carried out to investigate the relationship between the shape parameters and t^*, b^*, d, l. The value of the scale parameter θ is fixed at 1 in the numerical study. α range from 1.2 to 10, and β ranges from $1/10\alpha$ to α, in order to guarantee the condition $\alpha\beta \le 1$. Results are summarized in the flowing tables.

Table 6.1 Change points of MRL

β \ α	$1/10\alpha$	$1/5\alpha$	$3/10\alpha$	$2/5\alpha$	$1/2\alpha$	$3/5\alpha$	$7/10\alpha$	$4/5\alpha$	$9/10\alpha$	α
1.2	0.5841	0.6305	**0.6404**	0.6315	0.6082	0.5704	0.5144	0.4313	0.2983	0.0141
1.5	0.3307	0.3838	0.4029	**0.4036**	0.3906	0.3649	0.3252	0.2672	0.1796	0.0065
2	0.2012	0.2508	0.2730	**0.2796**	0.2746	0.2591	0.2324	0.1914	0.1279	0.0001
2.5	0.1516	0.1980	0.2207	**0.2295**	0.2278	0.2166	0.1952	0.1607	0.1051	0.0001
3	0.1253	0.1693	0.1920	**0.2018**	0.2017	0.1926	0.1737	0.1422	0.0902	0.0001
4	0.0977	0.1380	0.1600	0.1705	**0.1718**	0.1645	0.1479	0.1190	0.0701	0.0001
5	0.0829	0.1215	0.1419	0.1526	**0.1542**	0.1478	0.1320	0.1042	0.0571	0.0006
6	0.0738	0.1117	0.1296	0.1402	**0.1423**	0.1360	0.1209	0.0939	0.0482	0.0022
7	0.0689	0.1012	0.1201	0.1322	**0.1343**	0.1274	0.1127	0.0859	0.0412	0.0062
8	0.0657	0.0980	0.1179	0.1241	**0.1273**	0.1195	0.1055	0.0791	0.0375	0.0101
9	0.0625	0.0920	0.1105	0.1224	**0.1224**	0.1153	0.0995	0.0767	0.0322	0.0169
10	0.0585	0.0897	0.1083	0.1167	**0.1185**	0.1098	0.0965	0.0733	0.0323	0.0254

From table 6.1, we can see that the change point t^* of MRL decreases as α increases; while as β increase, it initially increases and then decreases, with the maximum of each row highlighted in the table. It is interesting to note that when α is large (larger than 3), the maximum of t^* is achieved when β=1/2α. It is also observed that when $\alpha\beta$ is close to 1 (the last column), t^* is very small, which means that the optimal burn-in time is very short or can be neglected.

Table 6.2 Change points of FRF

α \ β	1/10α	1/5α	3/10α	2/5α	1/2α	3/5α	7/10α	4/5α	9/10α	α
1.2	1.7818	1.7312	1.5600	1.4238	1.3048	1.192	1.0753	0.9412	0.7627	**0.4386**
1.5	1.3839	1.2709	1.1774	1.0942	1.0157	0.9375	0.8550	0.7621	0.6476	0.4801
2	1.0606	1.0069	0.9570	0.9088	0.8607	0.8110	0.7577	0.6980	**0.6267**	0.5319
2.5	0.9429	0.9089	0.8756	0.8422	0.8079	0.7718	0.7327	**0.6888**	0.6370	0.5707
3	0.8886	0.8639	0.8390	0.8135	0.7870	0.7587	**0.7278**	0.6931	0.6525	0.6017
4	0.8456	0.8296	0.8131	**0.7958**	**0.7775**	**0.7577**	0.736	0.7116	0.6833	0.6489
5	0.8336	**0.8217**	**0.8093**	0.7962	0.7821	0.7669	0.7502	0.7315	0.7098	0.6838
6	**0.8321**	0.8226	0.8126	0.802	0.7907	0.7783	0.7647	0.7495	0.7320	0.7112
7	0.8348	0.8269	0.8185	0.8096	0.8000	0.7896	0.7781	0.7653	0.7506	0.7334
8	0.8391	0.8323	0.8251	0.8175	0.8092	0.8002	0.7903	0.7792	0.7666	0.7518
9	0.8441	0.8382	0.8318	0.8251	0.8178	0.8098	0.8011	0.7914	0.7803	0.7674
10	0.8493	0.8440	0.8383	0.8323	0.8257	0.8186	0.8108	0.8021	0.7923	0.7809

In contrast to t^*, the change point b^* of FRF decreases as β increases; while as α increases, it initially decreases and then increases, with the minimum of each column highlighted. It is also observed that b^* does not change as much as t^* when the values of α and β are altered, so it is not as important as the latter in determining the pattern of relative difference d and the length of useful life l.

Table 6.3 Relative Difference between the Change Points

α \ β	$1/10\alpha$	$1/5\alpha$	$3/10\alpha$	$2/5\alpha$	$1/2\alpha$	$3/5\alpha$	$7/10\alpha$	$4/5\alpha$	$9/10\alpha$	α
1.2	0.6722	0.6358	0.5895	0.5565	0.5339	**0.5215**	0.5216	0.5417	0.6089	0.9679
1.5	0.7610	0.6980	0.6578	0.6311	0.6154	**0.6108**	0.6196	0.6494	0.7226	0.9865
2	0.8103	0.7509	0.7147	0.6924	0.6810	**0.6805**	0.6933	0.7257	0.7959	0.9999
2.5	0.8392	0.7821	0.7479	0.7274	**0.7180**	0.7194	0.7336	0.7667	0.8350	0.9999
3	0.8590	0.8040	0.7712	0.7520	**0.7436**	0.7461	0.7613	0.7948	0.8618	0.9999
4	0.8844	0.8337	0.8032	0.7857	**0.7790**	0.7828	0.7990	0.8328	0.8975	0.9998
5	0.9006	0.8521	0.8246	0.8083	**0.8028**	0.8073	0.8241	0.8575	0.9196	0.9991
6	0.9113	0.8642	0.8405	0.8252	**0.8200**	0.8252	0.8419	0.8748	0.9341	0.9969
7	0.9175	0.8776	0.8533	0.8367	**0.8322**	0.8387	0.8552	0.8877	0.9451	0.9916
8	0.9218	0.8823	0.8571	0.8482	**0.8426**	0.8507	0.8665	0.8984	0.9511	0.9865
9	0.9259	0.8903	0.8671	0.8517	**0.8503**	0.8576	0.8758	0.903	0.9587	0.978
10	0.9311	0.8938	0.8708	0.8598	**0.8565**	0.8659	0.8809	0.9087	0.9592	0.9675

The relative difference d increases as α increases; while as β increases, it initially decreases and then increases, with the minimum of each row highlighted. This pattern is reasonable in considering that b^* is not as volatile as t^*, and hence the change of $d = \dfrac{b^* - t^*}{b^*}$ is more dependent on the latter factor t^*.

Table 6.4 Length of the Useful Period

α \ β	1/10α	1/5α	3/10α	2/5α	1/2α	3/5α	7/10α	4/5α	9/10α	α
1.2	1.6419	1.3345	1.2104	1.1374	1.0921	1.064	1.0471	1.0365	1.0250	**1.0090**
1.5	0.8844	0.8065	0.7570	0.7229	0.6985	0.6807	0.6679	0.6596	**0.6573**	0.6786
2	0.5440	0.5220	0.5063	0.4952	0.4875	0.4830	**0.4818**	0.4847	0.4947	0.5243
2.5	0.4179	0.4099	0.4045	0.4013	**0.4002**	0.4012	0.4050	0.4125	0.4266	0.4560
3	0.3526	0.3500	**0.3489**	0.3493	0.3512	0.3549	0.3611	0.3707	0.3862	0.4143
4	**0.2847**	0.2861	0.2884	0.2917	0.2962	0.3022	0.3102	0.3212	0.3372	0.3626
5	**0.2486**	0.2514	0.2549	0.2592	0.2646	0.2713	0.2799	0.2912	0.3067	0.3299
6	**0.2254**	0.2288	0.2328	0.2375	0.2433	0.2503	0.2590	0.2701	0.2850	0.3065
7	**0.2088**	0.2124	0.2167	0.2216	0.2275	0.2345	0.2432	0.2541	0.2684	0.2885
8	**0.1961**	0.1999	0.2042	0.2092	0.2151	0.2221	0.2307	0.2413	0.2551	0.2740
9	**0.1859**	0.1898	0.1942	0.1992	0.2051	0.2120	0.2204	0.2307	0.2440	0.2621
10	**0.1776**	0.1814	0.1858	0.1908	0.1966	0.2035	0.2117	0.2218	0.2347	0.2520

The length of the useful period l decreases as α increases; while as β increases, it initially decreases and then increases. It is also observed that l has a similar pattern as the absolute difference between t^* and b^*, which is the length of the period in which FRF $h(t)$ does not change dramatically.

From figure 6.5, we also observe that the flat portion of $h(t)$ tends to be low with small α and small β. Therefore, there exists a trade-off between the length of the useful period and the level of $h(t)$ during this period. In order to achieve a longer useful period, a unit having an odd Weibull distributed lifetime with small α and large β is preferred, while if the level of the random failure rate during the useful period is of most interest, a product having an odd Weibull lifetime model with small α and small β is more desirable.

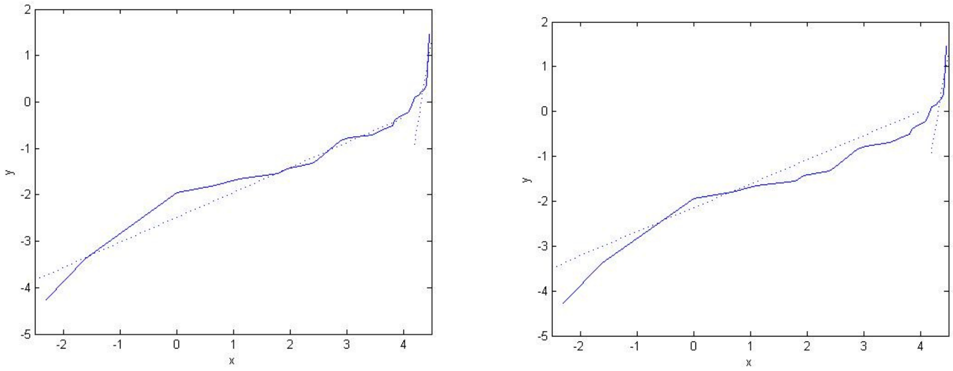
6.5 An Illustrative Example

The sample cited in Cooray (2006) from Aarset (1987) contains 50 device failure data. The TTT plot shows that the data have a bathtub-shaped failure rate. Therefore, the odd Weibull distribution may be suitable to model the data and the shape parameters should be in the region $\alpha > 1$ and $0 < \alpha\beta \leq 1$.

WPP plotting for the data is displayed in figure 6.6. Fit two straight lines to both sides of the curve:

$$L_1: \ y = 0.54x - 2.49, \ L_2: \ y = 7.64x - 33 \tag{6.31}$$

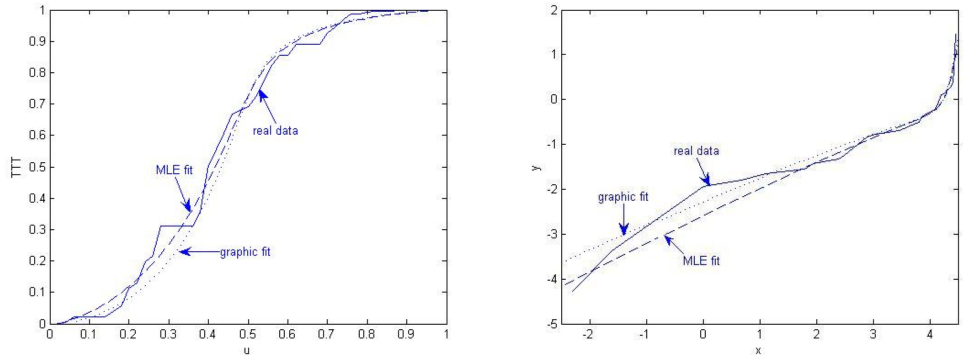
From L_1 and equation (6.17), $\tilde{\alpha}\tilde{\beta} = 0.54$. From L_2 and equation (6.18), $\tilde{\alpha} = 7.64$, and then $\tilde{\beta} = 0.071$, so L_3 is $y = 7.64x - 30.355$. The intersection point of L_1 and L_3 is $(3.925, -0.371)$. Since the point lies under the x-axis and the slope of L_3 is larger than L_1, we need to shift L_3 to the right and/or L_1 to the left. Empirical experience shows that shifting L_1 horizontally to the left by 0.638 lifts the intersection point to the x-axis and also ensures L_1 still fits the left tail of C well. See the following two figures.



(a) Fit to WPP L_1: y=0.54x-2.49, L_2: y=7.64x-33. (b) Fit to WPP L_1: y=0.54x-2.146, L_2: y=7.64x-33.

Figure 6.6 WPP and linear approximations

After the adjustment, WPP estimates of the model parameters can be obtained as $\tilde{\alpha} = 7.64, \tilde{\beta} = 0.071, \tilde{\theta} = 50.635$. These estimates are quite close to the MLE obtained in Cooray (2006) $\hat{\alpha} = 6.9657, \hat{\beta} = 0.0921, \hat{\theta} = 53.509$. The ML is thus the log-likelihood at the MLE point $L(\hat{\alpha}, \hat{\beta}, \hat{\theta}) = -215.88$. The TTT plot and WPP plot are illustrated in Figure 6.7.



(a) TTT plot of the device failure data. (b) WPP plot of the device failure data.

Figure 6.7 Modelling the Aarset (1987) device data with odd Weibull

To see whether the effort of modelling the data with a distribution having an additional shape parameter β to the 2-parameter Weibull distribution is worthwhile or not, we can do likelihood ratio test. The null hypothesis is $\beta = 1$, and the RML is defined as

$$\Lambda = -2\ln\left[L\left(\hat{\alpha},1,\hat{\theta}\right)/L\left(\hat{\alpha},\hat{\beta},\hat{\theta}\right)\right] \tag{6.32}$$

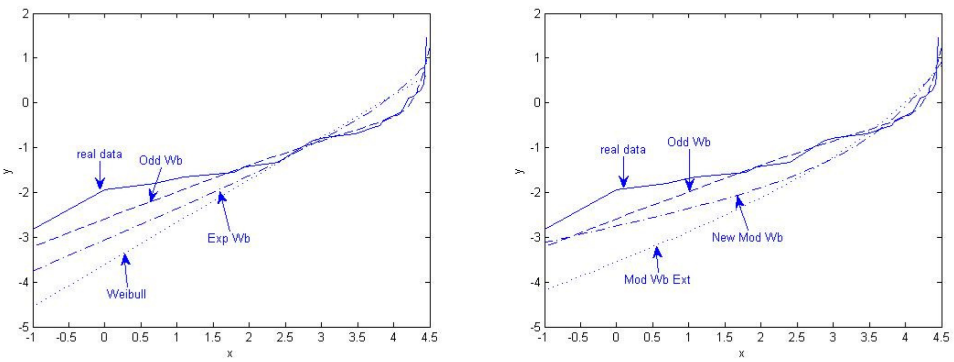
where $\left(\hat{\alpha},\hat{\theta}\right)$ are MLE of the parameters of the fitted Weibull distribution and $L\left(\hat{\alpha},1,\hat{\theta}\right)$ is the corresponding ML.

For this data set, $\left(\hat{\alpha},\hat{\theta}\right) = (0.949, 44.913)$ and $L\left(\hat{\alpha},1,\hat{\theta}\right) = -241.002$. Hence, the likelihood ratio is $\Lambda = 50.404$, and the corresponding p-value is $1.25*10^{-12}$. Under 99.9% significance level, we can reject the null hypothesis and conclude that the odd Weibull distribution provides a better fit to the data than the parameter Weibull distribution.

However, this conclusion may not be so convincing considering that the data exhibits a bathtub-shaped failure rate, while the Weibull distribution cannot produce such failure rate curve. To compare the goodness-of-fit of the odd Weibull distribution to this dataset with other bathtub-shaped Weibull-related distributions, we use the ML indexes. The exponentiated Weibull (Mudholkar and Srivastava (1993)), Weibull extension (Xie *et al*. (2002)) and modified Weibull (Lai *et al*. (2003)) are considered for the comparison.

Firstly, the MLE of the parameters of the exponentiated Weibull distribution obtained in Mudholkar and Srivastava (1993) are $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) = (4.69, 0.146, 91.023)$, and the corresponding ML is $L_1(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) = -229.114$. Secondly, the MLE of the parameters of the Weibull extension distribution obtained in Tang *et al*. (2003) are $(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) = (13.747, 0.588, 0.00876)$, and the corresponding ML is $L_2(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) = -231.647$. Finally, the MLE of the parameters of the modified Weibull distribution obtained in Ng (2005) are $(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) = (0.0624, 0.355, 0.02332)$, and the corresponding ML is $L_3(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) = \text{-}227.155$. All these MLs are smaller than the ML $L(\hat{\alpha}, \hat{\beta}, \hat{\theta}) = \text{-}215.800$ of odd Weibull fit. Therefore, if ML is considered as the indicator of goodness-of-fit, the odd Weibull performs the best among the four generalizations of the parameter Weibull distribution to model the lifetime data.

The following figure is plotted to illustrate the difference of fitting by the several lifetime distributions.



(a) WPP of odd Wb, Weibull and Exp Wb. (b) WPP of odd Wb, Mod Wb Ext and Mod Wb.

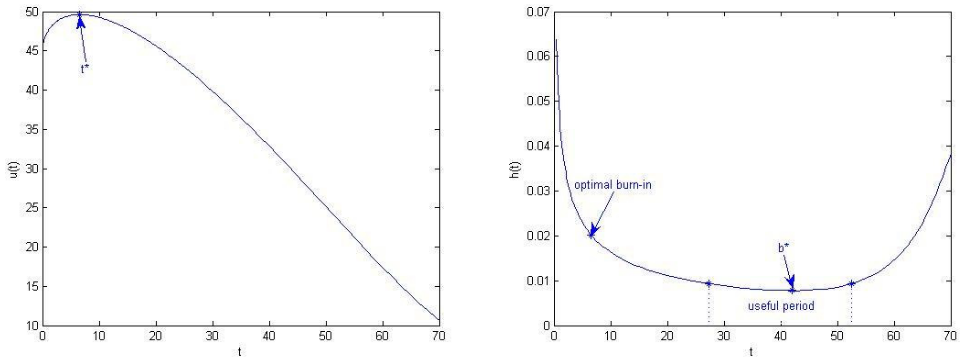
Figure 6.8 Comparison of fit among different distributions

It is easy to see from the above figure that the odd Weibull distribution provides far better fit than the other candidates to the empirical WPP.

With the odd Weibull parameter estimates obtained, the lifetime of the devices can be modeled

$$F(t) = 1 - \left\{ 1 + \left[\exp\left(\frac{t}{\hat{\theta}} \right)^{\hat{\alpha}} \right]^{\hat{\beta}} \right\}^{-1} \tag{6.33}$$

From (6.25), (6.26), (6.27) and (6.28), the change points of $h(t)$ and $\mu(t)$ are obtained numerically, $b^* = 41.9823$ and $t^* = 6.5607$, and hence $d = 0.8535$. So the optimal burn-in time based on the criterion of maximizing mean residual life is $t^* = 6.5607$, the useful period for a 20% tolerance index k is between time 27.2652 and 52.4744, and the length $l = 25.2092$.



(a) MRL $\mu(t)$ and its change point t^* (b) FRF $h(t)$, b^*, optimal burn-in time and useful period

Figure 6.9 MRL and FRF of the fitted odd Weibull model

6.6 Summary

In this chapter, the statistical properties of the newly proposed odd Weibull distribution are studied. This model is generated by evaluating randomness of the "odds of death" of a Weibull or inverse Weibull variable. WPP, the commonly used technique in Weibull analysis, is plotted for the distribution and used to obtain crude estimates of the parameters for a given sample data set. Finally, optimal burn-in and useful period related issues are discussed for the bathtub shaped failure rate curve. Numerical results support the conclusion that the random risk period of the bathtub shaped failure rate curve of the odd Weibull distribution is flat and long in most cases.

Chapter 7. Conclusion and Future Work

The main focus of the work presented in this thesis is to study the statistical properties of the Weibull models which can describe bathtub shaped failure rate data and parameter estimation problem associated with these models. This chapter summarizes the results of the research work and discusses their limitations and implications. Recommendations on further research and practical application are also presented.

7.1 Research Results

Weibull analysis is a powerful tool for analyzing lifetime data. Using the Weibull models to fit lifetime data is composed of data collection, model selection, parameter estimation and model application. In this thesis, a statistical characterization of the modified Weibull (Lai *et al.*, 2003) and odd Weibull (Cooray, 2006) is carried out and parameter estimation of the model parameters is discussed.

In Chapter 3, a detailed description of the important statistics and distribution functions of the modified Weibull distribution is presented. ML estimation procedure and usage of the likelihood-ratio test to make decisions regarding model choice is described. It is found that with the accelerating parameter λ, the modified Weibull distribution does not only extend the application of the Weibull and type-1 extreme value distributions to more monotonic FRF shapes, but also fits well to a variety of bathtub shaped failure rate data. In addition, due to the simple form of the distribution function, WPP and linear regression can both be applied to estimate the model parameters with a complete sample. All these benefits of the modified Weibull distribution makes it a good generalization of the Weibull distribution to model lifetime data, especially those with bathtub shaped failure rate.

Despite the convenience at application, graphic parameter estimation methods such as WPP are usually only able to produce very crude estimates when approximation or trial-and-error tests are needed. In such case, ML estimation would perform much better. The first half of Chapter 4 considers the log-likelihood function and Observed Fisher Information matrix for a class of distributions with a certain form. A technique is proposed to simplify the computation of the Observed Fisher Information matrix for the 3-parameter members of the class and another technique is proposed to decrease the number of unknown parameters in the log-likelihood function. Both techniques are useful in the consideration of computation and the latter one can also play a role in investigating the property of MLE for the members in the class.

Usually ML estimation is carried out by researchers to estimate model parameters of a statistical distribution without verifying the validity of doing so. However, sometimes this is risky that unexpected outcomes such as inability of convergence and multiple local maxima would be encountered. Fortunately, experiences show that MLE of the modified Weibull distribution exist and are unique. In the second half of Chapter 4, we examine and transform the log-likelihood function of the modified Weibull distribution via using the technique proposed earlier, and successfully prove the preferable properties of MLE of the modified Weibull parameters.

Under mild regularity conditions, MLE is asymptotically unbiased and the most efficient. However, for small size samples MLE is sometimes not as good as other estimation methods. Besides, for a parameter with small value as compared to others, bias of MLE could be so large that is several times of the true value. In such case, MCMC simulation provides less dispersed estimators, as well as easily constructed probability intervals. Chapter 5 narrates the details of obtaining MCMCE, compares the estimators with MLE and concludes that MCMC estimation for the parameters of the modified Weibull distribution is a good alternative to MLE for small size samples and can provide empirically exact probability intervals.

As a recently introduced Weibull model, the odd Weibull distribution is able to exhibit monotonic, unimodal and bathtub shaped failure rate. A detailed statistical characterization of this distribution is carried out in Chapter 6. WPP method is applied to estimate the model parameters, and it is shown that the estimation procedure is consistent and can achieve its unique stable point. Burn-in and related issues of the bathtub shaped failure rate curve are discussed, and the second portion of the curve is found to be flat and long, which makes the odd Weibull distribution a flexible and adaptable model for bathtub failure rate data.

7.2 Future Research

Rather than MLE and Bayesian estimation, there are still other parameter estimation methods. For the 3-parameter Weibull distribution, MPE seems to a good alternative and generalization of MLE. Future study of this method on the 3-parameter Weibull distribution would yield estimators that always exist and are consistent. For parameter estimation of the modified Weibull distribution, though multiple or nonlinear regression is thought to be crude, certain measures such as weighted least square may be taken to correct the error and reduce the bias. In addition, bias correction measures have been proposed to MLE of the Weibull parameters, so similar studies might be extended to MLE of the modified Weibull parameters.

The dimension decreasing technique proposed in Chapter 4 has only been applied to the Weibull and modified Weibull distributions, but the application can be extended to help investigate ML estimation of the parameters of other distributions such as the exponentiated Weibull (Mudholkar and Srivastava, 1993) and Weibull extension (Xie *et al*., 2002).

Confidence interval estimation is another important topic associated with the parameters of statistical distributions. Except for the MLE based normality approximation and MCMC simulation based empirical probability interval construction, the conditional (Lawless, 1973; Maswadah, 2003) and unconditional (Thoman *et al*., 1969) confidence intervals are interesting alternatives. The conditional method introduces a set of ancillary statistics, formulate conditional PDF to the parameters, and integrate the functions to get upper and lower bounds of the confidence intervals. The unconditional method makes use of the pivotal property of the parameters of the Weibull distribution and constructs the confidence intervals via Monte Carlo simulation. Chen (2004) and Chen and Chen (2009) did a good study of the simulation based confidence interval construction for the location parameter of the 3-parameter Weibull distribution, and their idea dates back to Thoman *et al*. (1969). Further discussions on the use of the methods for other members of the Weibull family would certainly be beneficial.

Combination of two or more models is usually a fast and easy way to generate new useful models, and this pattern of model building has straightforward physical or mechanical explanations, so is of vast application in real data modeling. For system life modeling, though combinations of the Weibull distribution has been studied extensively, few works are extended to the other members of the Weibull family. Bebbington *et al*. (2007[2]) studied the effect on MRL and the change points after adding a constant competing risk to a bathtub FRF. Because of the diversity of application and less difficulty in parameter estimation than other multi-parameter distributions, various combinations of different Weibull models are worth constructing and further exploring.

REFERENCES

Aarset, M. V. (1987). How to identify bathtub hazard rate, IEEE Transactions on Reliability, 36, 106-108.

Adamidis, K., Dimitrakopoulou, T. & Loukas, S. (2005). On an extension of the exponential-geometric distribution, Statistics & Probability Letters, 73, 259 – 269.

Ahmad, N., Bokhari, M. U., Quadri, S. M. K. & Khan, M. G. M. (2008). The exponentiated Weibull software reliability growth model with various testing-efforts and optimal release policy: A performance analysis, International Journal of Quality & Reliability Management, 25, 211-235.

Akahira, M. (1991). The 3/2th and 2nd order asymptotic efficiency of maximum probability estimators in non-regular cases, Annals of the Institute of Statistical Mathematics, 43, 181-195.

Ashour, S. K. & Afify, W. M. (2007). Statistical analysis of exponentiated Weibull family under type I progressive interval censoring with random removals, Journal of Applied Sciences Research, 3, 1851-1863.

Baftery, A. E. & Banfield, J. D. (1991). Stopping the Gibbs Sampler, the use of morphology, and other issues in spatial statistics, Annals of the Institute of Statistical Mathematics, 43, 32-43.

Bhattacharya P. & Bhattacharjee, R. (2009). A study on Weibull distribution for estimating the parameters, Wind Engineering, 33, 469-476.

Bain, L. J. (1978). Statistical analysis of reliability and life testing models, Marcel Dekker, New York.

Bain, L. J. & Antle, C. E. (1967). Estimation of parameters of the Weibull distribution, Technometrics, 9, 621-627.

Bairamov, I. (2006). Progressive type II censored order statistics for multivariate observations, Journal of Multivariate Analysis, 97, 797-809.

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.

Balakrishnan, N. & Mi, J. (2003). Existence and uniqueness of the MLE for normal distribution based on general progressively type II censored samples, Statistics & Probability Letters, 64, 407-414.

Balasooriya, U. & Low, C. K. (2004). Competing causes of failure and reliability tests for Weibull lifetimes under type I Progressive Censoring, IEEE Transactions on Reliability, 53, 29-36.

Banerjee, A. & Kundo, D. (2008). Inference Based on Type II Hybrid Censored Data From a Weibull Distribution, IEEE Transactions on Reliability, 57, 369-378.

Barlow, R. E. & Campo, R. Total time on test processes and applications to failure data analysis, In Barlow, R. E., Fussel, J. B. & Singpurwalla N. D. (1975). Reliability and Fault Tree Analysis, 451-481.

Bebbington, M., Lai, C. D. & Zitikis, R. (2006). Useful periods for lifetime distributions with bathtub shaped hazard rate functions. IEEE Transactions on Reliability,

Bebbington, M., Lai, C. D. & Zitikis, R. (2007)(1). A flexible Weibull extension, Reliability Engineering and System Safety, 92, 719-726.

Bebbington, M., Lai, C. D. & Zitikis, R. (2007)(2) . Reduction in mean residual life in the presence of a constant competing risk, Applied Stochastic Models in Business and Industry, 24, 51 – 63.

Bebbington, M., Lai, C. D. & Zitikis, R. (2008). Estimating the turning point of a bathtub-shaped failure distribution, Journal of Statistical Planning and Inference, 138, 1157-1166.

Belisle, C. (1998). Slow convergence of Gibbs sampler, The Canadian Journal of Statistics, 26, 629-641,

Benard, A. & Bos-Levenbach, E. C. (1953). The plotting of observations on probability paper, Statistica Neerlandica, 7, 163-173.

Bergman, B. & Klefsjo, B. (1984). The total time on test concept and its use in reliability theory, Operations Research, 32, 596-606.

Bergman, B. (1986). Estimation of Weibull parameters using a weight function, Journal of Material Science Letters, 5, 611-614.

Blischke, W. R. (1974). On nonregular estimation of the location parameter of the gamma and Weibull distributions, Communications in Statistics, 3, 1109 – 1129

Borzadaran, G. R. M., Yari, G. H. & Pasha, E. (2007). Information measures for some well-known families, Communications in Statistics – Theory and Methods, 36, 669-677.

Bousquet, N., Bertholon, H. & Celeux, G. (2006). An alternative competing risk model to the Weibull distribution for modelling aging in lifetime data analysis, Lifetime Data Analysis, 12, 481–504.

Canavos, G. C. & Tsokos, C. P. (1973). Bayesian estimation of life parameters in the Weibull distribution, Operations Research, 21, 755-763.

Cancho, V. G. & Bolfarine, H. (2001). Modeling the presence of immunes by using the exponentiated-Weibull model, Journal of Applied Statistics, 28, 659-671.

Cancho, V. G., Bolfarine, H. & Achcar, J. A. (1999). A Bayesian analysis for the exponentiated-Weibull distribution, Journal of Applied Statistical Science, 8, 227-242.

Cao, Q. V. & McCarty, S. M. (2006). New methods for estimating parameters of weibull functions to characterize future diameter distributions in forest stands, Proceedings of the 13th Biennial Southern Silvicultural Research Conference, 338-340.

Cao, Q. V. (2005). Predicting parameters of a Weibull function for modeling diameter distribution, Forest Science, 50, 682-685.

Carrasco, J. M. F., Ortega, E. M. M. & Paula, G. A. (2008)[(2)]. Log-modified Weibull regression models with censored data: sensitivities and residual analysis, Computational Statistics and Data Analysis, 52, 4021-4039.

Casella, G. & George, E. I. (1992). Explaining the Gibbs sampler, The American Statistician, 46, 167-174.

Castet, J.-F. & Saleh, J. H. (2009). Single versus mixture Weibull distributions for nonparametric satellite reliability, Reliability Engineering and System Safety, 95, 295-300.

Carta, J. A. & Ramirez, P. (2007). Analysis of two-component mixture Weibull statistics for estimation of wind speed distributions, Renewable Energy, 32, 518-531.

Casella, G. & George, E. I. (1992). Explaining the Gibbs Sampler, American Statistician, 46, 167-174.

Chen, D. & Chen, Z. (2009). Statistical inference about the location parameter of the three-parameter Weibull distribution, Journal of Statistical Computation and Simulation, 79, 215–225.

Chen, K. W., Papadopoulos, A. S. & Tamer, P. (1989). On Bayes estimation for mixtures of two Weibull distributions under type I censoring, Microelectronics Reliability, 29, 609-617.

Chen, Z. (2004). Exact confidence intervals and joint confidence regions for the parameters of the Weibull distribution, International Journal of Reliability, Quality and Safety Engineering, 11, 133-140.

Cheng, R. C. & Amin, N. A. K. (1983). Estimating Parameters in Continuous Univariate Distributions with a Shifted Origin, Journal of Royal Statistical Society B, 45, 394-403.

Chernoff, H. & Lehmann, E. L. (1954). The use of maximum likelihood estimates in χ^2 tests for goodness of fit, The Annals of Mathematical Statistics, 25, 579-586.

Choudhury, A. (2005). A simple derivation of moments of the exponentiated Weibull distribution, Metrika, 62, 17 – 22.

Cohen, A. C. & Whitten, B. J. (1982). Modified maximum likelihood and modified moment estimators for the three-parameter Weibull distribution, Communication in statistics – Theory and Methods, 11, 2631-2656.

Cohen, A. C. (1965). Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples, Technometrics, 7, 579-588.

Cohen, A. C., Whitten, B. J. & Ding, Y. (1984). Modified moment estimation for the three-parameter Weibull distribution, Journal of Quality Technology, 16, 159-167.

Cooray, K. (2006). Generalization of the Weibull distribution: the odd Weibull family, Statistical Modelling, 6, 265-277.

Cousineau, D., Brown, S. & Heathcote, A. (2004). Fitting Distributions Using Maximum Likelihood: Methods and Packages, Behavior Research Methods, Instruments, & Computers, 36, 742-756.

Cousineau, D. (2009). Fitting the Three-Parameter Weibull Distribution: Review and Evaluation of Existing and New Methods, IEEE Transactions on Dielectrics and Electrical Insulation, 16, 281-288.

Cousineau, D. (2009)[(1)]. Nearly unbiased estimators for the three-parameter Weibull distribution with greater efficiency than the iterative likelihood method, British Journal of Mathematical and Statistical Psychology, 62, 167–191.

Cowles, M. K. & Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review, Journal of the American Statistical Association, 91, 883-904.

Cran, G. W. (1988). Moment estimators for the 3-parameter Weibull distribution, IEEE Transactions on Reliability, 37, 360-363.

David, H. A. (1970). Note: on Chiang's proportionally assumption in the theory of competing risks, Biometrics, 26. 336-339.

De Haan, L. & Ferreira, A. (2006). Extreme Value Theory, Springer.

Dellaportas, P. & Wright, D. (1991). Positive embedded integration in Bayesian analysis, Statistics and Computing, 1, 1-12.

Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm, Journal of the Royal Statistical Society, 39, 1-38.

Dimitrakopoulou, T. D., Adamidis, K. & Loukas, S. (2007). A lifetime distribution with an upside-down bathtub-shaped hazard function, IEEE Transactions on Reliability, 56, 308-311.

Dodson, B. (2006). The Weibull analysis handbook, ASQ Quality Press, Milwaukee.

Drapella, A. (1993). The complementary Weibull distribution: unknown or just forgottern? Quality and Reliability Engineering International, 9, 383-385.

Drapella, A. (1999). An Improved Failure-Free Estimation Method, Quality and Reliability Engineering International, 15, 235-238.

Dubey, S. D. (1966). Hyper-Efficient Estimator of the Location Parameter of the Weibull Law, Naval Research Logistic Quarterly, 13, 253-264.

Dubey, S. D. (1967). Some percentile estimators for Weibull parameters, Technometrics, 9, 119-129.

Dudewicz, E. J. (1973). Maximum probability estimators for ranked means, Annals of the Institute of Statistical Mathematics, 25, 467-477.

Efron, B. & Hinkley, D. V. (1978). Assessing the accuracy of the maximum likelihood estimator: observed versus expected Fisher information, Biometrika, 65, 457-482.

Engelhardt, M. & Bain, L. J. (1974). Some results on point estimation for the two-parameter Weibull or Extreme-value distribution, Technometrics, 16, 49-56.

Farnum, N. R. & Booth, P. (1997). Uniqueness of maximum likelihood estimators of the 2-parameter Weibull distribution, IEEE Transactions on Reliability, 36, 523-525.

Ferarri, S. L. P., Silva, M. R. D. & Cribari-neto, F. (2007). Adjusted profile likelihoods for the Weibull shape parameter, Journal of Statistical Computation and Simulation, 77, 531-548.

Gaeddert, J. & Annamalai, A. (2005). New estimators for the Weibull fading parameters, Proceedings of Vehicular Technology Conference, 2, 1367-1371.

Gelfand A. E. & Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities, Journal of the American Statistical Association, 85, 398-409.

Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. (2004). Bayesian Data Analysis, 2nd edition, Chapman & Hall.

Gelman, A. & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences, Statistical Science, 7, 457-472.

Geman, S. & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 6, 721-741.

Gertsbakh, I. & Kagan, A. (1999). Characterization of the Weibull distribution by properties of the Fisher information under type I censoring, Statistics & Probability Letters, 42, 99-105.

Ghitany, M. E., Al-Hussaini, E. K. & Al-Jarallah, R. A. (2005). Marshall-Olkin extended weibull distribution and its application to censored data, Journal of Applied Statistics, 32, 1025 – 1034.

Ghitany, M. E., Al-Hussaini, E. K. & Alkhalfan, L. A. (2007). Marshall-Olkin extended lomax distribution and Its application to censored data, Communications in Statistics-Theory and Methods, 36, 1855 – 1866.

Ghitany, M.E. & Kotz, S. (2007). Reliability properties of extended linear failure-rate distributions, Probability in the Engineering and Informational Sciences, 21, 441-450.

Gilks, W. R. & Wild, P. (1992). Adaptive rejection sampling for Gibbs sampler, Applied Statistics, 41, 337-348.

Gilks, W. R., Best, N. G. & Tan, K. K. C. (1995). Adaptive rejection Metropolis sampling within Gibbs sampling, Applied Statistics, 44, 455-472.

Godambe, V. P. (1960). An optimum property of regular maximum likelihood estimation, The Annals of Mathematical Statistics, 31, 1208-1211.

Gong, G. & Samaniego, F. J. (1981). Pseudo maximum likelihood estimation: theory and applications, The Annals of Statistics, 9, 861-869.

Gong, Z. J. (2006). Estimation of mixed Weibull distribution parameters using the SCEM-UA algorithm: application and comparison with MLE in automotive reliability analysis, Reliability Engineering and System Safety, 91, 915-922.

Green, E. J., Roesch, F. A., Jr., Smith, A. F. M. & Strawderman, W. E. (1994). Bayesian estimation for the three-parameter Weibull distribution with tree diameter data, Biometrics, 50, 254-269.

Gupta, A., Mukherjee, B. & Upadhyay, S. K. (2008). Weibull extension model: a Bayes study using Markov chain Monte Carlo simulation, Reliability Engineering and System Safety, 93, 1434-1443.

Gupta, R. C. & Akman, H. O. (1995). Mean residual life functions for certain types of non-monotonic ageing, Communications in Statistics – Stochastic Models, 1995, 11, 219–225.

Gupta, R. C., Akman, O. & Lvin, S. (1999). A study of log-logistic model in survival analysis, Biometrical Journal, 41, 431-443.

Gupta, R. D. & Kundu, D. (2006). On the comparison of Fisher information of the Weibull and GE distributions, Journal of Statistical Planning and Inference, 136, 3130-3144.

Gurvich, M. R., DiBenedetto, A. T. & Ranade, S. V. (1997). A new statistical distribution for characterizing the random length of brittle materials, Journal of Materials Science, 32, 2559-2564.

Harter, H. L. & Moore, A. H. (1965). Maximum-likelihood estimation of the parameters of Gamma and Weibull populations from complete and censored samples, Technometrics, 7, 639-643.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications, Biometrika, 57, 97-109.

Heathcote, A., Brown, S. & Mewhort, D. J. K. (2002). Quantile Maximum Likelihood Estimation of Response Time Distributions, Psychonomic Bulletin & Review, 9, 394-401.

Heathcote, A., Brown, S. & Cousineau, D. (2004). QMPE: Estimating Lognormal, Wald and Weibull RT distributions with a parameter dependent lower bound, Behavior Research Methods, Instruments, & Computers, 36, 277-290.

Hirose, H. (1996). Maximum Likelihood Estimation in the 3-Parameter Weibull Distribution, IEEE Transactions on Dielectrics and Electrical Insulation, 3, 43-55.

Hirose, H. & Lai, T. L. (1997). Inference from grouped data in three-parameter Weibull models with applications to breakdown-voltage experiments, Technometrics, 39, 199-210.

Hirose, H. (1999). Bias correction for the maximum likelihood estimates in the two-parameter Weibull distribution, IEEE Transactions on Electrical Insulation, 6, 66-68.

Hirose, H. (2002). Maximum likelihood parameter estimation in the extended Weibull distribution and its applications to breakdown voltage estimation, 9, 524-536.

Hossain, A. & Howlader, H. A. (1996). Unweighted least squares estimation of Weibull parameters, Journal of Statistical Computation and Simulation, 54, 265-271.

Hung, W. L. (2001). Weighted Least-Squares Estimation of the Shape Parameter of the Weibull Distribution, Quality and Reliability Engineering International, 17, 467-469.

Hung, W. L. & Liu, Y. C. (2005). Estimation of Weibull parameters using a fuzzy least-squares method, International Journal of Uncertainty, Fuzziness and Knowledge -Based Systems, 12, 701-711.

Ibrahim, J. G., Chen, M. H. & Sinha, D. (2001). Bayesian Survival Analysis, Springer, NewYork.

Jennrich, P. I. & Sampson, P. F. (1976). Newton-Raphson and related algorithms for maximum likelihood variance component estimation, Technometrics, 18, 11-17.

Jiang, H., Xie, M. & Tang, L. C. (2008)(1). Markov chain Monte Carlo methods for parameter estimation of the modified Weibull distribution, Journal of Applied Statistics, 35, 647-658.

Jiang, H., Xie, M. & Tang, L. C. (2008)(2). On the odd Weibull distribution, Proceedings of the Institution of Mechanical Engineers, Part O, Journal of Risk and Reliability, 222, 583-594.

Jiang, H., Xie, M. & Tang, L. C. (2010). On MLEs of the parameters of a modified Weibull distribution for progressively type-2 censored samples, Journal of Applied Statistics, 34, 617-627.

Jiang, R. & Murthy, D. N. P. (1995). Modeling failure-data by mixture of 2 Weibull distributions: a graphical approach, IEEE Transactions on Reliability, 44, 477-488.

Jiang, R. & Murthy, D. N. P. (1997(1)). Parametric study of competing risk model involving two Weibull distributions, International Journal of Reliability, Quality and Safety Engineering, 4, 17-34.

Jiang, R. & Murthy, D. N. P. (1997(2)). Parametric study of multiplicative model involving two Weibull distributions, Reliability Engineering and System Safety, 55, 217-226.

Jiang, R. & Murthy, D. N. P. (1999). The exponentiated Weibull family: a graphic approach, IEEE Transactions on Reliability, 48, 68-72.

Jiang, R. & Murthy, D. N. P. (2003). Study of *n*-fold Weibull competing risk model, Mathematical and Computer Modelling, 38, 1259-1273.

Jiang, R., Murthy, D. N. P. & Ji, P. (2001). Models involving two inverse Weibull distributions, Reliability Engineering and System Safety, 73, 73-81.

Jiang, R., Zuo, M. J. & Murthy, D. N. P. (1999). Two sectional models involving two Weibull distributions, International Journal of Reliability, Quality and Safety Engineering, 6, 103-122.

Jiang, S. Y. & Kececioglu, D. (1992). Graphical representation of 2 mixed-Weibull distributions, IEEE Transactions on Reliability, 41, 241-247.

Jukić, D., Benšić, M. & Scitovski, R. (2008). On the existence of the nonlinear weighted least squares estimate for a three-parameter Weibull distribution, Computational Statistics and Data Analysis, 52, 4502-4511.

Kaminskiy, M. P. & Krivtsov, V. V. (2005). A simple procedure for Bayesian estimation of the Weibull distribution, IEEE Transactions on Reliability, 54, 612-616.

Kantar, Y. M. & Senoglu, B. (2008). A comparative study for the location and scale parameters of the Weibull distribution with given shape parameter, Computers & Geosciences, 34, 1900 – 1909.

Kao, J. H. K. (1958). Computer methods for estimating Weibull parameters in reliability studies, I. R. E. Transactions. Reliability and Quality Control, 13, 15-22.

Kao, J. H. K. (1959). A graphical estimation of mixed Weibull parameters in life-testing of electron tubes, Technometrics, 1, 389-407.

Khan, M. S., Pasha, G. R. & Pasha, A. H. (2008). Theoretical analysis of inverse Weibull distribution, WSEAS Transactions on Mathematics, 7, 30-38.

Kottas, A. (2006). Nonparametric Bayesian survival analysis using mixtures of Weibull distributions, Journal of Statistical Planning and Inference, 136, 578-596.

Kotz, S. & Nadarajah, S. (2000). Extreme value distributions: theory and applications, Imperial College Press, London.

Krishnamoorthy, K., Lin, Y. & Xia, Y. (2009). Confidence limits and prediction limits for a Weibull distribution based on the generalized variable approach, Journal of Statistical Planning and Inference, doi: 10.1016/j.jspi.2008.12.010.

Kuβ, U. (1972). Contributions to maximum probability estimators, Probability Theory and Related Fields, 24, 123-133.

Kumar, U., Klefsjo, B. & Granholm, S. (1989). Reliability investigation for a fleet of load haul dump machines in a Swedish mine, Reliability Engineering and System Safety, 16, 341-361.

Kundo, D. (2008). Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring, Technometrics, 50, 144-154.

Lai, C. D., Xie, M. & Murty, D. N. P. (2003). A modified Weibull distribution, IEEE Transactions on Reliability, 52, 33-37.

Lai, C. D., Zhang, L. Y. & Xie, M. (2004). Mean residual life and other properties of Weibull related bathtub shape failure rate distributions, International Journal of Reliability, Quality and Safety Engineering, 11, 113-132.

Lawless, J. F. (1973). Conditional versus unconditional confidence intervals for the parameters of the Weibull distribution, Journal of American Statistical Association, 68, 665-669.

Lemon, G. H. (1975). Maximum likelihood estimation for the three parameter Weibull distribution based on censored samples, Technometrics, 17, 247-254.

Lockhart, R. A. & Stephen, M. A. (1994). Estimation and Tests of Fit for the 3-Parameter Weibull Distribution, Journal of the Royal Statistical Society B, 56, 491-500.

Lu, H. L., Chen, C. H. & Wu, J. W. (2004) A note on weighted least-squares estimation of the shape parameter of the Weibull distribution, Quality and Reliability Engineering International, 20, 579-586.

Makelainen, T., Schmidt, K. & Styan, G. P. H. (1981). On the existence and uniqueness of the maximum likelihood estimate of a vector-valued parameter in fixed-size samples, The Annals of Statistics, 9, 758-767.

Mann, N. R., Schafer, R. E. & Singpurwalla, N. D. (1974). Methods for statistical analysis of reliability and life data, Wiley, New York.

Marković, D., Jukić, D. & Benšić, M. (2009). Nonlinear weighted least squares estimation of a three-parameter Weibull density with a nonparametric start, Journal of Computational and Applied Mathematics, 228, 304-312.

Marks, N. B. (2005). Estimation of Weibull parameters from common percentiles, Journal of Applied Statistics, 32, 17-24.

Marshall, A. W. & Olkin, I. (1997). A new method for adding a parameter to a family of distribution with application to the Exponential and Weibull families, Biometrika, 84, 641-652.

Maswadah, M. (2003). Conditional confidence interval estimation for the inverse weibull distribution based on censored generalized order statistics, Journal of statistical computation and simulation, 73, 887-898.

Mendenhall, W. & Hader, R. J. (1958). Estimation of parameters of mixed exponentially distribution failure time distributions from censored life test data, Biometrika, 45, 504-520.

Merganič, J. & Sterba, H. (2006). Characterisation of diameter distribution using the Weibull function: method of moments, European Journal of Forest Research, 125, 427-439

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N. & Teller, A. H. (1953). Equation of state calculations by fast computing machines, The Journal of Chemical Physics, 21, 1087-1092.

Mi, J. (2006). MLE of parameters of location-scale distribution for complete and partially grouped data, Journal of Statistical Planning and Inference, 136, 3565-3582.

Mittal, M. M. & Dahiya, R. C. (1989). Estimating the parameters of a truncated Weibull distribution, Communications in Statistics – Theory and Methods, 18, 2027-2042.

Montanari G. C., Mazzanti, G., Cacciari, M. & Fothergill, J. C. (1997). In search of convenient techniques for reducing bias in the estimation of Weibull parameters for uncensored test, IEEE Transactions on Electrical Insulation, 4, 306-313.

Mosler, K. & Scheicher, C. (2008). Homogeneity testing in a Weibull mixture model, Statistical Papers, 49, 315-332.

Mudholkar, G. S. & Hutson, A. D. (1996). The exponentiated Weibull family: some properties and a flood data application, Communications in Statistics – Theory and Methods, 25, 3059-3083.

Mudholkar, G. S. & Kollia, G. D. (1994). Generalized Weibull family: a structural analysis, Communications in Statistics – Theory and Methods, 23, 1149-1171.

Mudholkar, G. S. & Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data, IEEE Transactions on Reliability, 42, 299-302.

Mudholkar, G. S., Srivastava, D. K. & Freimer, M. (1995). The exponentiated Weibull family: a reanalysis of the bus-motor-failure data, Technometrics, 37, 436-445.

Mudholkar, G. S., Srivastava, D. K. & Kollia, G. D. (1996). A generalization of the Weibull distribution with application to the analysis of survival data, Journal of the American Statistical Association, 91, 1575-1583.

Murthy, D. N. P. & Jiang, R. (1997). Parametric study of sectional models involving two Weibull distributions, Reliability Engineering and System Safety, 56, 151-159.

Murthy, D. N. P., Bulmer, M. & Eccleston, J. A. (2004)(1). Weibull model selection for reliability modelling, Reliability Engineering and System Safety, 86, 257-267.

Murthy, D. N. P., Xie, M. & Jiang, R. (2004)(2). Weibull models, John Wiley & Sons, New Jersey.

Muth, E. J. (1977). The theory and applications of reliability. In Reliability Models with Positive Memory Derived From the Mean Residual Life Function, Academic Press, London.

Nadarajah, S. & Gupta, A. K. (2005). On the moments of the exponentiated Weibull distribution, Communications in Statistics – Theory and Methods, 34, 253-256.

Nadarajah, S. & Kotz, S. (2005). On some recent modifications of Weibull distribution, IEEE Transactions on Reliability, 54, 561-562.

Nadarajah, S. & Kotz, S. (2007). On the Moment of the Maxima of Weibull Random Variables, IEEE Transactions on Vehicular Technology, 56, 1467-1468.

Nagode, M. & Fajdiga, M. (2000). An improved algorithm for parameter estimation suitable for mixed Weibull distributions, International Journal of Fatigue, 22, 75-80.

Nandi, S. & Dewan, I. (2010). An EM algorithm for estimating the parameters of bivariate Weibull distribution under censoring, Computational Statistics and Data Analysis, 54, 1559-1569.

Nassar, M. M. & Eissa, F. H. (2003). On the exponentiated Weibull distribution, Communications in Statistic s – Theory and Methods, 32, 1317-1336.

Nassar, M. M. & Eissa, F. H. (2004). Bayesian estimation for the exponentiated Weibull model. Communication in statistics – Theory and Methods, 33, 2343-2362.

Nelson, W. B. (1982). Applied life Data Analysis, John Wiley, New York.

Ng, H. K. T. (2005). Parameter Estimation for a Modified Weibull Distribution, for Progressively Type II Censored Samples, IEEE Transactions on Reliability, 54, 374-380.

O'Connor, P. D. T. (2002). Practical Reliability Engineering, John Wiley & Sons, Chichester.

Offinger, R. (1998). Least Squares and Minimum Distance Estimation in the 3-Parameter Weibull and Frechet Models with Applications to River Drain Data, Advances in Stochastic Models for Reliability, Quality and Safety (Kahle *et al.*), Boston, Birkhuser, 81-97.

Ortega E.M.M., Cancho V.G. & Bolfarine H. (2006). Influence diagnostics in exponentiated-Weibull regression models with censored data, Statistics and Operations Research Transactions, 30, 171-196.

Pang, W. K., Forster, J. J. & Troutt, M. D. (2001). Estimation of wind speed distribution using Markov chain Monte Carlo techniques, Journal of Applied Meterology, 40, 1476-1484.

Pang, W. K., Hou, S. H. & Yu, W. T. (2007). On a proper way to select population failure distribution and a stochastic optimization method in parameter estimation, European Journal of Operational Research, 177, 604 – 611.

Pang, W. K., Leung, P. K., Huang, W. K. & Liu, W. (2005). On interval estimation of the coefficient of variation for the three-parameter Weibull, Lognormal and Gamma distribution: A simulation-based approach, European Journal of Operational Research, 164, 367-377.

Pascual, F. (2007). Accelerated life test planning with independent Weibull competing risks with known shape parameter, IEEE Transactions on Reliability, 56, 85-93.

Pascual, F. (2008). Accelerated life test planning with independent Weibull competing risks, IEEE Transactions on Reliability, 57, 435-444

Rekkas, M. & Wong, A. (2005). Third-order inference for theWeibull distribution, Computational Statistics & Data Analysis, 49, 499 – 525.

Roberts, G. O. & Smith, A. F. M. (1994). Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms, Stochastic Process and their Applications, 49, 207-216.

Rockette, H., Antle, C. & Klimko, L. A. (1974). Maximum likelihood estimation with the Weibull model, Journal of American Statistical Association, 69, 246-249.

Ross, R. (1996). Bias and standard deviation due to Weibull parameter estimation for small data sets, IEEE Transactions on Electrical Insulation, 3, 28-42.

Roussas, G. G. (1977). Asymptotic properties of the maximum probability estimates in Markov processes, Annals of the Institute of Statistical Mathematics, 29, 203-219.

Sankaram, P. G. & Jayakumar, K. (2007). On proportional odds models, Statistical Papers, DOI 10.1007/s00362-006-0042-3.

Singh, U., Gupta, P. K. & Upadhyay, S. K. (2005). Estimation of three-parameter exponentiated-Weibull distribution under type II censoring, Journal of Statistical Planning and Inference, 134, 350-372.

Sinha, S. K. & Sloan, J. A. (1988). Bayes estimation of the parameters and reliability function of the 3-parameter Weibull distribution, IEEE Transactions on Reliability, 37, 364-369.

Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases, Biometrika, 72, 67-90.

Smith, R. L. & Naylor, J. C. (1987). A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution, Applied Statistics, 36, 358-369.

Sultan, K. S., Ismail, M. A. & Al-Moisheer, A. S. (2007). Mixture of two inverseWeibull distributions: properties and estimation, Computational Statistics & Data Analysis, 51, 5377 – 5387.

Swain, J., Venkatraman, S. & Wilson, J. (1988). Least Squares Estimation of Distribution Function in Johnson's Translation System, Journal of Statistical Computation and Simulation, 29, 271 – 297.

Tan, Z. (2009). A new approach to MLE of Weibull distribution with interval data, Reliability Engineering and System Safety, 94, 394– 403.

Tang, L. C. (2003). Failure-Free Life in Reliability Modeling, International Journal of Industrial Engineering, 10, 332-338.

Tang, Y., Xie, M. & Goh, T. N. (2003). Statistical analysis of a Weibull extension model, Communications in Statistics – Theory and Methods, 32, 913-928.

Thoman, D. R., Bain, L. J. & Antle, C. E. (1969). Inference on the parameters of the Weibull distribution, Technometrics, 11, 445-460.

Thoman, D. R., Bain, L. J. & Antle, C. E. (1970). Maximum likelihood estimation, exact confidence intervals, and tolerance limits in the Weibull distribution, Technometrics, 12, 363-371.

Tiku, M. L. (1967). Estimating the Mean and Standard Deviation from a Censored Normal Sample, Biometrika, 54, 155-165.

Tiku, M.L. & Akkaya, A.D. (2004). Robust Estimation and Hypothesis Testing, New Age International, New Delhi.

Tiryakioglu, M. & Hudak, D. (2008). Unbiased estimates of the Weibull parameters by the linear regression method, Journal of Materials Science, 43, 1914-1919.

Touw, A. E. (2009). Bayesian estimation of mixed Weibull distributions, Reliability Engineering and System Safety, 94, 463-473

Tsionas, E. G. (2002). Bayesian analysis of finite mixtures of Weibull distributions, Communications in Statistics – Theory and Methods, 31, 37-48.

Upadhyay, S. K. & Peshwani, M. (2003). Choice between Weibull and Lognormal models: a simulation based Bayesian study, Communications in Statistics-Theory and Methods, 32, 381 – 405.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypothesis, Econometrica, 57, 307-333.

Wang, F. K. & Keats, J. B. (1995). Improved percentile estimation for the two-parameter Weibull distribution, Microelectronics Reliability, 35, 883-892.

Wang, F. K. (2000). A new model with bathtub-shaped failure rate using an additive Burr XII distribution, Reliability Engineering and System Safety, 70, 305-312.

Wang, R. H. & Fei, H. L. (2003). Uniqueness of the maximum likelihood estimate of the Weibull distribution tampered failure rate model, Communications in Statistics – Theory and Methods, 32, 2321-2338.

Watkins, A. J. (1996). On maximum likelihood estimation for the two parameter Weibull distribution, Microelectronics Reliability, 36, 595-603.

Wegner, H. (1976). On the existence of maximum probability estimators, Annals of the Institute of Statistical Mathematics, 28, 343-347.

Weibull, W. (1939). A statistical theory of the strength of material, Royal Swedish Institute for Engineering Research, 151, 1-45.

Weibull, W. (1951). A statistical distribution function of wide applicability, Journal of Applied Mechanics, 18, 293-297.

Weiss, L. & Wolfowitz, J. (1967). Maximum probability estimators, Annals of the Institute of Statistical Mathematics, 19, 193-206.

Weiss, L. & Wolfowitz, J. (1970). Maximum probability estimators and asymptotic sufficiency. Annals of the Institute of Statistical Mathematics, 22, 225-244.

White, J. S. (1969). The moments of log-Weibull order statistics, Technometrics, 11, 373-386.

Wu, D., Zhou, J. & Li, Y. (2006). Unbiased estimation of Weibull parameters with the linear regression method, Journal of the European Ceramic Society, 26, 1099–1105.

Wu, J. W., Lu, H. L., Chen, C. H. & Wu, C. H. (2004). Statistical inference about the shape parameter of the new two-parameter bathtub-shaped lifetime distribution, Quality and Reliability Engineering International, 20, 607-616.

Xie, M. & Lai, C. D. (1996). Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function, Reliability Engineering and System Safety, 52, 87-93.

Xie, M., Goh, T. N. & Tang, Y. (2004). On changing points of mean residual life and failure rate function for some generalized Weibull distribution, Reliability Engineering and System Safety, 84, 293-299.

Xie, M., Goh, T. N. & Tang, Y. (2004). On changing points of mean residual life and failure rate function for some generalized Weibull distributions, Reliability Engineering and System Safety, 84, 293–299.

Xie, M., Tang, Y. & Goh, T. N. (2002). A modified Weibull extension with bathtub-shaped failure rate function, Reliability Engineering and System Safety, 76, 279-285.

Yang Z. L., Xie, M. & Wong, A. C. M. (2007), A unified confidence interval for reliability related Weibull quantities, Journal of Statistical Computation and Simulation, 77, 365-378.

Yang, Z. L. & Lin, D. K. J. (2007). Improved maximum-likelihood estimation for the common shape parameter of several Weibull populations, Applied Stochastic Models in Business and Industry, 23, 373-383.

Yang, Z. L. & Xie, M. (2003). Efficient estimation on the Weibull shape parameter based on a modified profile likelihood, Journal of Statistical Computation and Simulation, 73, 115-123.

Yang, Z.L., See, S. P. & Xie, M. (2003). Transformation approaches for the construction of Weibull prediction interval, Computations Statistics and Data Analysis, 43, 357-368.

Zanakis, S. H. (1979). A simulation study of some simple estimators for the three-parameter Weibull distribution, Journal of Statistical Computation and Simulation, 9, 101-116.

Zellner, A. & Min, C. K. (1995). Gibbs sampler convergence criteria, Journal of the American Statistical Association, 90, 921-927.

Zhang, L. F., Xie, M. & Tang, L. C. (2006)[(1)]. Bias correction for the least squares estimator of the Weibull shape parameter with complete and censored data, Reliability Engineering and System Safety, 91, 930-939.

Zhang, L. F., Xie, M. & Tang, L. C. (2006)(2). Robust regression using probability plots for estimating the Weibull shape parameter, Quality and Reliability Engineering International, 22, 905-917.

Zhang, L. F., Xie, M. & Tang, L. C. (2007). A study of two estimation approaches for parameters of Weibull distribution based on WPP, Reliability Engineering and System Safety, 92, 360-368.

Zhang, T. L. & Xie. M. (2007). Failure data analysis with extended Weibull distribution, Communications in Statistics – Simulation and Computation, 36, 579-592.

Zhang, Y. & Meeker, W. Q. (2005). Bayesian life test planning for the Weibull distribution with given shape parameter, Metrika, 61, 237 – 249.

Zhao, X., Yu, Z. & Tong, H. (2008). A Bayesian approach to Weibull survival model for clinical randomized censoring trial based on MCMC simulation, The 2nd International Conference on Bioinformatics and Biomedical Engineering, 1181-1184.

Zheng, G. & Park, S. (2004). On the Fisher information in multiply censored and progressively censored data, Communications in Statistics – Theory and Methods, 33, 1821-1835.

Zheng, G. (2001). A characterization of the factorization of hazard function by the Fisher information under type II censoring with application to the Weibull family, Statistics & Probability Letters, 52, 249-253.

Zhou, Y. Y., Mi, J. & Guo, S. R. (2007). Estimation of parameters in logistic and log-logistic distribution with grouped data, Lifetime Data Analysis, 13, 421-429.