Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
`http://siba-ese.unisalento.it/index.php/ejasa/index`
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v9n1p227

Inference for exponential mean parameter under progressive Type-II censoring from imprecise lifetime

By Pak, Chatrabgoun

Published: 26 April 2016

Electronic Journal of Applied Statistical Analysis
Vol. 09, Issue 01, April 2016, 227-245
DOI: 10.1285/i20705948v9n1p227

Inference for exponential mean parameter under progressive Type-II censoring from imprecise lifetime

Abbas Pak*[a] and Omid Chatrabgoun[b]

[a] *Department of Computer Sciences, Faculty of Mathematical Sciences, Shahrekord University, P. O. Box 115, Shahrekord, Iran*
[b] *Department of Statistics, Malayer University, Malayer, Iran*

Published: 26 April 2016

Progressively Type-II censored sampling is an important method of obtaining data in lifetime studies. Statistical analysis of lifetime distributions under this censoring scheme is based on precise lifetime data. However, in real situations all observations and measurements of progressive Type-II censoring scheme are not precise numbers but more or less non-precise, also called fuzzy. In this paper, we consider the estimation of exponential mean parameter under progressive Type-II censoring scheme, when the lifetime observations are fuzzy and are assumed to be related to underlying crisp realization of a random sample. We propose a new method to determine the maximum likelihood estimate of the unknown mean parameter. In addition, a new numerical method for parameter estimation based on fuzzy data is provided. Using the parametric bootstrap method, we then discuss the construction of confidence intervals for the mean parameter. Monte Carlo simulations are performed to investigate the performance of all the different proposed methods. Finally, an illustrative example is also included.

keywords: Progressive Type-II censoring, Imprecise lifetime, Maximum likelihood estimation, Bootstrap confidence interval.

*Corresponding author:abbas.pak1982@gmail.com.

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

1 Introduction

The scheme of progressive censoring is of great value in life-testing experiments. Its allowance for the removal of life units from the experiment at various stages is an attractive feature as it will potentially save a lot for the experimenter in terms of cost and time. It can be described as follows. Suppose that n units are placed on a life test and the experimenter decides beforehand the quantity m, the number of units to be failed. Now at the time of the first failure, R_1 of the remaining $n-1$ surviving units are randomly removed from the experiment. Continuing on, at the time of the second failure, R_2 of the remaining $n-R_1-2$ units are randomly removed from the experiment. Finally, at the time of the mth failure, all the remaining $n-m-R_1-...-R_{m-1}(=R_m)$ surviving units are removed from the experiment.

Several authors considered different estimation procedures for lifetime distributions based on progressively Type-II censored data. See, for example Cohen (1963), Viveros and Balakrishnan (1994), Fernández (2004), Balakrishnan and Asgharzadeh (2005), Raqab and Madi (2011) and Al-Zahrani and Gindwan (2015). Their research results for estimating parameters of different lifetime distributions under progressive Type-II censoring are limited to precise data. However, in real situations all observations and measurements of continuous variables are not precise numbers but more or less non-precise. For instance, the lifetime observations may be reported as imprecise quantities such as: 'about $1000h$', 'approximately $1400h$', 'almost between $1000h$ and $1200h$', 'essentially less than $1200h$', and so on. This imprecision is different from variability and errors. The best up-to-date mathematical model for this imprecision are so-called non-precise (fuzzy) numbers. Classical statistical procedures and Bayesian inference are not appropriate to deal with such imprecise cases. Therefore, we need suitable statistical methodology to handle these data as well.

In recent years, several researchers pay attention to applying the fuzzy sets to estimation theory. Huang et al. (2006) proposed a new method to determine the membership function of the estimates of the parameters and the reliability function of multiparameter lifetime distributions. Coppi et al. (2006) presented some applications of fuzzy techniques in statistical analysis. Pak et al. (2013) and Pak et al. (2014) conducted a series of studies to develop the inferential procedures for the lifetime distributions on the basis of fuzzy data. In this paper, we discuss different methods of estimation for the parameter of exponential distribution on the basis of progressively Type-II censoring scheme when the available observations are described by means of fuzzy numbers. In Section 2, we first present in more detail the problem addressed in this paper. Some preliminary concepts about fuzzy numbers is included in this section. In Section 3, we obtain the maximum likelihood estimate (MLE) of the unknown mean parameter by using EM algorithm. A new parameter estimation method, called 'computational approach estimation' (CAE), is established in Section 4. By using the parametric bootstrap method, we then discuss the construction of confidence intervals for the unknown mean parameter in Section 5. In Section 6, a simulation study is carried out to assess the performance of the different proposed methods, and a practical example in life testing is analyzed for illustrative purposes.

2 Problem description

Consider a reliability experiment in which n identical units are placed on a life-test. Let $X_1, ..., X_n$ denote the lifetimes of these experimental units. Assume that these variables are independent and identically distributed as exponential distribution, $E(\theta)$, with probability density function (pdf)

$$f(x;\theta) = \frac{1}{\theta}\exp(-\frac{x}{\theta}) \quad ,x > 0, \theta > 0. \tag{1}$$

Prior to the experiment, a number $m < n$ is determined and the censoring scheme $(R_1, R_2, ..., R_m)$ with $R_i \geq 0$ and $\sum\limits_{i=1}^{m} R_i + m = n$ is specified. During the experiment, ith failure is observed and immediately after the failure, R_i functioning items are randomly removed from the test. Let $x_{1:m:n}, ..., x_{m:m:n}$ denote the m completely observed lifetimes. The likelihood function based on this progressively Type-II censored sample is (see Balakrishnan and Aggarwala, 2000)

$$L(\theta;\mathbf{x}) = \frac{C}{\theta^m}\exp(-\frac{\sum\limits_{i=1}^{m}(1+R_i)x_{i:m:n}}{\theta}), \tag{2}$$

where

$$C = n(n - R_1 - 1)(n - R_1 - R_2 - 2)...(n - R_1 - ... - R_{m-1} - m + 1).$$

Thus, the maximum likelihood estimator (MLE) of θ can be obtained by

$$\hat{\theta} = \frac{1}{m}(\sum_{i=1}^{m}(1+R_i)x_{i:m:n}).$$

However, in many fields of application, lifetime observations 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. One of the usual models of censored data is random interval-censorship. In this case, the event is only known to happen between two random examination times. The observations are thus of the form (U_i, V_i), $i = 1, ..., n$, and it is only known that $U_i \leq X_i \leq V_i$, where the X_i, $i = 1, ..., n$, are the (partially observed) survival times. Statistical analysis of lifetime distributions based on interval censored data has been discussed by Ng and Wang (2009) and Tan (2009), among others.

The problem addressed in this paper, is different from interval censoring. 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, consider a life-testing experiment in which

n identical batteries are placed on a test, and we are interested in the lifetime of these batteries. A tested battery may be considered as failed, or -strictly speaking- as non-conforming, when at least one value of its parameters falls beyond specification limits. In practice, however, the observer does not have the possibility to measure all parameters and is not able to define precisely the moment of failures, but rather he/she can only approximate them by means of the following imprecise quantities: "approximately lower than 1000 hours ", "approximately 1500 to 2000 hours", "approximately 2500 hours", "approximately 3000 hours", "approximately 3500 to 4000 hours", "approximately higher than 4500 hours", and so on. In order to model imprecise lifetimes, a generalization of real numbers is necessary. These lifetimes can be represented by fuzzy numbers. A fuzzy number is a subset, denoted by $\tilde{x}$, of the set of real numbers (denoted by $\mathbb{R}$) and is characterized by the so called membership function $\mu_{\tilde{x}}(.)$. Fuzzy numbers satisfy the following constraints (Dubois, 1980):

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

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

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

Among the various types of fuzzy numbers, LR-type fuzzy numbers (the triangular and trapezoidal fuzzy numbers are special cases of the LR-type fuzzy numbers) are most convenient and useful in describing fuzzy lifetime data. Therefore, we shall focus on the set of LR-type fuzzy numbers.

Definition 1 (Zimmermann, 2011, pp. 62). Let L (and R) be decreasing, shape functions from $\mathbb{R}^+$ to $[0,1]$ with $L(0) = 1$; $L(x) < 1$ for all $x > 0$; $L(x) > 0$ for all $x < 1$; $L(1) = 0$ or ($L(x) > 0$ for all x and $L(+\infty) = 0$). Then a fuzzy number $\tilde{x}$ is called of LR-type if for $c, \alpha > 0, \ \beta > 0$ in $\mathbb{R}$,

$$\mu_{\tilde{x}}(x) = \begin{cases} L(\frac{c-x}{\alpha}) & x \leq c \\ R(\frac{x-c}{\beta}) & x \geq c \end{cases}$$

where c is called the mean value of $\tilde{x}$ and α and β are called the left and right spreads, respectively. Symbolically, the LR-type fuzzy number is denoted by $\tilde{x} = (\alpha, c, \beta)$.

Definition 2 Suppose that $\tilde{x}_i = (\alpha_i, c_i, \beta_i), \ i = 1, ..., n$, be the LR-type fuzzy numbers. *The fuzzy mean value* of these numbers can be obtained as $\bar{\tilde{x}} = (\bar{\alpha}, \bar{c}, \bar{\beta})$ where

$$\bar{\alpha} = \frac{1}{n}\sum_{i=1}^{n}\alpha_i, \ \bar{c} = \frac{1}{n}\sum_{i=1}^{n}c_i \quad \text{and} \quad \bar{\beta} = \frac{1}{n}\sum_{i=1}^{n}\beta_i. \tag{3}$$

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

It must be noted that, our viewpoint in this paper is based on an *epistemic interpretation* of fuzzy data, which are assumed to "imperfectly specify a value that is existing and precise, but not measurable with exactitude under the given observation conditions"

(Gebhardt et al., 1998, p. 316). In this model, a fuzzy datum is thus seen as a possibility distribution associated to a precise realization of a random variable that has been only partially observed. In the next section, we introduce a generalization of the likelihood function under progressively Type-II fuzzy censored data and obtain the MLE of θ.

3 The MLE of mean parameter

Suppose that n independent units are put on a test and that the lifetime distribution of each unit is given by (1). Now consider the problem where under the progressively Type-II censoring scheme, failure times are not observed precisely and only partial information about them are available in the form of fuzzy numbers $\tilde{x}_i = (\alpha_i, c_i, \beta_i),\ i = 1, ..., m$, with the corresponding membership functions $\mu_{\tilde{x}_1}(.), ..., \mu_{\tilde{x}_m}(.)$. Let $c_{(1)} \leq c_{(2)} \leq ... \leq c_{(m)}$ denote the ordered values of the means of these fuzzy numbers. The lifetime of R_i surviving units, which are removed from the test after the ith failure, can be encoded as fuzzy numbers $\tilde{z}_{i1}, ..., \tilde{z}_{iR_i}$ with the membership functions

$$\mu_{\tilde{z}_{ij}}(z) = \begin{cases} 0 & z \leq c_{(i)} \\ 1 & z > c_{(i)} \end{cases}, \qquad j = 1, ..., R_i.$$

The fuzzy data $\tilde{\mathbf{w}} = (\tilde{x}_1, ..., \tilde{x}_m, \tilde{\mathbf{z}}_1, ..., \tilde{\mathbf{z}}_\mathbf{m})$, where $\tilde{\mathbf{z}}_i$ is a $1 \times R_i$ vector with $\tilde{\mathbf{z}}_i = (\tilde{z}_{i1}, \tilde{z}_{i2}, ..., \tilde{z}_{iR_i})$, for $i = 1, ..., m$, is thus the set of observed lifetimes. The corresponding observed-data log-likelihood function can be obtained, using Zadeh's definition of the probability of a fuzzy event (Zadeh, 1968), as

$$L_O(\tilde{\mathbf{w}}; \theta) = \sum_{i=1}^{m} \log \left\{ \int \frac{1}{\theta} \exp(-\frac{x}{\theta}) \mu_{\tilde{x}_i}(x) dx \right\} + \sum_{i=1}^{m} \sum_{j=1}^{R_i} \log \left\{ \int \frac{1}{\theta} \exp(-\frac{z}{\theta}) \mu_{\tilde{z}_{ij}}(z) dz \right\}. \tag{4}$$

Since the observed fuzzy data $\tilde{\mathbf{w}}$ can be seen as an incomplete specification of a complete data vector $\mathbf{w}$, the EM algorithm is applicable to obtain the maximum likelihood estimate of the parameter. The EM algorithm, introduced by Dempster et al. (1977), is a very popular tool to handle any missing or incomplete data situation. This algorithm is an iterative method which has two steps. In the E-step, it replaces any missing data by its expected value and in the M-step the log-likelihood function is maximized with the observed data and expected value of the incomplete data, producing an update of the parameter estimates. In the following, we describe the EM algorithm to determine the MLE of the parameter θ.

First of all, denote the lifetime of the failed and censored units by $\mathbf{X} = (X_1, ..., X_m)$ and $\mathbf{Z} = (\mathbf{Z}_1, ..., \mathbf{Z}_m)$, respectively, where $\mathbf{Z}_i$ is a $1 \times R_i$ vector with $\mathbf{Z}_i = (Z_{i1}, ...Z_{iR_i})$, for $i = 1, ..., m$. The combination of $(\mathbf{X}, \mathbf{Z}) = \mathbf{W}$ forms the complete lifetimes and the corresponding log-likelihood function is denoted by $L(\mathbf{W}; \theta)$. To perform the E-step of the algorithm, we need to compute the conditional expectation of the complete-data log likelihood $logL(\mathbf{W}; \theta)$ conditionally on the observed data $\tilde{\mathbf{w}}$, as follows:

$$E_{\theta^{(h)}}\left(logL(\mathbf{W};\theta) \mid \tilde{\mathbf{w}}\right) = -n\log\theta - \frac{1}{\theta}\left[\sum_{i=1}^{m} E_{\theta^{(h)}}(X_i \mid \tilde{x}_i) + \sum_{i=1}^{m}\sum_{j=1}^{R_i} E_{\theta^{(h)}}(Z_{ij} \mid \tilde{z}_{ij})\right], \tag{5}$$

where $\theta^{(h)}$ denotes the current fit of θ at iteration h. The conditional expectations $\alpha_i^{(h)} = E_{\theta^{(h)}}(X_i \mid \tilde{x}_i)$ and $\beta_{ij}^{(h)} = E_{\theta^{(h)}}(Z_{ij} \mid \tilde{z}_{ij})$ can be computed using

$$E_{\theta^{(h)}}(U \mid \tilde{u}) = \int u f(u \mid \tilde{u};\theta^{(h)})du, \tag{6}$$

where the conditional density of U given $\tilde{u}$ can be obtained from

$$f(u \mid \tilde{u};\theta^{(h)}) = \frac{\frac{1}{\theta^{(h)}}\exp(-\frac{u}{\theta^{(h)}})\mu_{\tilde{u}}(u)}{\int \frac{1}{\theta^{(h)}}\exp(-\frac{u}{\theta^{(h)}})\mu_{\tilde{u}}(u)du}. \tag{7}$$

The M-step then consists in finding $\theta^{(h+1)}$ which maximizes $E_{\theta^{(h)}}(\log L(\mathbf{W};\theta) \mid \tilde{\mathbf{w}})$. This is easily achieved by solving the likelihood equation. From

$$\frac{\partial}{\partial\theta}E_{\theta^{(h)}}(\log L(\mathbf{W};\theta) \mid \tilde{\mathbf{w}}) = 0, \tag{8}$$

we get

$$\theta^{(h+1)} = \frac{1}{n}\left[\sum_{i=1}^{m}\alpha_i^{(h)} + \sum_{i=1}^{m}\sum_{j=1}^{R_i}\beta_{ij}^{(h)}\right]. \tag{9}$$

The MLE of θ can be obtained by repeating the E- and M-steps until the difference $L_O(\tilde{\mathbf{w}};\theta^{(h+1)}) - L_O(\tilde{\mathbf{w}};\theta^{(h)})$ becomes smaller than some arbitrary small amount.

4 Computational approach estimation method

In this section we propose a new parameter estimation procedure called 'CAE'. Although the maximum likelihood estimate obtained in the preceding section is preferable, its computation requires repeated evaluation of E- and M-steps until convergence occurs. On the other hand, the CAE method provides not only the computational ease but also reasonable mean squared errors. This finding is further discussed in Section 6.

Let $\tilde{x}_i = (\alpha_i, c_i, \beta_i),\ i = 1, ..., m$, be the original progressively Type-II censored sample from the population given in (1), with $(R_1, R_2, ..., R_m)$ being the progressive censoring scheme. Grzegorzewski and Hryniewicz (2002) considered the generalization of exponential model which admits vagueness in lifetimes. They obtained a fuzzy estimator for the mean lifetime θ in the presence of vague lifetimes. However, in most applications, crisp results are required instead of fuzzy ones. So, we propose the following computational approach to obtain a crisp value as an estimate of θ .

Step 1: Order the means of fuzzy numbers $\tilde{x}_i = (\alpha_i, c_i, \beta_i), \quad i = 1, ..., m$, as $c_{(1)} < c_{(2)} < ... < c_{(m)}$.

Step 2: Obtain the fuzzy mean value, say $\bar{\tilde{x}}$, of the fuzzy numbers by using (3).

Step 3: Convert the fuzzy number $\bar{\tilde{x}}$ into a real value by using the center of gravity defuzzification technique (see Appendix A) as follows.

$$\approx{x}^* = \frac{\int x\mu_{\bar{\tilde{x}}}(x)dx}{\int \mu_{\bar{\tilde{x}}}(x)dx}.$$

Step 4: The new estimate of θ is then computed by:

$$\overset{\approx}{\theta} = \overset{\approx}{x}^* + \frac{1}{m}(\sum_{i=1}^{m} R_i c_{(i)}). \qquad (10)$$

5 Bootstrap confidence intervals

In this section, we discuss the construction of confidence intervals (CIS) for the unknown parameter θ using the two types of bootstrap methods, *viz.*, percentile bootstrap (Boot-p) method and the bias-corrected and accelerated (BCa) percentile bootstrap method; see Efron and Efron (1982) and Efron and Tibshirani (1994) for pertinent details.

Before we discuss the construction of the bootstrap confidence intervals for θ, the following algorithm is used to generate the bootstrap sample of fuzzy numbers based on the original progressively Type-II censored fuzzy sample $\tilde{x}_i = (\alpha_i, x_i, \beta_i), \quad i = 1, ..., m$.

Step 1: Given the original progressively Type-II censored fuzzy sample $\tilde{x}_1, ..., \tilde{x}_m$, compute the MLE of θ, say $\hat{\theta}$, using the iterative algorithm (9).

Step 2: Generate the progressively Type-II censored sample of size m, say $x_1^*, ..., x_m^*$, with the underlying distribution as $E(\hat{\theta})$. Define the $LR-$type fuzzy numbers $\tilde{x}_1^*, ..., \tilde{x}_m^*$ as $\tilde{x}_i^* = (\alpha_i, x_i^*, \beta_i), \quad i = 1, ..., m$.

Step 3: Based on the simulated progressively Type-II censored fuzzy sample in Step 2 , calculate the bootstrap MLE of θ, denoted by $\hat{\theta}^*$, from (9).

Step 4: Repeat step 2 and 3, M times. Then, arrange all bootstrapped values of $\hat{\theta}^*$ in ascending order to obtain the ordered bootstrap sample of $\hat{\theta}_1^* < \hat{\theta}_2^* < ... < \hat{\theta}_M^*$.
With the bootstrap sample generated as above, we propose the following two parametric bootstrap confidence interval for θ.

Boot-p confidence interval:

A two sided $100(1-\gamma)\%$ percentile bootstrap CI for θ is

$$(\hat{\theta}^*_{[M(\frac{\gamma}{2})]}, \hat{\theta}^*_{[M(1-\frac{\gamma}{2})]}).$$

BCa percentile bootstrap confidence interval:

A two sided $100(1-\gamma)\%$ BCa percentile bootstrap CI for θ is

$$\left(\hat{\theta}^*_{[\alpha M]}, \hat{\theta}^*_{[\beta M]}\right),$$

where

$$\alpha = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 - z_{\alpha/2}}{1 - \hat{a}(\hat{z}_0 - z_{\alpha/2})}\right)$$

and

$$\beta = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z_{\alpha/2}}{1 - \hat{a}(\hat{z}_0 + z_{\alpha/2})}\right).$$

Here, $\Phi(.)$ denotes the CDF of the standard normal distribution, z_α is the upper $\alpha-$point of the standard normal distribution and $[x]$ denote the integer part of x. The value of bias-correction $\hat{z}_0$ is given by

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\sum_{j=1}^{M} I(\hat{\theta}^*_j < \hat{\theta})}{M}\right),$$

and a good estimate of the acceleration factor $\hat{a}$ is suggested to be

$$\hat{a} = \frac{\sum_{j=1}^{m} (\hat{\theta}^{(\cdot)} - \hat{\theta}^{(j)})^3}{6\left\{\sum_{j=1}^{m} (\hat{\theta}^{(\cdot)} - \hat{\theta}^{(j)})^2\right\}^{3/2}},$$

where $\hat{\theta}^{(j)}$ is the MLE of θ based on the original Type-II censored fuzzy sample with the jth observation deleted for $j = 1, ..., m$, and

$$\hat{\theta}^{(\cdot)} = \frac{1}{m}\sum_{j=1}^{m} \hat{\theta}^{(j)}.$$

6 Numerical experiments

6.1 Simulation

In order to evaluate the performance of all the different methods discussed in the preceding sections, a Monte Carlo simulation study was conducted and its results are described in this section. First, for fixed $\theta = 1$ and different choices of n , m and $(R_1, ..., R_m)$, we have generated progressively censored sample $x_1, ..., x_m$ from exponential distribution using the method proposed by Balakrishnan and Sandhu (1995). Then we define fuzzy numbers $\tilde{x}_1, ..., \tilde{x}_m$ with the corresponding membership functions

$$\mu_{\tilde{x}_i}(x) = \begin{cases} \frac{x-(x_i-h_i)}{h_i} & x_i - h_i < x \leq x_i \\ \frac{x_i+h_i-x}{h_i} & x_i < x \leq x_i + h_i \end{cases}, \qquad i = 1, ..., r,$$

where $h_i = 0.05x_i$. This procedure simulates the situation where the observer has only approximate knowledge of the failure times, and can only provide a guess x_i and an interval of plausible values $[x_i - h_i, x_i + h_i]$. From these fuzzy numbers, we obtain the MLE of θ, using the iterative algorithm (9). We have used the initial estimate to be $\theta^{(0)} = \frac{1}{m}\sum_{i=1}^{m} x_i$ and the iterative process stops when the relative change of the log-likelihood becomes less than 10^{-6}. We also obtain the estimate of θ using the CAE method. The average values (AV) and mean squared errors(MSE) of the estimates based on 1000 replication are presented in Tables 2-4. We also compute the 95% CIs for θ using the Boot-p and BCa percentile bootstrap methods. The average confidence lengths and coverage probabilities of the confidence intervals are reported in Tables 5-7.

From Tables 2-4, the following observations are made. The performance of the MLEs are satisfactory in terms of AVs and MSEs. For fixed n as m increases, the MSEs decrease for all cases as expected. Similar observations are made for the estimates of θ obtained from the CAE method. It is also observed that, the MSEs of the estimates based on the CAE method are quite close to that of the MLEs. Note that the above estimation results can be attributed to the assumed fuzzy numbers. The rationales for such fuzzy numbers, which are characterized by the membership functions, may influence the estimate results.

Among the bootstrap methods for constructing confidence interval of θ, the BCa percentile bootstrap method is better than the Boot-p method with respect to the coverage probabilities. From Tables 5-7, we can see that the coverage probabilities of the BCa bootstrap CIs are close to the nominal level unless the effective relative sample fraction $(\frac{m}{n})$ is small, while the same of the Boot-p CIs are lower than the nominal level. The average confidence lengths of the Boot-p CIs are slightly smaller than the BCa bootstrap CIs. We also realize from Tables 5-7 that a larger effective sample size (m) eventually improves the probability coverages and lengths for both Boot-p and BCa bootstrap CIs.

6.2 Illustrative example

In a life-testing experiment the lifetimes (in $1000km$) of front disk brake pads on a randomly selected set of $n = 40$ cars are monitored by a dealer network. Suppose that the random variable of interest is distributed exponentially by an unknown mean parameter of θ. But, in practice measuring the lifetime of a disk may not yield an exact result. A disk may work perfectly over a certain period but be braking for some time, and finally be unusable at a certain time. So, the observed lifetimes are reported as fuzzy numbers given below. In fact, imprecision is formulated by fuzzy numbers $\tilde{x}_i = (h_i, x_i)$, where $h_i = 0.005x_i$, $i = 1, ..., 40$, with membership functions

$$\mu_{\tilde{x}_i}(x) = \begin{cases} \frac{x-(x_i-h_i)}{h_i} & x_i - h_i \le x \le x_i \\ 0 & x > x_i \end{cases}$$

DATA SET:

$(0.11, 22.6)$, $(0.11, 22.7)$, $(0.14, 28.4)$, $(0.15, 31.7)$, $(0.16, 33.8)$,
$(0.16, 33.9)$, $(0.17, 34.4)$, $(0.18, 36.7)$, $(0.19, 38.4)$, $(0.19, 38.8)$,
$(0.20, 40.0)$, $(0.20, 41.0)$, $(0.21, 42.2)$, $(0.21, 42.7)$, $(0.21, 42.8)$,
$(0.22, 45.1)$, $(0.22, 45.5)$, $(0.22, 45.9)$, $(0.23, 46.9)$, $(0.24, 48.8)$,
$(0.25, 50.6)$, $(0.25, 50.7)$, $(0.25, 50.2)$, $(0.25, 51.6)$, $(0.26, 52.1)$,
$(0.26, 53.6)$, $(0.27, 54.2)$, $(0.28, 56.4)$, $(0.28, 56.7)$, $(0.29, 59.0)$,
$(0.29, 59.8)$, $(0.30, 61.5)$, $(0.31, 62.4)$, $(0.32, 64.5)$, $(0.36, 73.1)$,
$(0.40, 80.6)$, $(0.40, 81.3)$, $(0.40, 81.7)$, $(0.43, 86.2)$, $(0.51, 102.5)$.

We consider progressively censored samples of size $m = 16$ from the above data using three different sampling schemes, namely: Scheme 1: $R_1 = ... = R_{m-1} = 0$ and $R_m = 24$, Scheme 2: $R_1 = 24$ and $R_2 = ... = R_m = 0$ and Scheme 3: $R_1 = ... = R_{m-1} = 1$ and $R_m = 9$. In all the three cases, we obtain the estimates of θ using the MLE and CAE procedures. We also compute the 95% confidence intervals using the Boot-p and BCa percentile bootstrap methods. All the results are summarized in Table 1.

Table 1: MLE, CAE, Boot-p and the BCa confidence intervals for Example 1.

Scheme	MLE	CAE	Boot-p	BCa
1	103.2946	103.4273	(57.3583,156.2837)	(69.4044,189.6155)
2	69.6673	69.7336	(38.5471,106.4726)	(48.1239,131.7655)
3	94.1784	94.2925	(53.4195,148.2681)	(61.3592,166.8601)

7 Conclusions

In this paper we have proposed different procedures for estimating the exponential distribution parameter under progressively Type-II censoring when the lifetime observations are fuzzy numbers. We have derived the MLE of the unknown parameter θ. Also, a computational approach method for estimating θ from fuzzy numbers is introduced. Then, We have presented two procedures for constructing confidence intervals for θ. Based on the results of simulation study, we see clearly that, as the effective sample size increases, the performances of the MLEs in terms of MSEs become better. The performance of the estimates of θ based on the CAE method are quite similar to the MLEs, but it can be easily seen that the CAE algorithm has no complicated likelihood equations involved for solving the solution as the EM algorithm. Therefore, it can be efficiently implemented through a computing program.. We also see that, compared to the ordinary percentile bootstrap CIs, the BCa percentile bootstrap CIs perform better in terms of the coverage probabilities. The coverage probabilities of the CIs based on the BCa percentile bootstrap method are quite close to the nominal level unless the effective relative sample fraction ($\frac{m}{n}$) is small.

Appendix A

Center of gravity defuzzification method
Defuzzyfication is the opposite process to the essence of idea of fuzzy sets. Moreover, defuzzyfication is the last step on fuzzy control system and fuzzy reasoning system. Finally, defuzzyfication operation reduces, fuzzy number to a single, crisp, numerical value, result carries the best information and makes kind of synthesis about this fuzzy number. A common and useful defuzzification technique is the center of gravity method. The center of gravity defuzzification technique can be expressed as

$$z^* = \frac{\int z\mu_{\tilde{C}}(z)dz}{\int \mu_{\tilde{C}}(z)dz}$$

where z^* is the defuzzified output, $\mu_{\tilde{C}}(z)$ is the membership function of the fuzzy set $\tilde{C}$.

Acknowledgement

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

Table 2: Average values (AV) and mean squared errors (MSE) of the estimates of θ based on the MLE and CAE methods for $n = 20$ and different sampling schemes.

m	Censoring Scheme	MLE		CAE	
		AV	MSE	AV	MSE
5	(0,0,0,0,15)	1.0099	0.1995	1.0072	0.2010
	(15,0,0,0,0)	0.9875	0.2152	0.9893	0.2168
	(0,15,0,0,0)	1.0086	0.2002	1.0091	0.2017
	(1,1,1,1,11)	0.9881	0.2010	0.9916	0.2024
	(2,2,2,2,7)	0.9871	0.2109	0.9835	0.2122
7	(0,...,0,13)	1.0079	0.1344	1.0081	0.1351
	(13,0,...,0)	0.9962	0.1302	0.9944	0.1309
	(0,13,0,...,0)	1.0083	0.1344	1.0091	0.1352
	(2,...,2,1)	0.9915	0.1410	0.9940	0.1417
	(1,...,1,7)	0.9937	0.1313	0.9921	0.1320
10	(0,...,0,10)	0.9969	0.1033	0.9973	0.1036
	(10,0,...,0)	0.9965	0.1036	0.9945	0.1045
	(0,10,0,...,0)	0.9956	0.1029	0.9976	0.1033
	(1,...,1)	1.0081	0.0978	1.0095	0.0971
15	(0,...,0,5)	0.9992	0.0638	0.9977	0.0643
	(5,0,...,0)	1.0030	0.0691	1.0042	0.0693
	(0,5,0,...,0)	1.0033	0.0709	1.0043	0.0711

Table 3: Average values (AV) and mean squared errors (MSE) of the estimates of θ based on the MLE and CAE methods for $n = 30$ and different sampling schemes.

m	Censoring Scheme	MLE		CAE	
		AV	MSE	AV	MSE
10	(0,0,0,0,20)	0.9926	0.0991	0.9934	0.0995
	(20,0,0,0,0)	0.9901	0.1009	0.9915	0.1011
	(0,20,0,0,0)	0.9905	0.1011	0.9911	0.1018
	(1,...,1,11)	1.0082	0.0949	1.0097	0.0954
	(2,...,2)	1.0068	0.0968	1.0081	0.0974
15	(0,...,0,15)	1.0058	0.0628	1.0074	0.0633
	(15,0,...,0)	0.9932	0.0666	0.9927	0.0670
	(0,15,0,...,0)	0.9920	0.0650	0.9925	0.0652
	(1,...,1)	0.9945	0.0644	0.9959	0.0646
20	(0,...,0,10)	1.0033	0.0513	1.0041	0.0515
	(10,0,...,0)	0.9997	0.0500	1.0014	0.0501
	(0,10,0,...,0)	0.9966	0.0475	0.9962	0.0476
25	(0,...,0,5)	0.9968	0.0425	0.9972	0.0426
	(5,0,...,0)	1.0003	0.0418	1.0011	0.0419
	(0,5,0,...,0)	1.0002	0.0379	1.0012	0.0379

Table 4: Average values (AV) and mean squared errors (MSE) of the estimates of θ based on the MLE and CAE methods for $n = 50$ and different sampling schemes.

m	Censoring Scheme	MLE		CAE	
		AV	MSE	AV	MSE
15	(0,...,0,35)	0.9923	0.0605	0.9918	0.0611
	(35,0,...,0)	0.9928	0.0652	0.9934	0.0657
	(0,35,0,...,0)	0.9944	0.0635	0.9927	0.0639
	(1,...,1,21)	1.0038	0.0699	1.0065	0.0695
	(2,...,2,7)	1.0040	0.0617	1.0035	0.0620
20	(0,...,0,30)	0.9937	0.0485	0.9941	0.0486
	(30,0,...,0)	1.0046	0.0497	1.0061	0.0500
	(0,30,0,...,0)	0.9940	0.0468	0.9936	0.0471
	(1,...,1,11)	1.0049	0.0491	1.0068	0.0493
30	(0,...,0,20)	0.9978	0.0345	0.9987	0.0346
	(20,0,...,0)	1.0048	0.0362	1.0058	0.0363
	(0,20,0,...,0)	1.0040	0.0334	1.0052	0.0335
40	(0,...,0,10)	0.9980	0.0258	0.9990	0.0258
	(10,0,...,0)	1.0038	0.0257	1.0053	0.0257
	(0,10,0,...,0)	0.9973	0.0270	0.9992	0.0270

Table 5: Average confidence lengths and the corresponding coverage probabilities of the Boot-p and the BCa confidence intervals for $n = 20$ and different sampling schemes.

m	Censoring Scheme	Boot-p		BCa	
		Ave. Len.	Cov. Pr.	Ave. Len.	Cov. Pr.
5	(0,0,0,0,15)	1.7219	0.897	2.1498	0.925
	(15,0,0,0,0)	1.7302	0.895	2.0000	0.933
	(0,15,0,0,0)	1.7379	0.898	1.9248	0.929
	(1,1,1,1,11)	1.7285	0.899	2.1431	0.931
	(2,2,2,2,7)	1.7234	0.892	2.1606	0.923
7	(0,...,0,13)	1.4750	0.916	1.8539	0.939
	(13,0,...,0)	1.4586	0.913	1.6452	0.937
	(0,13,0,...,0)	1.4668	0.910	1.6377	0.945
	(2,...,2,1)	1.4262	0.901	1.5764	0.943
	(1,...,1,7)	1.4301	0.914	1.7857	0.945
10	(0,...,0,10)	1.2220	0.919	1.5182	0.946
	(0,10,0,...,0)	1.2155	0.924	1.3338	0.948
	(1,...,1)	1.2317	0.921	1.3567	0.949
15	(0,...,0,5)	0.9990	0.928	1.2047	0.950
	(5,0,...,0)	0.9967	0.925	1.0760	0.952
	(0,5,0,...,0)	1.0007	0.935	1.0785	0.950

Table 6: Average confidence lengths and the corresponding coverage probabilities of the Boot-p and the BCa confidence intervals for $n = 30$ and different sampling schemes.

m	Censoring Scheme	Boot-p		BCa	
		Ave. Len.	Cov. Pr.	Ave. Len.	Cov. Pr.
5	(0,...,0,20)	1.2493	0.927	1.5507	0.937
	(20,0,...,0)	1.2253	0.929	1.3527	0.941
	(0,20,0,...,0)	1.2411	0.928	1.3592	0.936
	(1,...,1,11)	1.2118	0.919	1.4961	0.941
	(2,...,2)	1.2379	0.922	1.3684	0.939
15	(0,...,0,15)	1.0025	0.937	1.2171	0.945
	(15,0,...,0)	1.0278	0.936	1.1076	0.943
	(0,15,0,...,0)	0.9959	0.934	1.0719	0.943
	(1,...,1)	1.0020	0.935	1.0802	0.947
20	(0,...,0,10)	0.8682	0.940	1.0398	0.951
	(10,0,...,0)	0.8759	0.939	0.9302	0.952
	(0,10,0,...,0)	0.8708	0.940	0.9244	0.949
25	(0,...,0,5)	0.7791	0.942	0.9119	0.954
	(5,0,...,0)	0.7684	0.941	0.8084	0.955
	(0,5,0,...,0)	0.7794	0.944	0.8207	0.954

Table 7: Average confidence lengths and the corresponding coverage probabilities of the Boot-p and the BCa confidence intervals for $n = 50$ and different sampling schemes.

m	Censoring Scheme	Boot-p		BCa	
		Ave. Len.	Cov. Pr.	Ave. Len.	Cov. Pr.
15	(0,...,0,35)	1.0017	0.929	1.2226	0.935
	(35,0,...,0)	1.0130	0.925	1.0911	0.939
	(0,35,0,...,0)	1.0057	0.931	1.0816	0.941
	(1,...,1,21)	1.0084	0.929	1.2140	0.937
	(2,...,2,7)	1.0070	0.933	1.1696	0.935
20	(0,...,0,30)	0.8727	0.935	1.0464	0.942
	(30,0,...,0)	0.8720	0.940	0.9239	0.949
	(0,30,0,...,0)	0.8602	0.937	0.9136	0.944
	(1,...,11)	0.8689	0.940	1.0337	0.943
30	(0,...,0,20)	0.7121	0.941	0.8388	0.953
	(20,0,...,0)	0.7145	0.944	0.7475	0.952
	(0,20,0,...,0)	0.7091	0.942	0.7415	0.950
40	(0,...,0,10)	0.6113	0.948	0.7197	0.957
	(10,0,...,0)	0.6181	0.946	0.6394	0.954
	(0,10,0,...,0)	0.6152	0.946	0.6366	0.955

References

Al-Zahrani, B. M. and Gindwan, M. S. (2015). Estimating the parameter of the lindley distribution under progressive type-ii censored data. *Electronic Journal of Applied Statistical Analysis*, 8(1):100–111.

Balakrishnan, N. and Aggarwala, R. (2000). *Progressive censoring: theory, methods, and applications*. Springer Science & Business Media.

Balakrishnan, N. and Asgharzadeh, A. (2005). Inference for the scaled half-logistic distribution based on progressively type-ii censored samples. *Communications in StatisticsTheory and Methods*, 34(1):73–87.

Balakrishnan, N. and Sandhu, R. (1995). A simple simulational algorithm for generating progressive type-ii censored samples. *The American Statistician*, 49(2):229–230.

Cohen, A. C. (1963). Progressively censored samples in life testing. *Technometrics*, 5(3):327–339.

Coppi, R., Gil, M. A., and Kiers, H. A. (2006). The fuzzy approach to statistical analysis. *Computational statistics & data analysis*, 51(1):1–14.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38.

Dubois, D. J. (1980). *Fuzzy sets and systems: theory and applications*, volume 144. Academic press.

Efron, B. and Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*, volume 38. SIAM.

Efron, B. and Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.

Fernández, A. J. (2004). On estimating exponential parameters with general type ii progressive censoring. *Journal of Statistical Planning and Inference*, 121(1):135–147.

Gebhardt, J., Gil, M. A., and Kruse, R. (1998). Fuzzy set-theoretic methods in statistics. In *Fuzzy sets in decision analysis, operations research and statistics*, pages 311–347. Springer.

Grzegorzewski, P. and Hryniewicz, O. (2002). Computing with words and life data. *International Journal of Applied Mathematics and Computer Science*, 12(3):337–346.

Huang, H.-Z., Zuo, M. J., and Sun, Z.-Q. (2006). Bayesian reliability analysis for fuzzy lifetime data. *Fuzzy Sets and Systems*, 157(12):1674–1686.

Ng, H. K. T. and 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.

Pak, A., Parham, G. A., and Saraj, M. (2013). On estimation of rayleigh scale parameter under doubly type-ii censoring from imprecise data. *Journal of Data Science*, 11:305–322.

Pak, A., Parham, G. A., and Saraj, M. (2014). Inferences on the competing risk reliability problem for exponential distribution based on fuzzy data. *Reliability, IEEE*

Transactions on, 63(1):2–12.

Raqab, M. Z. and Madi, M. T. (2011). Inference for the generalized rayleigh distribution based on progressively censored data. *Journal of Statistical Planning and Inference*, 141(10):3313–3322.

Singpurwalla, N. D. and 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 & System Safety*, 94(2):394–403.

Viveros, R. and Balakrishnan, N. (1994). Interval estimation of parameters of life from progressively censored data. *Technometrics*, 36(1):84–91.

Zadeh, L. A. (1968). Probability measures of fuzzy events. *Journal of mathematical analysis and applications*, 23(2):421–427.

Zimmermann, H.-J. (2011). *Fuzzy set theoryand its applications.* Springer Science & Business Media.