Estimating the parameters of lifetime distributions under progressively Type-II censoring from fuzzy data

N. B. KHOOLENJANI AND O. CHATRABGOUN

Abstract

The problem of estimating lifetime distribution parameters under progressively Type-II censoring originated in the context of reliability. But traditionally it is assumed that the available data from this censoring scheme are performed in exact numbers. However, some collected lifetime 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. This paper deals with the estimation of lifetime distribution parameters under progressively Type-II censoring scheme when the lifetime observations are reported by means of fuzzy numbers. A new method is proposed to determine the maximum likelihood estimates of the parameters of interest. The methodology is illustrated with two popular models in lifetime analysis, the Rayleigh and Lognormal lifetime distributions.

Mathematics Subject Classification 2000: 62N02
Additional Key Words and Phrases: Progressive Type-II censoring, Fuzzy lifetime data, Maximum likelihood estimation, EM algorithm.

1 INTRODUCTION

In many life test studies, it is common that the lifetimes of some test units may not be able to be recorded exactly. For example, in Type-II censoring, the test ceases after a predetermined number of failures in order to save time or cost. Furthermore, some test units may have to be removed at different stages in the study for various reasons. This would lead to progressive censoring. Progressive Type-II censored sampling is an important method of obtaining data in lifetime studies. Some of the earlier work on progressive censoring was conducted by [Cohen (1963)], [Mann (1971)], and [Thomas and Wilson (1972)]. Recently, inferences under progressive Type-II censoring have been discussed by several authors. [Viveros and Balakrishnan (1994)] have derived explicit expressions for the best linear unbiased estimators (BLUEs) of the parameters of both one- and two-parameter exponential distributions, and also discussed conditional inference for general location-scale distributions. [Balakrishnan and Kannan (2000)], [Balakrishnan et. al. (2003)] and [Balakrishnan and Asgharzadeh (2005)] have discussed inference procedures based on MLEs and

10.1515/jamsi-2016-0004

approximate MLEs for logistic, normal and half-logistic distributions, respectively. [Pradhan and Kundu (2009)] have obtained maximum likelihood estimates of the parameters of generalized exponential distribution. [Kim and Han (2009)] have considered estimation of the scale parameter of the Rayleigh distribution under general progressive censoring. However, in real situations, lifetime of units sometimes can not be recorded or measured precisely due to machine errors, human errors or some unexpected situations. 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. The lack of precision of lifetime data may be described using fuzzy sets. The classical statistical estimation methods 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. (1991) presented some applications of fuzzy techniques in statistical analysis. [Denoeux (2011)] considered the maximum likelihood estimation based on fuzzy data using the EM algorithm. Pak et al. (2013),(2014) conducted a series of studies to develop the inferential procedures for the lifetime distributions on the basis of fuzzy data. In this paper, our objective is to study the maximum likelihood estimation procedure for the lifetime distribution parameters when the progressively Type-II censored data are reported in the form of fuzzy numbers. In Section 2, we first present in greater detail the problem addressed in this paper. Some preliminary concepts about fuzzy numbers is included in this section. In Section 3, we introduce a generalization of the likelihood function under progressive Type-II censoring and obtain the maximum likelihood estimates in general setting. Two popular models in lifetime analysis, via, Rayleigh and Lognormal distributions, are used to illustrate the proposed method, respectively, in Sections 4 and 5.

2 PROBLEM DESCRIPTION

Consider a reliability experiment in which n independent units are placed on a life-test. Let $X_1,\dots,X_n$ denote the lifetimes of these experimental units. As usual, it is assumed that $X_i,\ i = 1,\dots,n$ are independent and identically distributed with probability density function $f_X(x;\theta)$ and cumulative distribution function $F_X(x;\theta)$, where θ denotes the vector of parameters. 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 $\sum_{i=1}^{m} R_i + m = n$ 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 $x_{1:m:n}, \ldots, x_{m:m:n}$ denote the m completely observed lifetimes. The likelihood function based on this progressively Type-II censored sample is then (see [Balakrishnan and Aggarwala(2000)])

$$L(\theta) = C \prod_{i=1}^{m} f_X(x_{i:m:n};\theta)\left[1 - F_X(x_{i:m:n};\theta)\right]^{R_i}, \tag{1}$$

where

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

The maximum likelihood estimators are those values of θ which maximize (1).

Precisely reported lifetimes are common when data comes from specially designed life tests. In such a case a failure should be precisely defined, and all tested items should be continuously monitored. However, in real situations these test requirements might not be fulfilled. In these cases, it is sometimes impossible to obtain exact observations of lifetime. The obtained lifetime data may be imprecise most of the time. 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 and Prade(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., $B_\lambda(\tilde{x}) = [a_\lambda, b_\lambda]$, $\forall \lambda \in (0,1]$.

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 $\mu_{x_0}(x_0) = 1$. Usually, *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 observations. Therefore, we shall focus on the set of *LR*-type fuzzy numbers.

Definition 2.1. ([Zimmermann(1991), pp.62] 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 2.2. Let $(\mathbb{R}^n, \mathscr{A}, P)$ be a probability space in which $\mathscr{A}$ is the $\sigma-$field of Borel sets in $\mathbb{R}^n$ and P is a probability measure over $\mathbb{R}^n$. Then, A fuzzy event in $\mathbb{R}^n$ is a fuzzy subset $\tilde{A}$ of $\mathbb{R}^n$, whose membership function $\mu_{\tilde{A}}$ is Borel measurable. The probability of a fuzzy event $\tilde{A}$ is defined by:

$$P(\tilde{A}) = \int \mu_{\tilde{A}}(\mathbf{x}) dP.$$

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 and obtain the maximum likelihood estimate (MLE) of θ.

3 MAXIMUM LIKELIHOOD ESTIMATION

Suppose that n independent units are put on a test and that the lifetime distribution of each unit is given by $f(x;\theta)$. 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

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

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 likelihood function can be obtained, using Zadeh's definition of the probability of a fuzzy event ([Zadeh(1968)]), as

$$L_O(\tilde{\mathbf{w}};\theta) = \prod_{i=1} \int \mu_{\tilde{x}_i}(x) f(x;\theta) dx \prod_{i=1}^{m} \prod_{j=1}^{R_i} \int \mu_{\tilde{z}_{ij}}(z) f(z;\theta) dz$$

and the observed-data log likelihood is

$$L(\tilde{\mathbf{w}};\theta) = \sum_{i=1}^{m} \log\left\{ \int \mu_{\tilde{x}_i}(x) f(x;\theta) dx \right\} + \sum_{i=1}^{m} \sum_{j=1}^{R_i} \log\left\{ \int \mu_{\tilde{z}_{ij}}(z) f(z;\theta) dz \right\}. \quad (2)$$

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 estimates (MLE) of the parameters. 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 use the EM algorithm to determine the MLE of 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 $L(\mathbf{W};\theta)$.
The E-step of algorithm requires the calculation of

$$E\left(L(\mathbf{W};\theta) \mid \tilde{\mathbf{w}}, \theta^{(h)}\right), \quad (3)$$

which mainly involves the computation of the conditional expectation of functions of $\mathbf{X}$ and $\mathbf{Z}$ conditional on the observed values $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{z}}$, respectively, and the current value of the parameters. To this end, we need to determine the conditional probability of $\mathbf{X}$ and $\mathbf{Z}$ given $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{z}}$, respectively, from the following formula:

$$f(\mathbf{u} \mid \tilde{\mathbf{u}};\theta^{(h)}) = \frac{\mu_{\tilde{\mathbf{u}}}(\mathbf{u}) f(\mathbf{u};\theta^{(h)})}{\int \mu_{\tilde{\mathbf{u}}}(\mathbf{u}) f(\mathbf{u};\theta^{(h)}) d\mathbf{u}}. \quad (4)$$

In the M-step on the $(h+1)$th iteration of the algorithm, the value of θ which maximizes $E(L(\mathbf{W};\theta) \mid \tilde{w}, \theta^{(h)})$ will be used as the next estimate of $\theta^{(h+1)}$. The MLE of θ can be obtained by repeating the E- and M-step until convergence occurs. It is showed in ([Denoeux (2011)]) that the observed-data log-likelihood $L(\mathbf{W};\theta)$ is not decreased after an EM iteration. Hence, convergence to some value L^* is ensured as long as the sequence $L(\mathbf{W};\theta^{(h)})$ for $h = 0,1,...$ is bounded from above.

4 RAYLEIGH LIFETIME DATA

The Rayleigh distribution is a special case of the two parameter Weibull distribution and a suitable model for life testing studies. [Polovko (1968)] and [Dyer and Whisenand (1973)] demonstrated the importance of this distribution in electro vacuum devices and communication engineering.
The probability density function (p.d.f.) of the Rayleigh distribution is defined as

$$f(y;\sigma) = \frac{y}{\sigma^2}\exp(-\frac{y^2}{2\sigma^2})\ , \qquad y>0,\ \sigma>0. \tag{5}$$

The log-likelihood function based on the complete lifetimes $\mathbf{W}$ becomes proportional to

$$L(\mathbf{W};\sigma) \propto -2n\log\sigma - \frac{1}{2\sigma^2}\left[\sum_{i=1}^{m} x_i^2 + \sum_{i=1}^{m}\sum_{j=1}^{R_i} z_{ij}^2\right]. \tag{6}$$

In the E-step, one needs to compute

$$-2n\log\sigma - \frac{1}{2\sigma^2}\left[\sum_{i=1}^{m} E(X_i^2 \mid \tilde{x}_i;\sigma^{(h)}) + \sum_{i=1}^{m}\sum_{j=1}^{R_i} E(Z_{ij}^2 \mid \tilde{z}_{ij};\sigma^{(h)})\right], \tag{7}$$

where $\sigma^{(h)}$ denotes the current fit of σ at iteration h. The conditional expectations $\alpha_i^{(h)} = E(X_i^2 \mid \tilde{x}_i;\sigma^{(h)})$ and $\beta_{ij}^{(h)} = E(Z_{ij}^2 \mid \tilde{z}_{ij};\sigma^{(h)})$ can be computed using

$$E(U^2 \mid \tilde{u};\sigma^{(h)}) = \frac{\int u^2\mu_{\tilde{u}}(u)f(u;\sigma^{(h)})du}{\int \mu_{\tilde{u}}(u)f(u;\sigma^{(h)})du}. \tag{8}$$

Hence, in the $(h+1)th$ iteration, the values of $\sigma^{(h+1)}$ are computed by the following formula:

$$\hat{\sigma}^{(h+1)} = \left\{\frac{1}{2n}\left(\sum_{i=1}^{m}\alpha_i^{(h)} + \sum_{i=1}^{m}\sum_{j=1}^{R_i}\beta_{ij}^{(h)}\right)\right\}^{1/2}. \tag{9}$$

In order to assess the accuracy of the MLEs computed through the procedure described above, we have carried out a simulation study. First, for different choices of n , m , σ and $(R_1,...,R_m)$, we have generated progressively censored sample $x_1,...,x_m$ from Rayleigh distribution using the method proposed by [Balakrishnan and Sandhu (1995)]. Then we have defined fuzzy numbers $\tilde{x}_1,...,\tilde{x}_m$ with the

Table I. The average values (AV) and mean squared errors (MSE) for the MLE of σ for different sample sizes and different sampling schemes.

n	m	Censoring Scheme	$\sigma=1$		$\sigma=2$	
			AV	MSE	AV	MSE
20	5	(0,0,0,0,15)	0.9793	0.0497	1.9548	0.2108
20	5	(15,0,0,0,0)	0.9777	0.0474	1.9520	0.1994
20	5	(0,15,0,0,0)	0.9836	0.0511	1.9561	0.2049
20	10	(0,...,0,10)	0.9865	0.0258	1.9634	0.0996
20	10	(10,0,...,0)	0.9849	0.0234	1.9754	0.0963
20	10	(0,10,0,...,0)	0.9934	0.0231	1.9735	0.0911
20	15	(0,...,0,5)	0.9923	0.0162	1.9942	0.0662
20	15	(5,0,...,0)	0.9909	0.0174	1.9782	0.0662
20	15	(0,5,0,...,0)	0.9934	0.0160	1.9751	0.0672
30	10	(0,...,0,20)	0.9830	0.0239	1.9823	0.0945
30	10	(20,0,...,0)	1.0138	0.0233	1.9714	0.0923
30	10	(0,20,0,...,0)	0.9799	0.0227	1.9853	0.0910
30	15	(0,...,0,15)	0.9890	0.0154	1.9966	0.0642
30	15	(15,0,...,0)	0.9890	0.0166	1.9872	0.0627
30	15	(0,15,0,...,0)	0.9886	0.0152	1.9873	0.0661
30	20	(0,...,0,10)	0.9947	0.0130	1.9944	0.0519
30	20	(10,0,...,0)	0.9933	0.0127	1.9949	0.0509
30	20	(0,10,0,...,0)	1.0034	0.0131	1.9920	0.0478
50	20	(0,...,0,30)	0.9916	0.0120	1.9892	0.0491
50	20	(30,0,...,0)	0.9880	0.0122	1.9789	0.0486
50	20	(0,30,0,...,0)	0.9972	0.0124	1.9837	0.0472
50	25	(0,...,0,25)	0.9927	0.0102	1.9967	0.0404
50	25	(25,0,...,0)	0.9936	0.0099	1.9893	0.0407
50	25	(0,25,0,...,0)	1.0012	0.0100	1.9869	0.0407
50	30	(0,...,0,20)	0.9956	0.0074	2.0021	0.0330
50	30	(20,0,...,0)	0.9939	0.0084	1.9901	0.0345
50	30	(0,20,0,...,0)	0.9991	0.0080	1.9913	0.0315

corresponding membership functions

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

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 σ, using the iterative algorithm (9). We have used the initial estimate to be $\sigma^{(0)}=(\frac{1}{2m}\sum_{i=1}^{m}x_i^2)^{1/2}$. The iterative process stops when the relative change of the log-likelihood becomes less than 10^{-6}. The average values (AV) and mean squared errors(MSE) of the estimates based on 1000 replication are presented in Table 1. From this table we observe that, as the sample size increases or the effective sample size increases, the performances of the MLEs in terms of mean squared errors become better. 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.

EXAMPLE 1: A progressively Type-II censored sample from the data on the failure times of 23 ball bearings in endurance test is used to demonstrate the above estimation procedure. For this data set, [Raqab and Madi (2002)] indicated that the one-parameter Rayleigh distribution provides a satisfactory fit. The data are presented in Table 2. But, in practice 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. So, such data may be reported as imprecise quantities. Assume that imprecision of the failure times of ball bearings is formulated by fuzzy numbers $\tilde{x}_i = (h_i, x_i)$, where $h_i = 0.005x_i$, $i = 1, ..., 16$, with membership functions

$$\mu_{\tilde{x}_i}(x) = \begin{cases} \frac{x-(x_i-h_i)}{h_i} & x_i - h_i \le x \le x_i \\ 0 & x > x_i \end{cases} \quad , i = 1, ..., 16.$$

From these fuzzy data and using the starting value $\sigma^{(0)} = (\frac{1}{32}\sum_{i=1}^{16} x_i^2)^{1/2} = 40.0687$, which is the estimate of the parameter computed over the centers of each fuzzy numbers, the final MLE of σ is found from (9) to be $\hat{\sigma} = 48.8245$. Fig. 1 shows a plot of the observed-data log-likelihood as a function of $\sigma^{(h)}$. We can check that the MLE corresponds in this case to a global maximum of the observed-data log-likelihood.

Table II. Progressively censored sample for Example 2

i	1	2	3	4	5	6	7	8
x_i	17.88	28.92	33.00	41.52	42.12	45.60	48.48	51.84
R_i	2	0	0	1	0	0	1	0
i	9	10	11	12	13	14	15	16
x_i	51.96	54.12	55.56	67.80	68.64	68.88	84.12	93.12
R_i	0	0	1	0	0	0	0	2

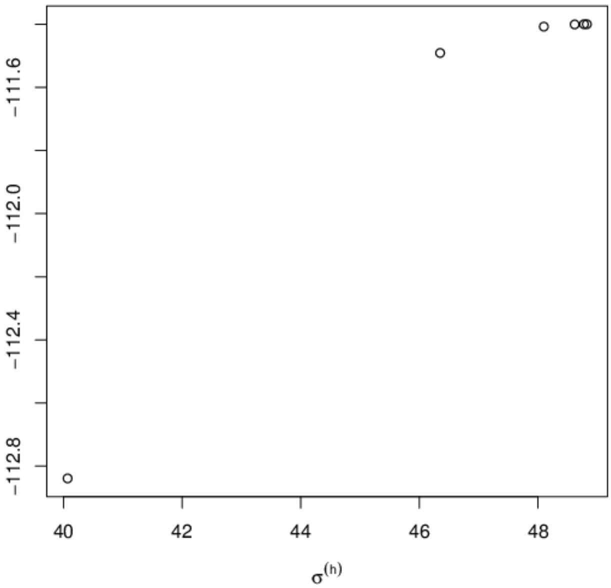

Fig. 1. Plot of the observed-data log-likelihood as a function of $\sigma^{(h)}$

5 LOGNORMAL LIFETIME DATA

Lognormal distribution is another commonly used lifetime distribution model in lifetime data analysis. Let X be the original lifetime variable that follows a Lognormal distribution with parameters μ and σ. The density of X is given by

$$f(y;\mu,\sigma) = \frac{1}{\sqrt{(2\pi)}\sigma y}\exp\left[-\frac{1}{2\sigma^2}(\log y-\mu)^2\right] \qquad ,y>0, \tag{10}$$

where μ and σ are the scale and shape parameters, respectively.
The log-likelihood function based on the complete lifetimes is proportional to

$$L(\mathbf{W};\mu,\sigma) \propto -n\log(\sigma) - \frac{1}{2\sigma^2}\left[\sum_{i=1}^{m}(\log x_i-\mu)^2 + \sum_{i=1}^{m}\sum_{j=1}^{R_i}(\log z_{ij}-\mu)^2\right]. \tag{11}$$

In the E-step, one requires to compute

$$n\log(\sigma) - \frac{1}{2\sigma^2}\left[\sum_{i=1}^{m}E((\log X_i)^2 \mid \tilde{x}_i;\mu^{(h)},\sigma^{(h)}) + \sum_{i=1}^{m}\sum_{j=1}^{R_i}E((\log Z_{ij})^2 \mid \tilde{z}_{ij};\mu^{(h)},\sigma^{(h)})\right]$$

$$+\frac{\mu}{\sigma^2}\left[\sum_{i=1}^{m} E(\log X_i \mid \tilde{x}_i;\mu^{(h)},\sigma^{(h)})+\sum_{i=1}^{m}\sum_{j=1}^{R_i} E(\log Z_{ij} \mid \tilde{z}_{ij};\mu^{(h)},\sigma^{(h)})\right]-\frac{n\mu^2}{2\sigma^2}. \quad (12)$$

The conditional expectations $\alpha_i^{(h)} = E((\log X_i)^2 \mid \tilde{x}_i;\mu^{(h)},\sigma^{(h)})$, $\gamma_i^{(h)} = E(\log X_i \mid \tilde{x}_i;\mu^{(h)},\sigma^{(h)})$, $\beta_{ij}^{(h)} = E((\log Z_{ij})^2 \mid \tilde{z}_{ij};\mu^{(h)},\sigma^{(h)})$ and $\eta_{ij}^{(h)} = E(\log Z_{ij} \mid \tilde{z}_{ij};\mu^{(h)},\sigma^{(h)})$ can be computed using

$$E((\log U)^2 \mid \tilde{u};\mu^{(h)},\sigma^{(h)}) = \frac{\int (\log u)^2 \mu_{\tilde{u}}(u) f(u;\mu^{(h)},\sigma^{(h)})du}{\int \mu_{\tilde{u}}(u) f(u;\mu^{(h)},\sigma^{(h)})du}, \quad (13)$$

and

$$E(\log U \mid \tilde{u};\mu^{(h)},\sigma^{(h)}) = \frac{\int (\log u)\, \mu_{\tilde{u}}(u) f(u;\mu^{(h)},\sigma^{(h)})du}{\int \mu_{\tilde{u}}(u) f(u;\mu^{(h)},\sigma^{(h)})du}. \quad (14)$$

From the usual results for complete data maximum likelihood estimation for lognormal distribution, the explicit formulas for the MLE of μ and σ are

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\log w_i,$$

$$\hat{\sigma} = \left[\frac{1}{n}\sum_{i=1}^{n}(\log w_i - \hat{\mu})^2\right]^{1/2}.$$

Therefore, in the $(h+1)$th iteration, the value of $\mu^{(h+1)}$ and $\sigma^{(h+1)}$ are computed by the following formulas:

$$\hat{\mu}^{(h+1)} = \frac{1}{n}\left[\sum_{i=1}^{m}\gamma_i^{(h)} + \sum_{i=1}^{m}\sum_{j=1}^{R_i}\eta_{ij}^{(h)}\right], \quad (15)$$

$$\hat{\sigma}^{(h+1)} = \left\{\frac{1}{n}\left[\sum_{i=1}^{m}\alpha_i^{(h)} + \sum_{i=1}^{m}\sum_{j=1}^{R_i}\beta_{ij}^{(h)}\right] - \left(\hat{\mu}^{(h+1)}\right)^2\right\}^{1/2}. \quad (16)$$

Table III. Simulated progressively censored sample from standard lognormal distribution

i	1	2	3	4	5	6	7	8
x_i	0.2721	0.2910	0.2929	0.3882	0.5594	0.5831	0.7041	0.8272
R_i	1	0	0	1	0	0	1	0
i	9	10	11	12	13	14	15	
x_i	0.8454	0.8963	1.1084	1.7867	2.0148	2.2213	4.2340	
R_i	0	0	0	0	0	0	2	

EXAMPLE 2: To illustrate experimentally the method presented in this section, we perform the following experiment. We first generated a progressively Type-II censored sample of size $m = 15$ from standard lognormal distribution. The data are presented in Table 3. Each realization of lifetimes was fuzzified by fuzzy numbers $\tilde{x}_1,...,\tilde{x}_m$ with the corresponding membership function

$$\mu_{\tilde{x}_i}(x) = \begin{cases} \exp\left(-(x_i - x)^2\right) & x \leq x_i \\ \exp\left(-(x - x_i)^2\right) & x > x_i \end{cases} \quad , i = 1,...,m.$$

Since the mean and standard deviation of the 15 observed sample points equal to -0.2015 and 0.8192, respectively, thus we can put $\mu^{(0)} = -0.2015$ and $\sigma^{(0)} = 0.8192$ as the starting values of the algorithm. After a few iterations, the estimates converge to the values $\hat{\mu} = 0.1276$ and $\hat{\sigma} = 1.0161$.

6 CONCLUSION

Although the maximum likelihood estimation method based on the progressively Type-II censored data has been studied extensively, traditionally it is assumed that the data available are performed in exact numbers. In real world situations, however, we deal with non-precise (fuzzy) data. Therefore, the conventional procedures used for estimating the unknown parameters of lifetime distributions will have to be adopted to the new situation. In this paper we have proposed a new method for obtaining maximum likelihood estimates of lifetime distribution parameters under progressively Type-II censoring scheme when the lifetime observations are fuzzy numbers. Two popular lifetime models, Rayleigh and Lognormal distributions, have been used to demonstrate how the the proposed method works. For the two cases, the subsequent guesses of the parameters are in explicit expression. The study of the applicability of the proposed approach in estimating the parameters of lifetime distributions under the other censoring schemes such as Hybrid Type-II and Hybrid progressive Type-II censoring are possible topics for further research.

ACKNOWLEDGEMENTS

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

REFERENCES

Balakrishnan, N., and Aggarwala, R. (2000), Progressive Censoring: Theory, Methods and Applications.Birkhauser, Boston.

Balakrishnan, N., and Asgharzadeh, A. (2005). Inference for the scaled half-logistic distribution based on progressively Type II censored samples. Communications in Statistics-Theory and Methods, 34, 73-87.

Balakrishnan, N., and Kannan, N. (2000). Point and interval estimation for the parameters of the logistic distribution based on progressively Type-II censored samples. In N. Balakrishnan, and C. R. Rao (Eds.), Handbook of statistics: Vol. 20 (pp. 431-456).

Balakrishnan, N., Kannan, N., Lin, C. T., and Ng, H. K. T. (2003). Point and interval estimation for the normal distribution based on progressively Type-II censored samples. IEEE Transactions on Reliability, 52, 90-95.

Balakrishnan, N. and Sandhu, R. A. (1995). A simple algorithm for generating progressively Type-II censored samples. The American Statistician, 49(2), 229-230.

Cohen, A. C. (1963). Progressively censored samples in life testing. Tecnometrics, Volume 5, 327-329.

Coppi, R., Gil, M.A. and Kiers, H.A.L., (2006). The fuzzy approach to statistical analysis. Computational Statistics and Data Analysis, 51(1), 114.

Denoeux, T. (2011). Maximum likelihood estimation from fuzzy data using the EM algorithm, Fuzzy Sets and Systems. 183(1), 72-91.

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, 39, 1-38.

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

Dyer, D. D. and Whisenand, C. W. (1973). Best linear estimator of the parameter of the Rayleigh distribution-Part I: Small sample theory for censored order statistics. IEEE Transactions on Reliability, 22, 27-34.

Gebhardt, J., Gil M.A. and Kruse R., (1998). Fuzzy set-theoretic methods in statistics, in:R.Slowinski(Ed.), Fuzzy Sets in Decision Analysis, Operations Research and Statistics, Kluwer Academic Publishers, Boston, pp.311-347.

Huang, H., Zuo, M. and Sun, Z., (2006). Bayesian reliability analysis for fuzzy lifetime data. Fuzzy Sets and Systems, 157, 16741686.

Kim C. and Han K. (2009). Estimation of the scale parameter of the Rayleigh distribution under general progressive censoring, Journal of the Korean Statistical Society, 38, 239-246.

Mann, N. R. (1971). Best linear invariant estimator for Weibull parameters under progressive censoring. Technometrics, 13, 521-533.

Pak, A., Parham, G.H. and Saraj, M., (2013). On estimation of Rayleigh scale parameter under doubly Type-II censoring from imprecise data. Journal of Data Science, 11, 303-320.

Pak, A., Parham, G.H. and Saraj, M., (2014). Inferences on the Competing Risk Reliability Problem for Exponential Distribution Based on Fuzzy Data. IEEE Transactions on reliability, 63(1), 2-13.

Polovko, A. M. (1968), Fundamentals of Reliability Theory. New York: Academic Press.

Pradhan B., and Kundu D. (2009). On progressively censored generalized exponential distribution, Test, 18, 497-515.

Raqab, M. Z. and Madi, M. T. (2002). Bayesian prediction of the total time on test using doubly censored Rayleigh data. Journal of Statistical Computation and Simulation, 72, 781-789.

Singpurwalla, N.D. and Booker, J.M. (2004). Membership functions and probability measures of fuzzy sets. Journal of the American Statistical Association, 99(467), 867877.

Thomas D. R., Wilson W. M. (1972) Linear order statistic estimation for the two-parameter Weibull and extreme value distributions from Type-II progressively censored samples. , Technometrics, 14 , 679-691.

Viveros, R., and Balakrishnan, N. (1994). Interval estimation of life characteristics from progressively censored data. Technometrics, 36, 84-91.

Zadeh, L. A. (1968). Probability measures of fuzzy events, Journal of Mathematical Analysis and Applications10, 421-427.

Zimmermann, H. J. (1991). Fuzzy set teory and its application, Kluwer, Dordrecht.

Nayereh Bagheri Khoolenjani
Department of Statistics,
University of Isfahan,
Isfahan, 81746-73441, Iran

Omid Chatrabgoun
Department of Statistics,
Malayer University,
Malayer, Iran