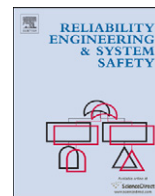




Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

A new approach to MLE of Weibull distribution with interval data

Zhibin Tan¹

Networks, Motorola Inc, Arlington Heights, Illinois 60004, USA

ARTICLE INFO

Article history:

Received 25 September 2007

Received in revised form

12 December 2007

Accepted 6 January 2008

Available online 3 May 2008

Keywords:

Weibull distribution

Exponential distribution

MLE

Censoring

Interval data

ABSTRACT

We study the two-parameter maximum likelihood estimation (MLE) problem for the Weibull distribution with consideration of interval data. Without interval data, the problem can be solved easily by regular MLE methods because the restricted MLE of the scale parameter β for a given shape parameter α has an analytical form, thus α can be efficiently solved from its profile score function by traditional numerical methods. In the presence of interval data, however, the analytical form for the restricted MLE of β does not exist and directly applying regular MLE methods could be less efficient and effective. To improve efficiency and effectiveness in handling interval data in the MLE problem, a new approach is developed in this paper. The new approach combines the Weibull-to-exponential transformation technique and the equivalent failure and lifetime technique. The concept of equivalence is developed to estimate exponential failure rates from uncertain data including interval data. Since the definition of equivalent failures and lifetimes follows EM algorithms, convergence of failure rate estimation by applying equivalent failures and lifetimes is mathematically proved. The new approach is demonstrated and validated through two published examples, and its performance in different conditions is studied by Monte Carlo simulations. It indicates that the profile score function for α has only one maximum in most cases. Such good characteristic enables efficient search for the optimal value of α .

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

The two-parameter Weibull distribution is widely used in reliability engineering and life data analysis because of its flexibility in modeling both increasing and decreasing failure rates. There are numerous papers and books dealing with various aspects of Weibull modeling, inference, applications, as well as parameter estimation. In general the parameter estimates of Weibull distribution can be found graphically via probability plotting paper, or analytically, either using least squares or maximum likelihood (ML) [1,2]. The probability plotting method is least mathematically intensive and can be best used for time to failure data with small size. The analytical least square method performs rank regression and is in essence the same methodology as the probability plotting method, except that it applies the principle of least squares to determine the line through the points, as opposed to just “eyeballing” it. Similar to the probability plotting method, the analytical least square method can be best used for time to failure data. Maximum likelihood estimation (MLE) works by developing a likelihood function based on the available data and finding the values of the

parameter estimates that maximize the likelihood function. The MLE method has many large sample properties that make it attractive for use. For instance, it is asymptotically consistent, efficient, and unbiased. With small data size, the MLE method can be badly biased. Study of and dealing with estimation bias can be found in [3–5]. Another advantage of the MLE method is that it can handle survival and interval data better than rank regression, particularly when dealing with a heavily censored data set with few exact failure times or when the censoring times are unevenly distributed.

In this paper we study MLE methods to estimate the shape parameter α and the scale parameter β in Weibull distribution with consideration of failure interval data. In general lifetime data can be put into several categories. Time to failure data, also called failure terminated data or complete data, give exact failure time. On the contrary, failure interval data only define a period in which a failure occurs. In this case, uncertainty exists because the exact failure time is unknown. Left censored data are a special case of failure interval data with the lower bound equal to zero. Though right censored data can also be treated as a special case of interval data with upper bound equal to infinity, it is usually more convenient to consider right censored data separately from failure interval data. In this paper we study MLE of Weibull distribution parameters from failure terminated data, right censored data, as well as failure interval data.

E-mail address: Zhibin.Tan@motorola.com¹ Zhibin Tan with Motorola.

Nomenclature $f(\cdot), F(\cdot), R(\cdot)$ pdf, cdf, and survival function α Weibull shape parameter β Weibull scale parameter λ exponential failure rate n number of failure terminated data x_i the i th failure terminated data, $i = 1, 2, \dots, n$ m number of right censored data c_i the i th right censored time, $i = 1, 2, \dots, m$ l number of failure intervals (a_i, b_i) the i th interval data, $a_i < b_i$, $i = 1, 2, \dots, l$ n^* equivalent failure x^* equivalent failure (test) time $\text{Log}(L)$ log likelihood function $G(\cdot)$ expectation of log likelihood function

When dealing with failure terminated and right censored data only, there exists an analytical form for the restricted MLE of β for a given α and α can be solved from the derived profile score function for α . In presence of interval data, regular MLE methods can still be applied, but there is no analytical solution for the restricted MLE of β for a given α . Though some iterative process can be developed from the regular MLE methods to obtain numerical estimate of β for a given α , convergence and effectiveness of the iterative process is not provable. Besides regular MLE methods, EM algorithms can be developed to deal with interval data. EM algorithm, first proposed by Dempster [6], is an elaborate technique for solving the MLE problem when the observed data is incomplete or has missing values. EM algorithm is remarkable because of its simplicity, generality, and convergence. The more detailed information regarding EM theorem is available in [7,8]. As we shall see in later sections, for Weibull distribution with interval data, there is no close form for the expectation of the log likelihood, and Monte Carlo simulation approach works but with certain limitations. Another alternative approach to dealing with interval data in estimating Weibull parameters is the genetic algorithm. The genetic algorithm is a subset of evolutionary algorithms that model biological processes to optimize highly complex cost functions. The method was developed by Holland [10] and finally popularized by Goldberg [11]. Advantages of a genetic algorithm include that it does not require derivative information and can optimize parameters with extremely complex cost surfaces [9]. It is very straightforward to develop a genetic algorithm to derive MLE of Weibull parameters by simply taking the log likelihood function as the cost function. Some concerns in using genetic algorithms include setting algorithm parameters like population size, ranges for estimated parameters, etc., which have impact on estimation performance.

In this paper we develop a new approach that is more efficient and effective in handling interval data in MLE of the Weibull parameters. The new approach combines the Weibull-to-exponential transformation technique and the equivalent failure and lifetime techniques. The concept of equivalence is developed to estimate exponential failure rates from uncertain data including interval data. Since the definition of equivalent failures and lifetimes follows EM algorithms, convergence of failure rate estimation by applying equivalent failures and lifetimes is mathematically proved. Equivalent failures and lifetimes for exponential distribution have been successfully applied to estimate component reliability from system masking data [14,15]. The Weibull-to-exponential transformation technique transforms Weibull random variables to exponential ones for which simple and analytical methods for statistical analysis and confidence intervals are available. This method should be used with caution because of the effect of mis-specification of the shape parameter, which has been studied in [12,13]. On one hand the effect on the quantities that only depend on the scale parameter could be very large even for a small mis-specification in the shape parameter; on the other hand, for quantities such as mean time to failure (MTTF), percentile lifetime, and mission

reliability that depend on both the scale and shape parameters, the joint effect could be small. In the proposed new approach, however, there is no such mis-specification effect because the shape and scale parameters are considered together. In fact, the scale parameter as a function of the shape parameter is just an intermediate result, and eventually only the profile score function for the shape parameter is considered.

This paper is organized as follows. In Section 2, we discuss the MLE problem for Weibull distribution, including different methods for complete data, right censored data, as well interval data. In Section 3, we establish some basis for the new approach. The new approach is explained and outlined in Section 4. Examples of applying the new approach are also presented in this section. Simulation study and some discussions are given in Section 5. Finally we conclude in Section 6.

2. Two-parameter MLE of Weibull distribution

In this paper we study MLE of Weibull distribution with the shape parameter α (alpha) and the scale parameter β (beta). The pdf, cdf, and survival function of Weibull distribution are in form of

$$f(x) = \frac{\alpha x^{\alpha-1}}{\beta^\alpha} e^{-(x/\beta)^\alpha} \quad (1)$$

$$F(x) = 1 - e^{-(x/\beta)^\alpha} \quad (2)$$

$$R(x) = e^{-(x/\beta)^\alpha} \quad (3)$$

2.1. The case of failure terminated data and right censored data

With n failure terminated data $x_1, x_2, \dots, x_n$ and m right censored data $c_1, c_2, \dots, c_m$, the log likelihood function is

$$\text{Log}(L) = \sum_{i=1}^n \log(f(x_i)) + \sum_{i=1}^m \log(R(c_i)) \quad (4)$$

$$= n \log(\alpha) + (\alpha - 1) \sum_{i=1}^n \log(x_i) - \alpha n \log(\beta) - \sum_{i=1}^n \left(\frac{x_i}{\beta}\right)^\alpha - \sum_{i=1}^m \left(\frac{c_i}{\beta}\right)^\alpha \quad (5)$$

By taking the derivatives of the log likelihood function with respect to α and β and setting them equal to zero, we have the following two equations:

$$\begin{aligned} \frac{\partial \text{Log}(L)}{\partial \alpha} &= \frac{n}{\alpha} + \sum_{i=1}^n \log(x_i) - n \log(\beta) - \sum_{i=1}^n \left(\frac{x_i}{\beta}\right)^\alpha \log\left(\frac{x_i}{\beta}\right) \\ &\quad - \sum_{i=1}^m \left(\frac{c_i}{\beta}\right)^\alpha \log\left(\frac{c_i}{\beta}\right) = 0 \end{aligned} \quad (6)$$

Download English Version:

<https://daneshyari.com/en/article/805977>

Download Persian Version:

<https://daneshyari.com/article/805977>

[Daneshyari.com](https://daneshyari.com)