Revista Colombiana de Estadística
Diciembre 2013, volumen 36, no. 2, pp. 337 a 356

Inference for the Weibull Distribution Based on Fuzzy Data

Inferencia para la distribución Weibull basada en datos difusos

Abbas Pak[1,a], Gholam Ali Parham[1,b], Mansour Saraj[2,c]

[1]Department of Statistics, Faculty of Mathematical Sciences and Computer, Shahid Chamran University of Ahvaz, Ahvaz, Iran

[2]Department of Mathematics, Faculty of Mathematical Sciences and Computer, Shahid Chamran University of Ahvaz, Ahvaz, Iran

Abstract

Classical estimation procedures for the parameters of Weibull distribution are based on precise data. It is usually assumed that observed data are precise real numbers. However, some collected data might be imprecise and are represented in the form of fuzzy numbers. Thus, it is necessary to generalize classical statistical estimation methods for real numbers to fuzzy numbers. In this paper, different methods of estimation are discussed for the parameters of Weibull distribution when the available data are in the form of fuzzy numbers. They include the maximum likelihood estimation, Bayesian estimation and method of moments. The estimation procedures are discussed in details and compared via Monte Carlo simulations in terms of their average biases and mean squared errors. Finally, a real data set taken from a light emitting diodes manufacturing process is investigated to illustrate the applicability of the proposed methods.

***Key words*:** Bayesian estimation, EM algorithm, Fuzzy data analysis, Maximum likelihood principle.

Resumen

Los procedimientos clásicos de estimación para los parámetros de la distribución Weibull se encuentran basados en datos precisos. Se asume usualmente que los datos observados son números reales precisos. Sin embargo, algunos datos recolectados podrían ser imprecisos y ser representados en la forma de números difusos. Por lo tanto, es necesario generalizar los métodos de estimación estadísticos clásicos de números reales a números difusos. En este artículo, diferentes métodos de estimación son discutidos para los

[a]PhD Student. E-mail: a-pak@scu.ac.ir
[b]Associate professor. E-mail: parham-g@scu.ac.ir
[c]Associate professor. E-mail: seraj.a@scu.ac.ir

parámetros de la distribución Weibull cuando los datos disponibles están en la forma de números difusos. Estos incluyen la estimación por máxima verosimilitud, la estimación Bayesiana y el método de momentos. Los procedimientos de estimación se discuten en detalle y se comparan vía simulaciones de Monte Carlo en términos de sesgos promedios y errores cuadráticos medios.

***Palabras clave*:** algoritmo EM, análisis de datos difusos, estimación Bayesiana, principio de máxima verosimilitud.

1. Introduction

The Weibull distribution was originally proposed by Waloddi Weibull back in 1937 for estimating machinery lifetime. Nowadays, the Weibull distribution is a broadly used in statistical model in engineering and life-time data analysis. The probability density function (pdf) and the cumulative distribution function (cdf) of a two-parameter Weibull random variable X can be written as

$$f(x;\alpha,\lambda) = \alpha\lambda x^{\alpha-1}\exp(-\lambda x^{\alpha}), \qquad x > 0 \tag{1}$$

and

$$F(x;\alpha,\lambda) = 1 - \exp(-\lambda x^{\alpha}), \qquad x > 0 \tag{2}$$

respectively, where $\lambda > 0$ is the scale and $\alpha > 0$ is the shape parameter. Several authors have addressed inferential issues for the parameters of a Weibull distribution; among others, Al-Baidhani & Sinclair (1987) compared the generalized least squares, maximum likelihood, and the two mixed method of estimating the parameters of a Weibull distribution. Qiao & Tsokos (1994) introduced an effective iterative procedure for the estimation. Watkins (1994) discussed maximum likelihood estimation for the two parameter Weibull distribution when the data for analysis contains both times to failure and censored times in operation. Marks (2005) considered the estimation of Weibull distribution parameters using the symmetrically located percentiles from a sample. Helu, Abu-Salih & Alkam (2010) proposed different methods of estimation for the parameters of Weibull distribution based on different sampling schemes-namely, simple random sample, ranked set sample, and modified ranked set sample.

The above inference techniques are limited to precise data. In real world situations, the data sometimes can not be measured and recorded precisely due to machine errors, human errors or some unexpected situations. The two types of such data namely, censored data and truncated data are widely used in practice. Censored data typically arise when an event of interest, such as a disease or a failure, is only partially observed, because information is gathered at certain examination times. Two usual models are random right-censorship and random interval-censorship. In the first case, the observations are assumed to be of the form $Y_i = \min(X_i, W_i)$, $i = 1, \ldots, n$, where the X_i are the (partially observed) survival times, and the W_i are the censoring times. In this model, both survival and censoring times are assumed to be random, and mutually independent. Estimating the parameters of Weibull distribution from such data have been considered by several authors. See, for example Ageel (2002), Balakrishnan & Kateri (2008), Nandi & Dewan (2010), Joarder, Krishna & Kundu (2011), Banerjee & Kundu (2012), and Lin, Chou & Huang (2012). In the case of so-called random interval censored data, the event is only known to happen between two random examination times. The observations are thus of the form (U_i, V_i), $i = 1, \dots, n$, and it is only known that $U_i \leq X_i \leq V_i$ for all i. Here again it is customary to assume independence between survival times X_i and censoring interval endpoints. Statistical analysis of Weibull distribution based on interval censored data has been discussed by Ng & Wang (2009) and Tan (2009), among others. Truncation is similar to but distinct from the concept of censoring. When the existence of the unseen "observation " is not known for observations that fall outside the particular range, the data that are observed are said to be truncated. Recently, Balakrishnan & Mitra (2012) developed the EM algorithm for the estimation of the parameters of the Weibull distribution based on left truncated and right censored data.

The problem addressed in this paper, is different from censoring and trancation. We are not concerned with imprecision arising from random inspection times, but with the situation in which the result of a random experiment is reported from the observer to the statistician with some imprecision, arising from its limited perception or recollection of the precise numerical values. For instance, the lifetime of some shaft may be reported as imprecise quantities such as: "about $1,000h$", "approximately $1,400h$", "almost between $1,000h$ and $1,200h$", "essentially less than $1,200h$", and so on. The lack of precision of such data can be described using fuzzy sets. The classical statistical estimation methods are not appropriate to deal with fuzzy sets. Therefore, the conventional procedures used for estimating the parameters of Weibull distribution will have to be adapted to the new situation. The main aim of this paper is to develop the inferential procedures for the two-parameter Weibull distribution when the available data are in the form of fuzzy numbers. In Section 2, we review the fundamental notation and basic definitions of fuzzy set theory. In Section 3, we first introduce a generalized likelihood function based on fuzzy data. We then discuss the computation of maximum likelihood estimates (MLEs) of the parameters α and λ by using the Newton-Raphson (NR) and Expectation Maximization (EM) algorithms, in Section 4. In Section 5, the Bayes estimates of the unknown parameters are obtained by using the approximation form of Tierney & Kadane (1986) under the assumption of Gamma priors. The estimation via method of moments is provided in Section 6. A Monte Carlo simulation study is presented in Section 7, which provides a comparison of all estimation procedures developed in this paper and one real data set is analyzed for illustrative purposes.

2. Basic Definition of Fuzzy Sets

To appreciate the nature of a fuzzy set, let us consider the following hypothetical example taken from Gertner & Zhu (1996). Consider an experiment characterized by a probability space $\mathcal{S} = (\mathcal{X}, \mathcal{B}_{\mathcal{X}}, P_\theta)$, where $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$ is a measurable space and P_θ belongs to a specified family of probability measures $\{P_\theta, \theta \in \Theta\}$ on $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$.

Any indicator function $I_A : \mathcal{X} \to \{0, 1\}$, defined by

$$I_A(x) = \begin{cases} 1 & x \in A, \\ 0 & x \notin A, \end{cases}$$

characterizes a crisp subset A in $\mathcal{X}$. For example, if $\mathcal{X} = \{x_i, i = 1, \dots, n\}$, represents all trees in a forest stand, then $A = \{x,\ x\text{'s age} \leq 40 \text{ yr}\}$ is its subset. So if tree x_3 is 27 yr old, $x_3 \in A$ and $I_A(x_3) = 1$; and if x_{239}'s age equals 56, $x_{239} \notin A$ and $I_A(x_{239}) = 0$. However, when referring to a "young tree", the set above described becomes a fuzzy set. Now relate each tree to its youthfulness by assigning a value between 1, representing absolutely young, and 0, representing absolutely not young, as the membership degree describing the subjective uncertainty of a tree being considered young. For instance, $\mu_{\text{young}}(x_3) = 0.9$, since x_3 will most likely be allocated into a younger class, whereas $\mu_{\text{young}}(x_{239}) = 0.49$ for x_{239} seems neither very young nor very old compared to other older trees in that stand. Thus, similar to crisp sets, a fuzzy subset $\widetilde{A}$ in $\mathcal{X}$ is characterized by a membership function $\mu_{\tilde{A}}(x)$ which associates with each point x in $\mathcal{X}$ a real number in the interval $[0, 1]$, with the value of $\mu_{\tilde{A}}(x)$ at x representing the "grade of membership"of x in $\widetilde{A}$. We hereafter assume that the sample space $\mathcal{X}$ is a set in a Euclidean space and $\mathcal{B}_{\mathcal{X}}$ is the smallest Borel $\sigma-$field on $\mathcal{X}$. A fuzzy event in $\mathcal{X}$ is a fuzzy subset $\tilde{A}$ of $\mathcal{X}$, whose membership function $\mu_{\tilde{A}}$ is Borel measurable. Many examples of fuzzy samples and observations appear in social and natural sciences. These occur when the linguistic concepts or propositions cannot be precisely defined, or accurate measurements of variables are not possible or necessary.

Example 1. An investigator is interested in analyzing the amount of an adverse substance extracted from a special brand of cigarettes. Assume that the investigator has not a mechanism of measurement which is sufficiently precise to determine exactly the amount of adverse substance of cigarettes, but rather he can only approximate them by means of imprecise observations, for instance, "The amount of adverse substance of cigarette is approximately 30 to 40 milligrams". A fuzzy approach lies in expressing the preceding observation as a fuzzy event $\widetilde{A}$ such as that defined, for instance, by the membership function (Figure 1).

$$\mu_{\widetilde{A}}(x) = \begin{cases} \frac{x-20}{10} & 20 \leq x \leq 30, \\ 1 & 30 \leq x \leq 40, \\ \frac{50-x}{10} & 40 \leq x \leq 50, \\ 0 & otherwise, \end{cases}$$

The notion of probability was extended to fuzzy events by Zadeh (1968) as follows.

Definition 1. Let $(\mathbb{R}^n, \mathcal{A}, P)$ be a probability space in which $\mathcal{A}$ is the $\sigma-$field of Borel sets in $\mathbb{R}^n$ and P is a probability measure over $\mathbb{R}^n$. Then, the probability of a fuzzy event $\widetilde{A}$ in $\mathbb{R}^n$ is defined by:

$$P(\widetilde{A}) = \int \mu_{\widetilde{A}}(\boldsymbol{x}) dP. \tag{3}$$

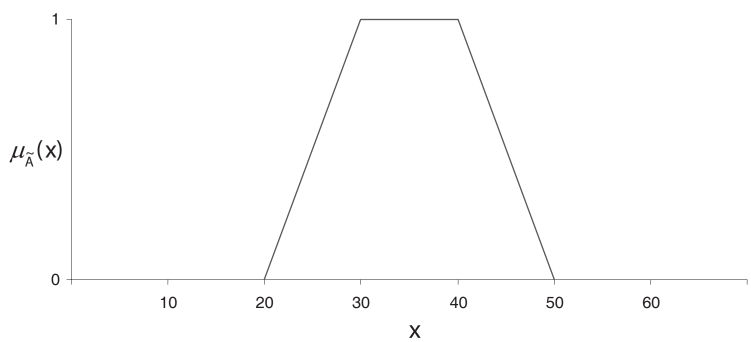

FIGURE 1: Fuzzy approach of the imprecise observation "approximately 30 to 40 ".

In particular, assume that P is the probability distribution of a continuous random variable $\boldsymbol{Y}$ with p.d.f. $g(\boldsymbol{Y})$. The conditional density of $\boldsymbol{Y}$ given $\widetilde{A}$ is given by

$$g(\boldsymbol{y} \mid \widetilde{A}) = \frac{\mu_{\widetilde{A}}(\boldsymbol{y})g(\boldsymbol{y})}{\int \mu_{\widetilde{A}}(\boldsymbol{u})g(\boldsymbol{u})d\boldsymbol{u}}. \tag{4}$$

The set consisting of all observable events from the experiment $\mathcal{S}$ determines a fuzzy information system (f.i.s.) associated with it, which is defined as follows.

Definition 2. (Tanaka ?). A fuzzy information system $\widetilde{\mathcal{S}}$ associated with the experiment $\mathcal{S}$ is a fuzzy partition $\mathcal{F} = \{\widetilde{x}_1, \dots, \widetilde{x}_K\}$ of $\mathcal{X}$, i.e., a set of K fuzzy events on $\mathcal{X}$ satisfying the orthogonality condition

$$\sum_{k=1}^{K} \mu_{\widetilde{x}_k}(x) = 1,$$

where $\mu_{\widetilde{x}_k}$ denotes the membership function of $\widetilde{x}_k$.

We now examine a brief example illustrating the preceding concept:

Example 2. To evaluate the problem of psychological depression in a population, there is no exact method that can measure and express the exact value for the severity of the disease in each person and, so measurement results may be reported by means of the following fuzzy observations: $\widetilde{x}_1$ = "approximately lower than 20", $\widetilde{x}_2$ ="approximately 25 to 30", $\widetilde{x}_3$ = "approximately 35 ", $\widetilde{x}_4$ ="approximately 40 to 45", $\widetilde{x}_5$ = "approximately 50", $\widetilde{x}_6$ = "approximately higher than 55", which are characterized by the membership functions

$$\mu_{\widetilde{x}_1}(x) = \begin{cases} 1 & x \leq 20, \\ \frac{25-x}{5} & 20 \leq x \leq 25, \\ 0 & otherwise, \end{cases} \qquad \mu_{\widetilde{x}_2}(x) = \begin{cases} \frac{x-20}{5} & 20 \leq x \leq 25, \\ 1 & 25 \leq x \leq 30, \\ \frac{35-x}{5} & 30 \leq x \leq 35, \\ 0 & otherwise, \end{cases}$$

$$\mu_{\widetilde{x}_3}(x) = \begin{cases} \frac{x-30}{5} & 30 \le x \le 35, \\ \frac{40-x}{5} & 35 \le x \le 40, \\ 0 & otherwise, \end{cases} \qquad \mu_{\widetilde{x}_4}(x) = \begin{cases} \frac{x-35}{5} & 35 \le x \le 40, \\ 1 & 40 \le x \le 45, \\ \frac{50-x}{5} & 45 \le x \le 50, \\ 0 & otherwise, \end{cases}$$

$$\mu_{\widetilde{x}_5}(x) = \begin{cases} \frac{x-45}{5} & 45 \le x \le 50, \\ \frac{55-x}{5} & 50 \le x \le 55, \\ 0 & otherwise, \end{cases} \qquad \mu_{\widetilde{x}_6}(x) = \begin{cases} \frac{x-50}{5} & 50 \le x \le 55, \\ 1 & x \ge 55, \\ 0 & otherwise, \end{cases}$$

respectively, (see Fig.2). Clearly, a f.i.s. $\widetilde{\mathcal{S}} = \{\widetilde{x}_1, ..., \widetilde{x}_7\}$ can be immediately constructed by defining $\mu_{\widetilde{x}_7} = 1 - \sum_{i=1}^{6} \mu_{\widetilde{x}_i})$

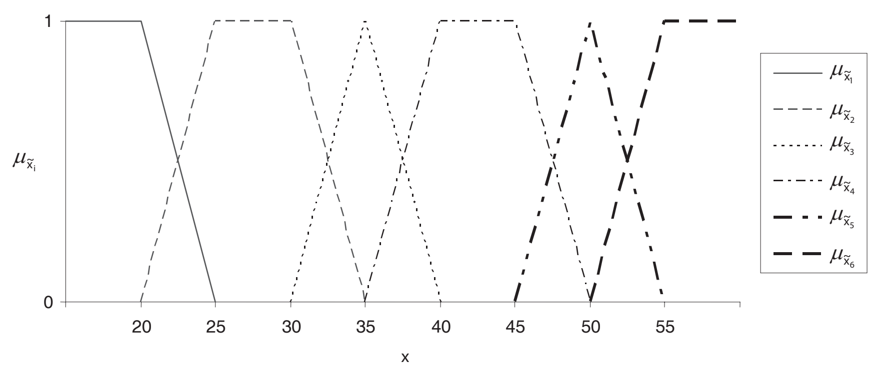

FIGURE 2: Membership functions of the fuzzy observations $\widetilde{x}_1$, $\widetilde{x}_2$, $\widetilde{x}_3$, $\widetilde{x}_4$, $\widetilde{x}_5$ and $\widetilde{x}_6$.

For more details about the membership functions and probability measures of fuzzy sets, one can refer to Singpurwalla & Booker (2004).

In order to model imprecise data, a generalization of real numbers is necessary. These data can be represented by fuzzy numbers. A fuzzy number is a subset, denoted by $\widetilde{x}$, of the set of real numbers (denoted by $\mathbb{R}$) and is characterized by the so called membership function $\mu_{\widetilde{x}}(.)$. Fuzzy numbers satisfy the following constraints (see Dubois & Prade (1980)):

(1) $\mu_{\widetilde{x}} : \mathbb{R} \longrightarrow [0, 1]$ is Borel-measurable;

(2) $\exists x_0 \in \mathbb{R} : \mu_{\widetilde{x}}(x_0) = 1$;

(3) The so-called $\lambda-$cuts $(0 < \lambda \le 1)$, defined as $B_\lambda(\widetilde{x}) = \{x \in \mathbb{R} : \mu_{\widetilde{x}}(x) \ge \lambda\}$, are all closed intervals, i.e., $B_\lambda(\widetilde{x}) = [a_\lambda, b_\lambda], \ \forall \lambda \in (0, 1]$.

With the definition of a fuzzy number given above, an exact (non-fuzzy) number can be treated as a special case of a fuzzy number. For a non-fuzzy real observation $x_0 \in \mathbb{R}$, its corresponding membership function is $\mu_{x_0}(x_0) = 1$.

Among the various types of fuzzy numbers, the triangular and trapezoidal fuzzy numbers are most convenient and useful in describing fuzzy data. For triangular membership functions, the triangular fuzzy number can be defined as $\widetilde{x} = (a, b, c)$ and its membership function is defined by the following expression:

$$\mu_{\widetilde{x}}(x) = \begin{cases} \frac{x-a}{b-a} & a \leq x \leq b, \\ \frac{c-x}{c-b} & b \leq x \leq c, \\ 0 & otherwise. \end{cases}$$

The trapezoidal fuzzy number can be defined as $\widetilde{x} = (a, b, c, d)$ with membership function

$$\mu_{\widetilde{x}}(x) = \begin{cases} \frac{x-a}{b-a} & a \leq x \leq b, \\ 1 & b \leq x \leq c, \\ \frac{d-x}{d-c} & c \leq x \leq d, \\ 0 & otherwise. \end{cases}$$

3. Fuzzy Data and the Likelihood Function

Suppose that $X_1, \ldots, X_n$ is a random sample of size n from Weibull distribution with pdf given by (1). Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ denotes the corresponding random vector. If a realization $\boldsymbol{x} = (x_1, \ldots, x_n)$ of $\boldsymbol{X}$ was known exactly, we could obtain the complete-data likelihood function as

$$L(\alpha, \lambda; \boldsymbol{x}) = \alpha^n \lambda^n \exp(-\lambda \sum_{i=1}^{n} x_i^{\alpha}) \prod_{i=1}^{n} x_i^{\alpha-1} \tag{5}$$

Now consider the problem where $\boldsymbol{x}$ is not observed precisely and only partial information about $\boldsymbol{x}$ is available in the form of a fuzzy subset $\widetilde{\boldsymbol{x}}$ with the Borel measurable membership function $\mu_{\widetilde{\boldsymbol{x}}}(\boldsymbol{x})$. In this setting, the fuzzy observation $\widetilde{\boldsymbol{x}}$ can be understood as encoding the observer's partial knowledge about the realization $\boldsymbol{x}$ of random vector $\boldsymbol{X}$, and the membership function $\mu_{\widetilde{\boldsymbol{x}}}$ is seen as a possibility distribution interpreted as a soft constraint on the unknown quantity $\boldsymbol{x}$. The fuzzy set $\widetilde{\boldsymbol{x}}$ can be considered to be generated by a two-step process:

1. A realization $\boldsymbol{x}$ is drawn from $\boldsymbol{X}$;
2. The observer encodes his/her partial knowledge of $\boldsymbol{x}$ in the form of a possibility distribution $\mu_{\widetilde{\boldsymbol{x}}}$.

It must be noted that, in this model, only step 1 is considered to be a random experiment. Step 2 implies gathering information about $\boldsymbol{x}$ and modeling this information as a possibility distribution.

Example 3. Consider a life-testing experiment in which n identical ball bearings are placed on test, and we are interested in the lifetime of these ball bearings. The unknown lifetime x_i of ball bearing i may be regarded as a realization of a random variable X_i induced by random sampling from a total population of ball bearings. In practice, however, measuring the lifetime of a ball bearing may not yield an exact result. A ball bearing may work perfectly over a certain period but be braking for some time, and finally be unusable at a certain time. Assume that two intervals are determined for the lifetime of each ball bearing i as follows:

- an interval $[a_i, d_i]$ certainly containing x_i;
- an interval $[b_i, c_i]$ containing highly plausible values for x_i.

This information may be encoded as a trapeozoidal fuzzy number $\widetilde{x}_i = (a_i, b_i, c_i, d_i)$ with support $[a_i, d_i]$ and core $[b_i, c_i]$, interpreted as a possibility distribution constraining the unknown value x_i. Information about $\boldsymbol{x}$ may be represented by the joint possibility distribution

$$\mu_{\widetilde{\boldsymbol{x}}}(\boldsymbol{x}) = \mu_{\widetilde{x}_1}(x_1) \times \dots \times \mu_{\widetilde{x}_n}(x_n). \tag{6}$$

Once $\widetilde{\boldsymbol{x}}$ is given, and assuming its membership function to be the Borel measurable, we can compute its probability according to Zadeh's definition of the probability of a fuzzy event. By using the expression (3), the observed-data likelihood function can then be obtained as

$$L_O(\alpha, \lambda; \widetilde{\boldsymbol{x}}) = P(\widetilde{\boldsymbol{x}}; \alpha, \lambda) = \int f(\boldsymbol{x}; \alpha, \lambda) \mu_{\widetilde{\boldsymbol{x}}}(\boldsymbol{x}) d\boldsymbol{x}. \tag{7}$$

Since the data vector $\boldsymbol{x}$ is a realization of an independent identically distributed (i.i.d.) random vector $\boldsymbol{X}$, and assuming the joint membership function $\mu_{\widetilde{\boldsymbol{x}}}(\boldsymbol{x})$ to be decomposable as in (6), the likelihood function (7) can be written as:

$$L_O(\alpha, \lambda; \widetilde{\boldsymbol{x}}) = \prod_{i=1}^{n} \int \alpha \lambda x^{\alpha-1} \exp(-\lambda x^{\alpha}) \mu_{\widetilde{x}_i}(x) dx, \tag{8}$$

and the observed-data log likelihood is

$$\begin{aligned} L^*(\alpha, \lambda; \widetilde{\boldsymbol{x}}) &= \log L_O(\alpha, \lambda; \widetilde{\boldsymbol{x}}) \\ &= n(\log \alpha + \log \lambda) + \sum_{i=1}^{n} \log \int x^{\alpha-1} \exp(-\lambda x^{\alpha}) \mu_{\widetilde{x}_i}(x) dx. \end{aligned} \tag{9}$$

4. Maximum Likelihood Estimation

The idea behind maximum likelihood parameter estimation is to determine the parameters that maximize the probability (likelihood) of the sample data. From a statistical point of view, the method of maximum likelihood is considered to be more robust and yields estimators with good statistical properties. In other words, maximum likelihood methods are versatile and apply to most models and to different types of data. The maximum likelihood estimate of the parameters α

and λ can be obtained by maximizing the log-likelihood $L^*(\alpha,\lambda;\widetilde{\boldsymbol{x}})$. Equating the partial derivatives of the log-likelihood (9) with respect to α and λ to zero, the resulting two equations are:

$$\frac{\partial}{\partial\alpha}L^*(\alpha,\lambda;\widetilde{\boldsymbol{x}})=\frac{n}{\alpha}+\sum_{i=1}^{n}\frac{\int(x^{\alpha-1}-\lambda x^{2\alpha-1})\log x\exp(-\lambda x^{\alpha})\,\mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha-1}\exp(-\lambda x^{\alpha})\,\mu_{\widetilde{x}_i}(x)dx}=0 \quad (10)$$

and

$$\frac{\partial}{\partial\lambda}L^*(\alpha,\lambda;\widetilde{\boldsymbol{x}})=\frac{n}{\lambda}-\sum_{i=1}^{n}\frac{\int x^{2\alpha-1}\exp(-\lambda x^{\alpha})\,\mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha-1}\exp(-\lambda x^{\alpha})\,\mu_{\widetilde{x}_i}(x)dx}=0. \quad (11)$$

Since there are no closed form of the solutions to the likelihood equations (10) and (11), an iterative numerical search can be used to obtain the MLEs. In the following, we describe the NR method and the EM algorithm to determine the MLEs of the parameters α and λ.

4.1. NR Algorithm

NR algorithm is a direct approach for estimating the relevant parameters in a likelihood function. In this algorithm, the solution of the likelihood equation is obtained through an iterative procedure. Let $\boldsymbol{\theta}=(\alpha,\lambda)^T$ be the parameter vector. Then, at the $(h+1)$th step of iteration process, the updated parameter is obtained as

$$\boldsymbol{\theta}^{(h+1)}=\boldsymbol{\theta}^{(h)}-\left[\frac{\partial^2L^*(\boldsymbol{\theta};\widetilde{\boldsymbol{x}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\,|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(h)}}\right]^{-1}\left[\frac{\partial L^*(\boldsymbol{\theta};\widetilde{\boldsymbol{x}})}{\partial\boldsymbol{\theta}}\,|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(h)}}\right] \quad (12)$$

where

$$\frac{\partial L^*(\boldsymbol{\theta};\widetilde{\boldsymbol{x}})}{\partial\boldsymbol{\theta}}=\begin{pmatrix}\frac{\partial L^*(\alpha,\lambda;\widetilde{\boldsymbol{x}})}{\partial\alpha}\\ \frac{\partial L^*(\alpha,\lambda;\widetilde{\boldsymbol{x}})}{\partial\lambda}\end{pmatrix}$$

and

$$\frac{\partial^2L^*(\boldsymbol{\theta};\widetilde{\boldsymbol{x}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}=\begin{pmatrix}\frac{\partial^2L^*(\alpha,\lambda;\widetilde{\boldsymbol{x}})}{\partial\alpha^2} & \frac{\partial^2L^*(\alpha,\lambda;\widetilde{\boldsymbol{x}})}{\partial\lambda\partial\alpha}\\ \frac{\partial^2L^*(\alpha,\lambda;\widetilde{\boldsymbol{x}})}{\partial\lambda\partial\alpha} & \frac{\partial^2L^*(\alpha,\lambda;\widetilde{\boldsymbol{x}})}{\partial\lambda^2}\end{pmatrix}$$

The second-order derivatives of the log-likelihood with respect to the parameters, required for proceeding with the NR method, are obtained as follows.

$$\begin{aligned}\frac{\partial^2}{\partial\alpha^2}L^*(\alpha,\lambda;\widetilde{\boldsymbol{x}}) &= -\frac{n}{\alpha^2}\\ &\quad+\sum_{i=1}^{n}\Big\{\frac{\int(\lambda^2x^{3\alpha-1}-\lambda x^{2\alpha-1})(\log x)^2\exp(-\lambda x^{\alpha})\,\mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha-1}\exp(-\lambda x^{\alpha})\,\mu_{\widetilde{x}_i}(x)dx}\\ &\qquad+\frac{\int(x^{\alpha-1}-2\lambda x^{2\alpha-1})(\log x)^2\exp(-\lambda x^{\alpha})\,\mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha-1}\exp(-\lambda x^{\alpha})\,\mu_{\widetilde{x}_i}(x)dx}\Big\}\\ &\quad-\sum_{i=1}^{n}\left[\frac{\int(x^{\alpha-1}-\lambda x^{2\alpha-1})\log x\exp(-\lambda x^{\alpha})\,\mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha-1}\exp(-\lambda x^{\alpha})\,\mu_{\widetilde{x}_i}(x)dx}\right]^2\end{aligned}$$

$$\begin{aligned}\frac{\partial^2}{\partial\lambda^2}L^*(\alpha,\lambda;\widetilde{\boldsymbol{x}}) &= -\frac{n}{\lambda^2}+\sum_{i=1}^{n}\{\frac{\int x^{3\alpha-1}\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha-1}\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}\\ &\quad -[\frac{\int x^{2\alpha-1}\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha-1}\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}]^2\},\end{aligned}$$

$$\begin{aligned}\frac{\partial^2}{\partial\lambda\partial\alpha}L^*(\alpha,\lambda;\widetilde{\boldsymbol{x}}) &= -\sum_{i=1}^{n}\frac{\int(2x^{2\alpha-1}-\lambda x^{3\alpha-1})\log x\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha-1}\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}\\ &\quad +\sum_{i=1}^{n}\{\frac{\int(1-\lambda x^{\alpha})x^{\alpha-1}\log x\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha-1}\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}\\ &\quad \times\frac{\int x^{2\alpha-1}\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha-1}\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}\}\end{aligned}$$

The iteration process then continues until convergence, i.e., until $\| \boldsymbol{\theta}^{(h+1)}-\boldsymbol{\theta}^{(h)} \|<\varepsilon$, for some pre-fixed $\varepsilon>0$. The maximum likelihood estimate of (α,λ) via NR algorithm is thereafter refereed as "$(\hat{\alpha}_{NR},\hat{\lambda}_{NR})$" in this paper.

It should be pointed out that the second-order derivatives of the log-likelihood are required at every iteration in the NR method. Sometimes the calculation of the derivatives based on fuzzy data can be rather tedious. Another viable alternative to the NR algorithm is the well-known EM algorithm. In the following, we discuss how that can be used to determine the MLEs in this case.

4.2. EM Algorithm

The EM algorithm is a broadly applicable approach to the iterative computation of maximum likelihood estimates and useful in a variety of incomplete-data problems. Since the observed fuzzy data $\widetilde{\boldsymbol{x}}$ can be seen as an incomplete specification of a complete data vector $\boldsymbol{x}$, the EM algorithm is applicable to obtain the maximum likelihood estimates of the unknown parameters. In the following, we use the fuzzy EM algorithm (see Denoeux (2011)) to determine the MLEs of α and λ.

From the Eq. (5), the log-likelihood function for the complete data vector $\boldsymbol{x}$ becomes:

$$\log L(\alpha,\lambda;\boldsymbol{x})=n\log\alpha+n\log\lambda+(\alpha-1)\sum_{i=1}^{n}\log x_i-\lambda\sum_{i=1}^{n}x_i^{\alpha} \tag{13}$$

Taking the derivative with respect to α and λ, respectively, on (13), the following likelihood equations are obtained:

$$\frac{n}{\lambda}=\sum_{i=1}^{n}x_i^{\alpha} \tag{14}$$

and

$$\frac{n}{\alpha}=\lambda\sum_{i=1}^{n}x_i^{\alpha}\log x_i-\sum_{i=1}^{n}\log x_i \tag{15}$$

Therefore the EM algorithm is given by the following iterative process:

1. Given starting values of α and λ, say $\alpha^{(0)}$ and $\lambda^{(0)}$ and set $h = 0$.
2. In the $(h+1)$th iteration,

   - The E-step requires to compute the following conditional expectations using the expression (4):

$$E_{1i} = E_{\alpha^{(h)},\lambda^{(h)}}(X^{\alpha} \mid \widetilde{x}_i) = \frac{\int x^{2\alpha^{(h)}-1} \exp\left(-\lambda^{(h)} x^{\alpha^{(h)}}\right) \mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha^{(h)}-1} \exp\left(-\lambda^{(h)} x^{\alpha^{(h)}}\right) \mu_{\widetilde{x}_i}(x)dx}$$

$$E_{2i} = E_{\alpha^{(h)},\lambda^{(h)}}(\log X \mid \widetilde{x}_i) = \frac{\int x^{\alpha^{(h)}-1} \log x \exp\left(-\lambda^{(h)} x^{\alpha^{(h)}}\right) \mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha^{(h)}-1} \exp\left(-\lambda^{(h)} x^{\alpha^{(h)}}\right) \mu_{\widetilde{x}_i}(x)dx}$$

$$\begin{aligned} E_{3i} &= E_{\alpha^{(h)},\lambda^{(h)}}(X^{\alpha} \log X \mid \widetilde{x}_i) \\ &= \frac{\int x^{2\alpha^{(h)}-1} \log x \exp\left(-\lambda^{(h)} x^{\alpha^{(h)}}\right) \mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha^{(h)}-1} \exp\left(-\lambda^{(h)} x^{\alpha^{(h)}}\right) \mu_{\widetilde{x}_i}(x)dx} \end{aligned}$$

   and the likelihood equations (14) and (15) are replaced by

$$\frac{n}{\lambda} = \sum_{i=1}^{n} E_{1i}, \tag{16}$$

   and

$$\frac{n}{\alpha} = \lambda \sum_{i=1}^{n} \left[E_{3i} - E_{2i}\right]. \tag{17}$$

   - The M-step requires to solve the Eqs. (16) and (17), and obtain the next values, $\lambda^{(h+1)}$ and $\alpha^{(h+1)}$, of λ and α, respectively, as follows:

$$\lambda^{(h+1)} = \frac{n}{\sum\limits_{i=1}^{n} E_{1i}}$$

$$\alpha^{(h+1)} = \left\{ \frac{1}{n} \lambda^{(h+1)} \sum_{i=1}^{n} \left[E_{3i} - E_{2i}\right] \right\}^{-1}$$

3. Checking convergence, if the convergence occurs then the current $\alpha^{(h+1)}$ and $\lambda^{(h+1)}$ are the maximum likelihood estimates of α and λ via EM algorithm; otherwise, set $h = h + 1$ and go to Step 2.

The maximum likelihood estimate of (α, λ) via EM algorithm is thereafter refereed as "$(\hat{\alpha}_{EM}, \hat{\lambda}_{EM})$" in this paper.

5. Bayesian Estimation

In recent decades, the Bayes viewpoint, as a powerful and valid alternative to traditional statistical perspectives, has received frequent attention for statistical inference. In this section, we consider the Bayesian estimation under the assumptions that α and λ have independent gamma priors with the pdfs

$$\pi_1(\alpha) = \frac{d^c}{\Gamma(c)} \alpha^{c-1} \exp(-\alpha d), \qquad \alpha > 0 \tag{18}$$

and

$$\pi_2(\lambda) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp(-\lambda b), \qquad \lambda > 0 \tag{19}$$

with the parameters $\alpha \sim Gamma(c, d)$ and $\lambda \sim Gamma(a, b)$. Based on the above priors, the joint posterior density function of α and λ given the data can be written as follows:

$$\pi(\alpha, \lambda \mid \widetilde{\boldsymbol{x}}) = \frac{\pi_1(\alpha)\pi_2(\lambda)\ell(\alpha, \lambda; \widetilde{\boldsymbol{x}})}{\int_0^\infty \int_0^\infty \pi_1(\alpha)\pi_2(\lambda)\ell(\alpha, \lambda; \widetilde{\boldsymbol{x}}) d\alpha d\lambda} \tag{20}$$

where

$$\ell(\alpha, \lambda; \widetilde{\boldsymbol{x}}) = \alpha^{(n+c-1)} \lambda^{(n+a-1)} \exp(-\alpha d) \exp(-\lambda b) \prod_{i=1}^{n} \int x^{\alpha-1} \exp\left(-\lambda x^\alpha\right) \mu_{\widetilde{x}_i}(x) dx$$

is the likelihood function based on the fuzzy sample $\widetilde{\boldsymbol{x}}$. Then, under a squared error loss function, the Bayes estimate of any function of α and λ, say $g(\alpha, \lambda)$, is

$$\begin{aligned} E(g(\alpha, \lambda) \mid \widetilde{\boldsymbol{x}}) &= \frac{\int_0^\infty \int_0^\infty g(\alpha, \lambda)\pi_1(\alpha)\pi_2(\lambda)\ell(\alpha, \lambda; \widetilde{\boldsymbol{x}}) d\alpha d\lambda}{\int_0^\infty \int_0^\infty \pi_1(\alpha)\pi_2(\lambda)\ell(\alpha, \lambda; \widetilde{\boldsymbol{x}}) d\alpha d\lambda} \\ &= \frac{\int_0^\infty \int_0^\infty g(\alpha, \lambda) e^{Q(\alpha, \lambda)} d\alpha d\lambda}{\int_0^\infty \int_0^\infty e^{Q(\alpha, \lambda)} d\alpha d\lambda} \end{aligned} \tag{21}$$

where $Q(\alpha, \lambda) = \ln[\pi_1(\alpha)\pi_2(\lambda)] + \ln \ell(\alpha, \lambda; \widetilde{\boldsymbol{x}}) \equiv \rho(\alpha, \lambda) + L(\alpha, \lambda)$. Note that Eq. (21) cannot be obtained analytically; therefore, in the following we adopt Tierney and Kadane's approximation for computing the Bayes estimates.

Setting $H(\alpha, \lambda) = Q(\alpha, \lambda)/n$ and $H^*(\alpha, \lambda) = \left[\ln g(\alpha, \lambda) + Q(\alpha, \lambda)\right]/n$, the expression in (21) can be reexpressed as

$$E(g(\alpha, \lambda) \mid \widetilde{\boldsymbol{x}}) = \frac{\int_0^\infty \int_0^\infty e^{nH^*(\alpha, \lambda)} d\alpha d\lambda}{\int_0^\infty \int_0^\infty e^{nH(\alpha, \lambda)} d\alpha d\lambda} \tag{22}$$

Following Tierney & Kadane (1986), Eq. (22) can be approximated as the following form:

$$\hat{g}_{BT}(\alpha,\lambda)=\left[\frac{\det\Sigma^*}{\det\Sigma}\right]^{1/2}\exp\left\{n\left[H^*(\bar{\alpha}^*,\bar{\lambda}^*)-H(\bar{\alpha},\bar{\lambda})\right]\right\} \tag{23}$$

where $(\bar{\alpha}^*,\bar{\lambda}^*)$ and $(\bar{\alpha},\bar{\lambda})$ maximize $H^*(\alpha,\lambda)$ and $H(\alpha,\lambda)$, respectively, and Σ^* and Σ are the negatives of the inverse Hessians of $H^*(\alpha,\lambda)$ and $H(\alpha,\lambda)$ at $(\bar{\alpha}^*,\bar{\lambda}^*)$ and $(\bar{\alpha},\bar{\lambda})$, respectively.

In our case, we have

$$\begin{aligned}H(\alpha,\lambda) &= \frac{1}{n}\{k+(n+c-1)\log\alpha+(n+a-1)\log\lambda-\alpha d\\ &\quad -\lambda b+\sum_{i=1}^{n}\log\int x^{\alpha-1}\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx\}.\end{aligned}$$

where k is a constant; therefore, $(\bar{\alpha},\bar{\lambda})$ can be obtained by solving the following two equations

$$\begin{aligned}\frac{\partial}{\partial\alpha}H(\alpha,\lambda) &= \frac{1}{n}\{\frac{n+c-1}{\alpha}-d\\ &\quad +\sum_{i=1}^{n}\frac{\int(x^{\alpha-1}-\lambda x^{2\alpha-1})\log x\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha-1}\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}\}\end{aligned}$$

$$\frac{\partial}{\partial\lambda}H(\alpha,\lambda)=\frac{1}{n}\left\{\frac{n+a-1}{\lambda}-b-\sum_{i=1}^{n}\frac{\int x^{2\alpha-1}\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha-1}\exp\left(-\lambda x^{\alpha}\right)\mu_{\widetilde{x}_i}(x)dx}\right\}$$

and, from the second derivatives of $H(\alpha,\lambda)$, the determinant of the negative of the inverse Hessian of $H(\alpha,\lambda)$ at $(\bar{\alpha},\bar{\lambda})$ is given by

$$\det\Sigma=(H_{11}H_{22}-H_{12}^2)^{-1}$$

where

$$\begin{aligned}H_{11} &= \frac{1}{n}\{-\frac{n+c-1}{\bar{\alpha}^2}\\ &\quad +\sum_{i=1}^{n}(\frac{\int(\bar{\lambda}^2x^{3\bar{\alpha}-1}-\bar{\lambda}x^{2\bar{\alpha}-1})(\log x)^2\exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right)\mu_{\widetilde{x}_i}(x)dx}{\int x^{\bar{\alpha}-1}\exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right)\mu_{\widetilde{x}_i}(x)dx}\\ &\quad +\frac{\int(x^{\bar{\alpha}-1}-2\bar{\lambda}x^{2\bar{\alpha}-1})(\log x)^2\exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right)\mu_{\widetilde{x}_i}(x)dx}{\int x^{\bar{\alpha}-1}\exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right)\mu_{\widetilde{x}_i}(x)dx})\\ &\quad -\sum_{i=1}^{n}[\frac{\int(x^{\bar{\alpha}-1}-\bar{\lambda}x^{2\bar{\alpha}-1})\log x\exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right)\mu_{\widetilde{x}_i}(x)dx}{\int x^{\bar{\alpha}-1}\exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right)\mu_{\widetilde{x}_i}(x)dx}]^2\ \}\end{aligned}$$

$$\begin{aligned}H_{22} &= \frac{1}{n}\{-\frac{n+a-1}{\bar{\lambda}^2}+\sum_{i=1}^{n}(\frac{\int x^{3\bar{\alpha}-1}\exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right)\mu_{\widetilde{x}_i}(x)dx}{\int x^{\bar{\alpha}-1}\exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right)\mu_{\widetilde{x}_i}(x)dx}\\ &\quad -[\frac{\int x^{2\bar{\alpha}-1}\exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right)\mu_{\widetilde{x}_i}(x)dx}{\int x^{\bar{\alpha}-1}\exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right)\mu_{\widetilde{x}_i}(x)dx}]^2)\}\end{aligned}$$

$$
\begin{aligned}
H_{12} \quad = \quad & \frac{1}{n}\{-\sum_{i=1}^{n} \frac{\int (2x^{2\bar{\alpha}-1} - \bar{\lambda}x^{3\bar{\alpha}-1}) \log x \exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right) \mu_{\widetilde{x}_i}(x)dx}{\int x^{\bar{\alpha}-1} \exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right) \mu_{\widetilde{x}_i}(x)dx} \\
& + \sum_{i=1}^{n} \Big(\frac{\int (1-\bar{\lambda}x^{\bar{\alpha}})x^{\bar{\alpha}-1} \log x \exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right) \mu_{\widetilde{x}_i}(x)dx}{\int x^{\bar{\alpha}-1} \exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right) \mu_{\widetilde{x}_i}(x)dx} \\
& \times \frac{\int x^{2\bar{\alpha}-1} \exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right) \mu_{\widetilde{x}_i}(x)dx}{\int x^{\bar{\alpha}-1} \exp\left(-\bar{\lambda}x^{\bar{\alpha}}\right) \mu_{\widetilde{x}_i}(x)dx} \Big)\}
\end{aligned}
$$

Now, following the same arguments with $g(\alpha, \lambda) = \alpha$ and λ, respectively, in $H^*(\alpha, \lambda)$, $\hat{\alpha}_{BT}$ and $\hat{\lambda}_{BT}$ in Equation (23) can then be obtained in a straightforward manner.

6. Method of Moments

It is well-known that the kth moment of the Weibull distribution with pdf (1) is

$$E(X^k) = \lambda^{-\frac{k}{\alpha}} \Gamma(1 + \frac{k}{\alpha})$$

where $\Gamma(.)$ is the complete Gamma function.

By equating the first and the second sample moments to the corresponding population moments, the following equations can be used to find the estimates of moment method.

$$\lambda^{-\frac{1}{\alpha}} \Gamma(1 + \frac{1}{\alpha}) = \frac{1}{n} \sum_{i=1}^{n} E_{\alpha,\lambda}(X \mid \widetilde{x}_i) \tag{24}$$

$$\lambda^{-\frac{2}{\alpha}} \Gamma(1 + \frac{2}{\alpha}) = \frac{1}{n} \sum_{i=1}^{n} E_{\alpha,\lambda}(X^2 \mid \widetilde{x}_i) \tag{25}$$

Since the closed form of the solutions to Eqs. (24) and (25) could not be obtained, an iterative numerical process to obtain the parameter estimates is described as follows:

1. Let the initial estimates of α and λ, say $\alpha^{(0)}$ and $\lambda^{(0)}$ with $h = 0$.
2. In the $(h+1)$th iteration, we first compute

$$E_{\alpha^{(h)},\lambda^{(h)}}(X^r \mid \widetilde{x}_i) = \frac{\int x^{\alpha^{(h)}+r-1} \exp\left(-\lambda^{(h)} x^{\alpha^{(h)}}\right) \mu_{\widetilde{x}_i}(x)dx}{\int x^{\alpha^{(h)}-1} \exp\left(-\lambda^{(h)} x^{\alpha^{(h)}}\right) \mu_{\widetilde{x}_i}(x)dx}, \qquad r = 1, 2.$$

3. Based on equations (24) and (25), solve the following equation for α

$$\frac{\left[\sum_{i=1}^{n} E_{\alpha^{(h)},\lambda^{(h)}}(X \mid \widetilde{x}_i)\right]^2}{n\left[\sum_{i=1}^{n} E_{\alpha^{(h)},\lambda^{(h)}}(X^2 \mid \widetilde{x}_i)\right]} = \frac{\left[\Gamma(1+\frac{1}{\alpha})\right]^2}{\left[\Gamma(1+\frac{2}{\alpha})\right]}$$

to obtain the solution as $\alpha^{(h+1)}$.

4. The solution for λ, say $\lambda^{(h+1)}$, is obtained through the following equation

$$\lambda^{(h+1)} = \left\{ \frac{n\Gamma(1 + (1/\alpha^{(h+1)}))}{\sum_{i=1}^{n} E_{\alpha^{(h)},\lambda^{(h)}}(X \mid \widetilde{x}_i)} \right\}^{\alpha^{(h+1)}}$$

5. Setting $h = h + 1$, repeat steps 2 to 4 until convergence occurs and denote the method of moment estimates as $\hat{\alpha}_M$ and $\hat{\lambda}_M$.

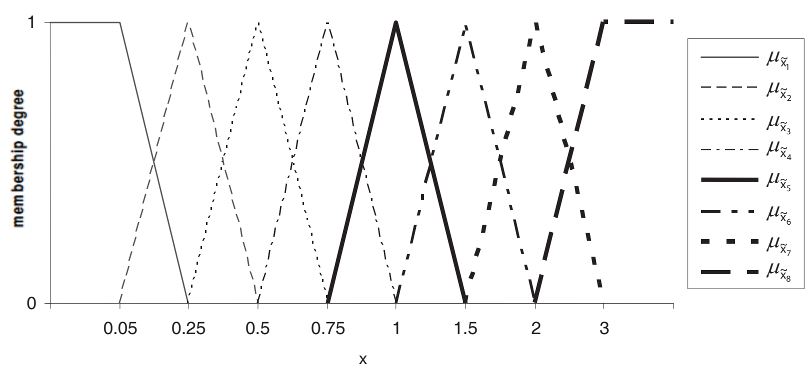

FIGURE 3: Fuzzy information system used to encode the simulated data.

7. Numerical Experiments

7.1. Simulation

In this section, we present some experimental results, mainly to observe how the different methods behave for different sample sizes. We obtain the estimates of the unknown parameters α and λ using the three methods provided in the preceding sections. The computations are performed using R 2.14.0 (R Development Core Team (2011)), which is a non-commercial, open source software package for statistical computing and graphics. First, for different sets of parameter values namely; $(\alpha, \lambda) = (0.5, 1), (1, 1), (2, 1)$, and various choices of n, we have generated i.i.d. random samples, say $\boldsymbol{x}$, from the Weibull distribution. Each realization of $\boldsymbol{x}$ was made fuzzy, using the f.i.s. shown in Fig.3, corresponding to the membership functions

$$\mu_{\widetilde{x}_1}(x) = \begin{cases} 1 & x \leq 0.05, \\ \frac{0.25-x}{0.2} & 0.05 \leq x \leq 0.25, \\ 0 & otherwise, \end{cases} \qquad \mu_{\widetilde{x}_2}(x) = \begin{cases} \frac{x-0.05}{0.2} & 0.05 \leq x \leq 0.25, \\ \frac{0.5-x}{0.25} & 0.25 \leq x \leq 0.5, \\ 0 & otherwise, \end{cases}$$

$$\mu_{\widetilde{x}_3}(x)=\begin{cases}\frac{x-0.25}{0.25} & 0.25\leq x\leq 0.5,\\ \frac{0.75-x}{0.25} & 0.5\leq x\leq 0.75,\\ 0 & otherwise,\end{cases}\qquad \mu_{\widetilde{x}_4}(x)=\begin{cases}\frac{x-0.5}{0.25} & 0.5\leq x\leq 0.75,\\ \frac{1-x}{0.25} & 0.75\leq x\leq 1,\\ 0 & otherwise,\end{cases}$$

$$\mu_{\widetilde{x}_5}(x)=\begin{cases}\frac{x-0.75}{0.25} & 0.75\leq x\leq 1,\\ \frac{1.5-x}{0.5} & 1\leq x\leq 1.5,\\ 0 & otherwise,\end{cases}\qquad \mu_{\widetilde{x}_6}(x)=\begin{cases}\frac{x-1}{0.5} & 1\leq x\leq 1.5,\\ \frac{2-x}{0.5} & 1.5\leq x\leq 2,\\ 0 & otherwise,\end{cases}$$

$$\mu_{\widetilde{x}_7}(x)=\begin{cases}\frac{x-1.5}{0.5} & 1.5\leq x\leq 2,\\ 3-x & 2\leq x\leq 3,\\ 0 & otherwise,\end{cases}\qquad \mu_{\widetilde{x}_8}(x)=\begin{cases}x-2 & 2\leq x\leq 3,\\ 1 & x\geq 3,\\ 0 & otherwise.\end{cases}$$

Then the estimates of α and λ for the fuzzy sample were computed using the maximum likelihood method (via NR and EM algorithms), the moments method and a Bayesian procedure. For computing the Bayes estimates, we have assumed that λ and α have $Gamma(a, b)$ and $Gamma(c, d)$ priors respectively. To make the comparison meaningful, it is assumed that the priors are non-informative, and they are $a = b = c = d = 0$. Note that in this case the priors are non-proper also. Press (2001) suggested to use very small non-negative values of the hyperparameters in this case, and it will make the priors proper. We have tried $a = b = c = d = 0.0001$. The results are not significantly different than the corresponding results obtained using non-proper priors, and are not reported due to space. From now on, the estimates of parameters obtained by using NR algorithm, EM algorithm, Bayesian procedure and moments method will be denoted by NR, EM, BET and MME, respectively. The average biases (AB) and mean squared errors (MSE) of the estimates over $5,000$ replications are presented in Tables 1-2.

From the experiments, we found that using the NR or EM algorithm for the computation of maximum likelihood estimates of α and λ give similar estimation results, but EM is computationally slower. For small and moderate sample sizes, the Bayesian procedure gives the most precise parameter estimates as shown by ABs and MSEs in Tables 1-2. For large sample sizes ($n = 100, 200$ and 500), the performance of the MLEs, MMEs and Bayes estimates are almost identical. For all the methods, it is observed that as the sample size increases, the biases and MSEs of the estimates decrease as expected.

7.2. Application example

In order to demonstrate the application of proposed methods, let us consider a case study on the light emitting diodes (LED) manufacturing process that focuses on the luminous intensities of LED sources. The process distribution has been justified and has been shown to be fairly close to the Weibull distribution. A sample of size $n = 30$ is taken from the stable process. Since the data given by

Table 1: MSE of the estimates of α and λ for different sample sizes.

n	α	λ	Estimation of α				Estimation of λ			
			NR	EM	BET	MME	NR	EM	BET	MME
15	0.5	1	0.0619	0.0620	0.0594	0.0705	0.0870	0.0871	0.0836	0.092
	1	1	0.0987	0.0988	0.0830	0.1246	0.1129	0.1130	0.1091	0.1191
	2	1	0.1263	0.1264	0.1129	0.1380	0.1465	0.1465	0.1421	0.1483
20	0.5	1	0.0558	0.0559	0.0512	0.0631	0.0727	0.0728	0.0639	0.0792
	1	1	0.0942	0.0943	0.0744	0.1193	0.1088	0.1089	0.0966	0.1139
	2	1	0.1017	0.1018	0.0922	0.1240	0.1226	0.1227	0.1182	0.1259
30	0.5	1	0.0366	0.0367	0.0341	0.0394	0.0489	0.0489	0.0422	0.0519
	1	1	0.0614	0.0614	0.0488	0.0828	0.0691	0.0692	0.0646	0.0707
	2	1	0.0721	0.0722	0.0630	0.0895	0.0843	0.0844	0.0819	0.0895
50	0.5	1	0.0285	0.0286	0.0257	0.0335	0.0365	0.0365	0.0342	0.0386
	1	1	0.0361	0.0362	0.0331	0.0451	0.0427	0.0427	0.0419	0.0430
	2	1	0.0488	0.0489	0.0425	0.0536	0.0572	0.0572	0.0558	0.0587
70	0.5	1	0.0214	0.0215	0.0208	0.0232	0.0305	0.0306	0.0291	0.0318
	1	1	0.0282	0.0282	0.0225	0.0346	0.0338	0.0339	0.0328	0.0345
	2	1	0.0327	0.0328	0.0311	0.0387	0.0478	0.0478	0.0460	0.0491
100	0.5	1	0.0154	0.0154	0.0152	0.0156	0.0227	0.0228	0.0220	0.0236
	1	1	0.0191	0.0192	0.0187	0.0195	0.0284	0.0285	0.0282	0.0289
	2	1	0.0270	0.0270	0.0263	0.0271	0.0395	0.0395	0.0390	0.0397
200	0.5	1	0.0104	0.0104	0.0098	0.0109	0.0174	0.0175	0.0168	0.0179
	1	1	0.0127	0.0128	0.0120	0.0134	0.0211	0.0211	0.0202	0.0218
	2	1	0.0214	0.0214	0.0209	0.0225	0.0356	0.0356	0.0348	0.0360
500	0.5	1	0.0055	0.0055	0.0051	0.0058	0.0118	0.0118	0.0113	0.0122
	1	1	0.0086	0.0086	0.0085	0.0088	0.0173	0.0174	0.0161	0.0179
	2	1	0.0142	0.0142	0.0139	0.0153	0.0235	0.0235	0.0230	0.0238

luminous intensity of a particular LED inevitably have some degree of imprecision, the luminous intensities of diodes are reported in the form of lower and upper bounds as well as a point estimate, which are as follows:

DATA SET:

$$
\begin{aligned}
&(2.163, 2.738, 3.068), (5.972, 6.353, 8.150), (1.032, 1.971, 2.642),\\
&(0.628, 0.964, 1.735), (2.995, 3.442, 5.066), (3.766, 5.814, 6.212),\\
&(0.974, 1.839, 2.045), (4.352, 5.206, 5.988), (3.920, 4.762, 6.121),\\
&(1.375, 2.195, 3.086), (0.618, 0.839, 2.217), (4.575, 6.050, 6.734),\\
&(1.027, 1.218, 3.116), (6.279, 8.156, 9.435), (2.821, 3.409, 5.272),\\
&(7.125, 8.470, 9.044), (5.443, 6.231, 7.395), (1.766, 2.190, 2.638),\\
&(7.155, 8.013, 8.352), (0.830, 1.288, 2.541), (3.590, 4.169, 4.899),\\
&(5.965, 7.344, 8.019), (3.177, 3.600, 4.213), (4.634, 5.780, 7.058),\\
&(7.261, 8.325, 8.871), (2.247, 2.990, 4.128), (6.032, 7.746, 8.529),\\
&(4.065, 5.312, 7.480), (5.434, 7.093, 7.655), (1.336, 2.750, 3.284).
\end{aligned}
$$

In our approach, each triplet is modeled by a triangular fuzzy number $\widetilde{x}_i$, and is interpreted as a possibility distribution related to an unknown value x_i, itself a realization of a random variable X_i. For this data, we employ NR and

Table 2: AB of the estimates of α and λ for different sample sizes.

n	α	λ	Estimation of α				Estimation of λ			
			NR	EM	BET	MME	NR	EM	BET	MME
15	0.5	1	0.1272	0.1273	0.0734	0.1533	0.1180	0.1181	0.1092	0.1230
	1	1	0.1381	0.1382	-0.0783	0.1698	0.1291	0.1292	0.1262	0.1322
2		1	0.1914	0.1915	0.1527	0.2038	0.1570	0.1571	0.1503	0.1637
20	0.5	1	0.1091	0.1092	-0.0617	0.1326	0.0931	0.0931	0.0865	0.1026
	1	1	0.1354	0.1355	-0.0699	0.1633	0.1205	0.1206	0.1177	0.1298
	2	1	0.1775	0.1775	0.1344	0.1851	0.1427	0.1428	0.1321	0.1485
30	0.5	1	0.0922	0.0923	0.0591	0.1130	0.0778	0.0779	0.0631	0.0840
	1	1	0.1228	0.1228	-0.0621	0.1417	0.1086	0.1087	0.1059	0.1152
	2	1	0.1439	0.1439	0.1223	0.1507	0.1162	0.1163	0.1137	0.1218
50	0.5	1	0.0754	0.0755	0.0518	0.0908	0.0620	0.0621	0.0582	0.0685
	1	1	0.0917	0.0918	-0.0571	0.1275	0.0927	0.0927	0.0905	0.0996
	2	1	0.1254	0.1255	0.1033	0.1445	0.1067	0.1067	0.1013	0.1151
70	0.5	1	0.0628	0.0629	0.0435	0.0711	0.0514	0.0514	0.0507	0.0536
	1	1	0.0887	0.0887	0.0494	0.1065	0.0833	0.0834	0.0821	0.0875
	2	1	0.1057	0.1058	-0.0932	0.1126	0.0983	0.0983	0.0970	0.0994
100	0.5	1	0.0413	0.0413	0.0408	0.0419	0.0459	0.0459	0.0455	0.0463
	1	1	0.0438	0.0438	0.0426	0.0440	0.0648	0.0648	0.0642	0.0655
	2	1	0.0906	0.0907	0.0896	0.0918	0.0952	0.0952	0.0948	0.0961
200	0.5	1	0.0287	0.0288	0.0281	0.0290	0.0317	0.0318	0.0314	0.0318
	1	1	0.0349	0.0349	0.0345	0.0353	0.0573	0.0573	0.0570	0.0574
	2	1	0.0855	0.0856	0.0851	0.0859	0.0736	0.0737	0.0733	0.0738
500	0.5	1	0.0211	0.0212	0.0207	0.0225	0.0244	0.0244	0.0241	0.0245
	1	1	0.0267	0.0268	0.0260	0.0271	0.0408	0.0409	0.0404	0.0412
	2	1	0.0762	0.0762	0.0758	0.0766	0.0553	0.0554	0.0550	0.0557

EM algorithms to compute the ML estimates. The stopping criterion is based on the difference between the two consecutive iterates, with a tolerance value $\varepsilon = 10^{-6}$. The final MLEs are $(\hat{\alpha}_{NR}, \hat{\lambda}_{NR}) = (2.1094, 0.0318)$ and $(\hat{\alpha}_{EM}, \hat{\lambda}_{EM}) = (2.1095, 0.0319)$. Also, by using the procedure presented in section 6, the moment estimate of (α, λ) becomes $(\hat{\alpha}_M, \hat{\lambda}_M) = (2.1257, 0.0374)$. For computing the Bayes estimate, we assume that both α and λ have a $Gamma(0.0001, 0.0001)$ prior. Therefore, using the Tierney and Kadane's approximation, the Bayes estimate of the parameters becomes $(\hat{\alpha}_{BT}, \hat{\lambda}_{BT}) = (2.1036, 0.0287)$.

8. Conclusions

Some work has been done in the past on the estimation of Weibull distribution parameters based on complete and censored samples. But, traditionally it is assumed that the available data are performed in exact numbers. However, some collected data might be imprecise and are represented in the form of fuzzy numbers. Therefore, we need suitable statistical methodology to handle these data as well. In this paper, we have discussed different estimation procedures for the Weibull distribution when the obtained data are fuzzy numbers. They include the maximum likelihood method (via NR and EM algorithms), a Bayesian procedure and the method of moments. We have then carried out a simulation study to assess the performance of all these procedures. The recommendations of an estimator based on minimum biases and MSEs are as follows:

i) For small and moderate sample sizes, the performance of the Bayes estimates is generally best followed by the MLEs and then the MMEs. Thus, it would seem reasonable to recommend the use of the Bayesian procedure for estimating the unknown parameters α and λ.

ii) For large sample sizes, the three estimation procedures behave in similar manner.

Acknowledgments

The authors are thankful to the referees for their valuable comments which led to a considerable improvement in the presentation of this article.

[Recibido: febrero de 2013 — Aceptado: septiembre de 2013]

References

Ageel, M. I. (2002), 'A novel means of estimating quantiles for 2-parameter Weibull distribution under the right random censoring model', *Journal of Computational and Applied Mathematics* **149**(2), 373–380.

Al-Baidhani, P. A. & Sinclair, C. (1987), 'Comparison of methods of estimation of parameters of the Weibull distribution', *Communications in Statistics-Simulation and Computation* **16**(2), 373–384.

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**(17), 2971–2975.

Balakrishnan, N. & Mitra, D. (2012), 'Left truncated and right censored Weibull data and likelihood inference with an illustration', *Computational Statistics and Data Analysis* **56**, 4011–4025.

Banerjee, A. & Kundu, D. (2012), 'Inference based on type-II hybrid censored data from a Weibull distribution', *IEEE Transactions on Reliability* **57**(2), 369–378.

Denoeux, T. (2011), 'Maximum likelihood estimation from fuzzy data using the EM algorithm', *Fuzzy Sets and Systems* **183**(1), 72–91.

Dubois, D. & Prade, H. (1980), *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, New York.

Gertner, G. Z. & Zhu, H. (1996), 'Bayesian estimation in forest surveys when samples or prior information are fuzzy', *Fuzzy Sets and Systems* **77**, 277–290.

Helu, A., Abu-Salih, M. & Alkam, O. (2010), 'Bayes estimation of Weibull distribution parameters using ranked set sampling', *Communications in Statistics-Theory and Methods* **39**(14), 2533–2551.

Joarder, A., Krishna, H. & Kundu, D. (2011), 'Inferences on Weibull parameters with conventional type-I censoring', *Computational Statistics and Data Analysis* **55**, 1–11.

Lin, C., Chou, C. & Huang, Y. (2012), 'Inference for the Weibull distribution with progressive hybrid censoring', *Computational Statistics and Data Analysis* **56**, 451–467.

Marks, N. B. (2005), 'Estimation of Weibull parameters from common percentiles', *Journal of Applied Statistics* **32**(1), 17–24.

Nandi, S. & Dewan, I. (2010), 'An EM algorithm for estimating the parameters of bivariate Weibull distribution under random censoring', *Computational Statistics and Data Analysis* **54**(6), 1559–1569.

Ng, H. K. T. & Wang, Z. (2009), 'Statistical estimation for the parameters of Weibull distribution based on progressively type-I interval censored sample', *Journal of Statistical Computation and Simulation* **79**(2), 145–159.

Press, S. J. (2001), *The Subjectivity of Scientists and the Bayesian Approach*, Wiley, New York.

Qiao, O. & Tsokos, C. P. (1994), 'Parameter estimation of the Weibull probability distribution', *Mathematics and Computers in Simulation* **37**, 47–55.

R Development Core Team (2011), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
*http://www.R-project.org

Singpurwalla, N. D. & Booker, J. M. (2004), 'Membership functions and probability measures of fuzzy sets', *Journal of the American Statistical Association* **99**(467), 867–877.

Tan, Z. (2009), 'A new approach to MLE of Weibull distribution with interval data', *Reliability Engineering and System Safety* **94**(2), 394–403.

Tierney, L. & Kadane, J. B. (1986), 'Accurate approximations for posterior moments and marginal densities', *Journal of the American Statistical Association* **81**, 82–86.

Watkins, A. J. (1994), 'On maximum likelihood estimation for the two parameter Weibull distribution', *Microelectronics Reliability* **36**(5), 595–603.

Zadeh, L. A. (1968), 'Probability measures of fuzzy events', *Journal of Mathematical Analysis and Applications* **10**, 421–427.