Brazilian Journal of Probability and Statistics
2017, Vol. 31, No. 3, 502–515
DOI: 10.1214/16-BJPS321

Statistical inference for the parameter of Lindley distribution based on fuzzy data

Abbas Pak

Shahrekord University

Abstract. In many practical situations, we face data which are not only random but vague as well. To deal with these two types of uncertainties, it is necessary to incorporate fuzzy concept into statistical technique. In this paper, we investigate the maximum likelihood estimation and Bayesian estimation for Lindley distribution when the available observations are reported in the form of fuzzy data. We employ the EM algorithm to determine the maximum likelihood estimate (MLE) of the parameter and construct approximate confidence interval by using the asymptotic normality of the MLE. In the Bayesian setting, we use an approximation based on the Laplace approximation as well as a Markov Chain Monte Carlo technique to compute the Bayes estimate of the parameter. In addition, the highest posterior density credible interval of the unknown parameter is obtained. Extensive simulations are performed to compare the performances of the different proposed methods.

1 Introduction

The Lindley distribution specified by the probability density function (p.d.f.)

$$f(x;\theta) = \frac{\theta^2}{1+\theta}(1+x)\exp(-\theta x); \qquad x > 0; \theta > 0, \tag{1.1}$$

and cumulative distribution function (c.d.f.)

$$F(x) = 1 - \frac{1+\theta+\theta x}{1+\theta}\exp(-\theta x); \qquad x > 0, \theta > 0 \tag{1.2}$$

was introduced by Lindley (1958) as a new distribution useful to analyze lifetime data especially in applications modeling stress-strength reliability. From now on Lindley distribution with parameter θ will be denoted by Lindley(θ). Several authors have addressed inferential issues for Lindley distribution parameter based on complete and censored samples; for example, Ghitany, Al-Mutairi and Nadarajah (2008) studied the properties of the Lindley distribution under a carefully mathematical treatment. They also showed in a numerical example that the Lindley distribution gives better modeling than the one based on the Exponential distribution. Gupta et al. (2011) discussed reliability estimation in Lindley distribution

Key words and phrases. Fuzzy data analysis, maximum likelihood principle, Bayesian estimation, asymptotic confidence interval.

Received February 2016; accepted May 2016.

with progressively type II right censored sample. Mazucheli and Achcarb (2011) considered a competing risks model when the causes of failure follow the Lindley distribution. Gupta and Singh (2012) discussed the classical and Bayesian analysis of the hybrid censored lifetime data under the assumption that the lifetimes follow Lindley distribution.

The above inference techniques for estimating the parameter of Lindley distribution are based on precisely defined crisp data. However, in many practical situations we face data which are not only random but vague as well. Randomness involves only uncertainties in the outcomes of an experiment; vagueness, on the other hand, involves uncertainties in the meaning of the data. As an example, consider a case study on the ball bearing manufacturing process that focuses on the lifetime of ball bearings. A ball bearing may work perfectly over a certain period but be braking for some time, and finally be unusable at a certain time. Therefore, the lifetime of each ball bearing may be reported by means of vague statements such as "approximately lower than 17", "approximately 33 to 35", "approximately 47 to 90 but near to 90", "approximately higher than 125", and so on. In this situation, randomness occurs when the ball bearings are selected at random and vagueness is due to limited ability of the observer to describe the lifetime of ball bearings using numbers. To deal with both types of uncertainties—*randomness* and *vagueness*, it is necessary to incorporate fuzzy concept into statistical technique.

In recent years, many papers on generalization of classical statistical methods to analysis of fuzzy data have been published. Wu (2004) discussed the Bayesian estimation on lifetime data under fuzzy environments. Gil, López-Diaz and Ralescu (2006) presented a backward analysis on the interpretation, modelling and impact of the concept of fuzzy random variable. Viertl (2006) studied generalization of classical statistical inference methods for univariate fuzzy data. Zarei et al. (2012) considered the Bayesian estimation of failure rate and mean time to failure based on vague set theory in the case of complete and censored data sets. Very recently, Pak, Parham and Saraj (2013) conducted a series of studies to develop the inferential procedures for the lifetime distributions on the basis of fuzzy data.

In this paper, we study different estimation procedures for the parameter of Lindley distribution when the available information from the experiments are described by means of fuzzy numbers. We first describe the construction of fuzzy data from imprecise observations, and then discuss the computation of maximum likelihood estimate of the parameter θ. Based on fuzzy data, there is no closed form for the MLE; therefore, we employ EM algorithm to determine the maximum likelihood estimate. We also construct the approximate confidence interval of the unknown parameter by using the asymptotic distribution of the MLE. We further consider the Bayesian inference of the parameter of Lindley distribution. Since the Bayes estimate cannot be obtained in explicit form, we provide an approximation, namely Tierney and Kadane's approximation, as well as a Markov Chain Monte Carlo (MCMC) technique to compute the Bayes estimate and construct the highest posterior density (HPD) credible interval of the parameter θ.

The rest of this paper is organized as follows. In Section 2, we obtain the maximum likelihood estimate of the parameter θ and also construct the approximate confidence interval by using asymptotic normality of the MLE. The Bayesian analyses are provided in Section 3. A Monte Carlo simulation study is presented in Section 4, which provides a comparison of all estimation procedures developed in this paper. In the following, we first review the main definitions of fuzzy sets and some of the formula used in this paper.

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 $A = \{x, x\text{'s age} \le 40 \text{ yr}\}$ 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 $\tilde{A}$ 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 $\tilde{A}$. A fuzzy event in $\mathcal{X}$ is a fuzzy subset $\tilde{A}$ of $\mathcal{X}$, whose membership function $\mu_{\tilde{A}}$ is Borel measurable. As an example, consider an opinion poll during which a number of individuals are questioned on their perception of the relative length of different line segments with respect to a fixed longer segment that was used as a standard for comparison. The answer given by the individual may be vague statement such as "The length of the line segment is approximately 30 to 40". A fuzzy approach lies in expressing this observation as a fuzzy event $\tilde{A}$ such as that defined, for instance, by the membership function

$$\mu_{\tilde{A}}(x) = \begin{cases} \dfrac{x-20}{10}, & 20 \le x \le 30, \\ 1, & 30 \le x \le 40, \\ \dfrac{50-x}{10}, & 40 \le x \le 50, \\ 0, & \text{otherwise} \end{cases}$$

(see Figure 1).

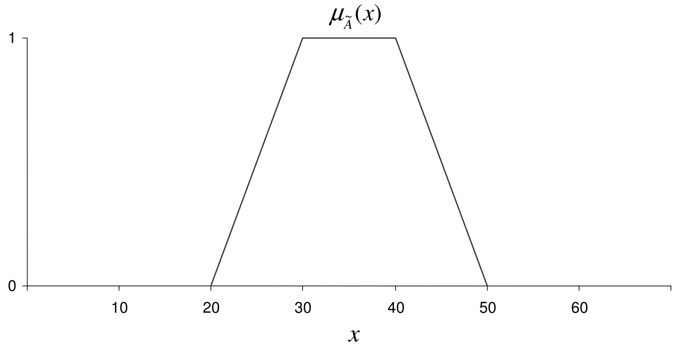

Figure 1 *Fuzzy approach of the imprecise observation "approximately* 30 *to* 40*".*

Definition 1 (Zadeh (1968)). Let $(\mathbb{R}^n, \mathcal{A}, P)$ be a probability space in which $\mathcal{A}$ is the σ-field of Borel sets in $\mathbb{R}^n$ and P is a probability measure over $\mathbb{R}^n$. The probability of a fuzzy event $\tilde{A}$ is defined by:

$$P(\tilde{A}) = \int \mu_{\tilde{A}}(\mathbf{x})\, dP. \tag{1.3}$$

Two fuzzy events $\tilde{A}$ and $\tilde{B}$ in the probability space $(\mathbb{R}^n, \mathcal{A}, P)$ are said to be independent if $P(\tilde{A}\tilde{B}) = P(\tilde{A})P(\tilde{B})$ where $\tilde{A}\tilde{B}$ is the fuzzy subset of $\mathbb{R}^n$ with membership function $\mu_{\tilde{A}\tilde{B}}(\mathbf{x}) = \mu_{\tilde{A}}(\mathbf{x})\mu_{\tilde{B}}(\mathbf{x})$.

The conditional probability of $\tilde{A}$ given $\tilde{B}$ is defined by

$$P(\tilde{A}|\tilde{B}) = \frac{P(\tilde{A}\tilde{B})}{P(\tilde{B})}$$

provided $P(\tilde{B}) > 0$.

In particular, assume that P is the probability distribution of a continuous random variable $\mathbf{X}$ with probability density function $g(\mathbf{x})$. For a crisp subset A and a fuzzy subset $\tilde{B}$, the above conditional probability becomes

$$P(A|\tilde{B}) = \frac{\int \mu_A(\mathbf{x})\mu_{\tilde{B}}(\mathbf{x})g(\mathbf{x})\, d\mathbf{x}}{\int \mu_{\tilde{B}}(\mathbf{u})g(\mathbf{u})\, d\mathbf{u}} = \int_A \frac{\mu_{\tilde{B}}(\mathbf{x})g(\mathbf{x})\, d\mathbf{x}}{\int \mu_{\tilde{B}}(\mathbf{u})g(\mathbf{u})\, d\mathbf{u}}.$$

The conditional density of $\mathbf{X}$ given $\tilde{B}$ can thus be defined as

$$g(\mathbf{x}|\tilde{B}) = \frac{\mu_{\tilde{B}}(\mathbf{x})g(\mathbf{x})}{\int \mu_{\tilde{B}}(\mathbf{u})g(\mathbf{u})\, d\mathbf{u}}. \tag{1.4}$$

Definition 2 (See Shafiq and Viertl (2014)). 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 $\mu_{\tilde{x}}(\cdot)$. Fuzzy numbers satisfy the following constraints:

(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 λ-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., $B_\lambda(\tilde{x}) = [a_\lambda, b_\lambda]$, $\forall \lambda \in (0, 1]$.

Examples of membership functions to characterize fuzzy numbers are triangular and trapezoidal fuzzy numbers. A triangular fuzzy number is defined as $\tilde{x} = (\nu, \omega, \delta)$ with the corresponding membership function

$$\mu_{\tilde{x}}(x) = \begin{cases} \dfrac{x - \nu}{\omega - \nu}, & \nu \leq x \leq \omega, \\ \dfrac{\delta - x}{\delta - \omega}, & \omega \leq x \leq \delta, \\ 0, & \text{otherwise.} \end{cases}$$

The trapezoidal fuzzy number can be defined as $\tilde{x} = (\delta, \nu, \omega, \eta)$ with the corresponding membership function

$$\mu_{\tilde{x}}(x) = \begin{cases} \dfrac{x - \delta}{\nu - \delta}, & \delta \leq x \leq \nu, \\ 1, & \nu \leq x \leq \omega, \\ \dfrac{\eta - x}{\eta - \omega}, & \omega \leq x \leq \eta, \\ 0, & \text{otherwise.} \end{cases}$$

Example 1. Consider a life-testing experiment in which n identical light emitting diodes (LEDs) are placed on test. A tested LED may be considered as failed, or—strictly speaking—as nonconforming, 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 a failure. So, he/she provides an interval $[\nu_i, \omega_i]$ which certainly contains the lifetime of LED i and an interval $[\delta_i, \eta_i]$ which contains highly plausible values for that lifetime. This information may be encoded as a trapezoidal fuzzy number $\tilde{x}_i = (\delta_i, \nu_i, \omega_i, \eta_i)$ with the corresponding membership function

$$\mu_{\tilde{x}_i}(x_i) = \begin{cases} \dfrac{x_i - \delta_i}{\nu_i - \delta_i}, & \delta_i \leq x_i \leq \nu_i, \\ 1, & \nu_i \leq x_i \leq \omega_i, \\ \dfrac{\eta_i - x_i}{\eta_i - \omega_i}, & \omega_i \leq x_i \leq \eta_i, \\ 0, & \text{otherwise.} \end{cases}$$

In this case randomness arises from the selection of LEDs as well as environmental factors which influence the perception by the observer. In contrast, fuzziness arises from the meaning of the reported failure times.

2 Maximum likelihood estimation

Suppose that n identical units are placed on a life test with the corresponding lifetimes $X_1, \dots, X_n$. It is assumed that these variables are independent and identically distributed as Lindley(θ). Let $\mathbf{X} = (X_1, \dots, X_n)$ denotes the vector of lifetimes. If a realization $\mathbf{x}$ of $\mathbf{X}$ was known exactly, we could obtain the complete-data likelihood function as

$$\ell(\theta) = \left(\frac{\theta^2}{1+\theta}\right)^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \prod_{i=1}^n (1 + x_i). \tag{2.1}$$

Consider the situation where the available information about $\mathbf{x}$ can not be exactly perceived, but that rather it may be assimilated with fuzzy numbers $\tilde{x}_1, \dots, \tilde{x}_n$ with the corresponding membership functions $\mu_{\tilde{x}_1}(\cdot), \dots, \mu_{\tilde{x}_m}(\cdot)$. Then, by using expression (1.3), we can obtain the likelihood function of θ as

$$L(\theta) = \prod_{i=1}^n \int \frac{\theta^2}{1+\theta}(1+x)\exp(-\theta x)\mu_{\tilde{x}_i}(x)\,dx \tag{2.2}$$

and the corresponding log-likelihood function $L^*(\theta) = \log L(\theta)$ becomes

$$L^*(\theta) = 2n\log\theta - n\log(1+\theta) + \sum_{i=1}^n \log \int (1+x)\exp(-\theta x)\mu_{\tilde{x}_i}(x)\,dx. \tag{2.3}$$

The maximum likelihood estimate of the parameter θ can be computed as any value maximizing the observed-data log-likelihood (2.3). Equating the derivative of the log-likelihood L^* with respect to θ to zero, we have

$$\frac{\partial}{\partial\theta}L^*(\theta) = \frac{2n}{\theta} - \frac{n}{1+\theta} - \sum_{i=1}^n \frac{\int x(1+x)\exp(-\theta x)\mu_{\tilde{x}_i}(x)\,dx}{\int (1+x)\exp(-\theta x)\mu_{\tilde{x}_i}(x)\,dx} = 0. \tag{2.4}$$

To achieve estimation via ML method, it is not easy to solve the equation (2.4), directly. However, similar to the proof of Theorem 1 in Pak, Parham and Saraj (2013), one can easily check that the likelihood equation (2.4) has a unique solution. In this case, an iterative numerical search can be used to obtain the MLE. Therefore, in the following, we describe the EM algorithm to determine the maximum likelihood estimate of the parameter θ.

The EM algorithm approaches the problem of maximizing the observed-data log likelihood $L^*(\theta)$ by proceeding iteratively with the complete-data log likelihood

$$\log\ell(\theta) = 2n\log\theta - n\log(1+\theta) + \sum_{i=1}^n \log(1+x_i) - \theta\sum_{i=1}^n x_i.$$

Each iteration of the algorithm involves two steps called the expectation step (E-step) and the maximization step (M-step). The E-step requires the calculation of

$$
\begin{aligned}
&E_{\theta^{(h)}}\big[\log \ell(\theta)|\tilde{x}_1,\ldots,\tilde{x}_n\big] \\
&\quad = c(\mathbf{Y}) + 2n\log\theta - n\log(1+\theta) - \theta\left(\sum_{i=1}^{n} E_{\theta^{(h)}}[X_i|\tilde{x}_i]\right),
\end{aligned}
\tag{2.5}
$$

where $c(\mathbf{Y})$ does not depend on θ, and $\theta^{(h)}$ denotes the current fit of θ at iteration h. By using Relation (1.4), the conditional expectation $E_{\theta^{(h)}}[X_i|\tilde{x}_i]$ for $i = 1,\ldots,n$, can be computed as:

$$
E_{\theta^{(h)}}[X_i|\tilde{x}_i[= \frac{\int x(1+x)\exp(-\theta x)\mu_{\tilde{x}_i}(x)\,dx}{\int (1+x)\exp(-\theta x)\mu_{\tilde{x}_i}(x)\,dx}, \qquad i = 1,\ldots,n. \tag{2.6}
$$

The M-step then consists in finding $\theta^{(h+1)}$ which maximizes $E_{\theta^{(h)}}[\log \ell(\theta)|\tilde{x}_1,\ldots,\tilde{x}_n]$. This is easily achieved by solving the likelihood equation. From

$$
\frac{\partial}{\partial\theta} E_{\theta^{(h)}}\big[\log \ell(\theta)|\tilde{x}_1,\ldots,\tilde{x}_n\big] = 0,
$$

we get

$$
\hat{\theta}^{(h+1)} = \frac{-(\alpha^{(h)}-1) + \sqrt{(\alpha^{(h)}-1)^2 + 8\alpha^{(h)}}}{2\alpha^{(h)}}, \tag{2.7}
$$

where

$$
\alpha^{(h)} = \frac{1}{n}\sum_{i=1}^{n} E_{\theta^{(h)}}[X_i|\tilde{x}_i].
$$

The MLE of θ can be obtained by repeating the E- and M-steps until convergence.

Once the maximum likelihood estimate of θ is obtained, we can use the asymptotic normality of the MLEs to compute the approximate $100(1-\alpha)\%$ confidence interval for the parameter as follows:

$$
\hat{\theta} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{1}{I(\hat{\theta})}}. \tag{2.8}
$$

Here, $z_{\frac{\alpha}{2}}$ is an upper percentile of the standard normal variate and the observed Fisher information $I(\hat{\theta})$ is obtained as

$$
I(\hat{\theta}) = -\frac{\partial^2}{\partial\theta^2} L^*(\theta)|_{\theta=\hat{\theta}},
$$

where

$$
\begin{aligned}
\frac{\partial^2}{\partial\theta^2} L^*(\theta) = -\frac{2n}{\theta^2} + \frac{n}{(1+\theta)^2} + \sum_{i=1}^{n}\Bigg\{ &\frac{\int x^2(1+x)\exp(-\theta x)\mu_{\tilde{x}_i}(x)\,dx}{\int (1+x)\exp(-\theta x)\mu_{\tilde{x}_i}(x)\,dx} \\
&- \left[\frac{\int x(1+x)\exp(-\theta x)\mu_{\tilde{x}_i}(x)\,dx}{\int (1+x)\exp(-\theta x)\mu_{\tilde{x}_i}(x)\,dx}\right]^2\Bigg\}.
\end{aligned}
$$

3 Bayesian estimation

In this section, we describe the Bayes estimate of the unknown parameter as well as the corresponding highest posterior density credible interval. In the Bayesian estimation, unknown parameter is assumed to behave as random variable with distribution commonly known as prior probability distribution. Therefore, as a conjugate prior for θ, we take the Gamma(a, b) density of the form

$$\pi(\theta) \propto \theta^{a-1} \exp(-\theta b), \qquad \theta > 0, \tag{3.1}$$

where $a > 0$ and $b > 0$. By combining (2.2) with (3.1), the joint density function of the data and θ becomes

$$\pi(\text{data}, \theta) \propto \theta^{2n+a-1} \frac{\exp(-\theta b)}{(1+\theta)^n} \prod_{i=1}^{n} \int (1+x) \exp(-\theta x) \mu_{\tilde{x}_i}(x)\, dx. \tag{3.2}$$

Thus, the posterior density function of θ given the data can be obtained as

$$\pi(\theta|\text{data}) = \frac{\pi(\text{data}, \theta)}{\int_0^\infty \pi(\text{data}, \theta)\, d\theta}. \tag{3.3}$$

It is well known that the Bayes estimate of any function of θ, say $h(\theta)$, under squared error loss function is the posterior mean which is obtained by

$$\int_0^\infty \pi(\theta|\text{data}) h(\theta)\, d\theta. \tag{3.4}$$

The Eqs. (3.3) and (3.4) do not simplified to nice closed forms due to the complex form of the likelihood function. Therefore, we use Tierney and Kadane's approximation as well as MCMC method for computing the Bayes estimate of θ.

3.1 Tierney and Kadane's approximation

We first rewrite the expression in (3.4) as

$$\int_0^\infty \pi(\theta|\text{data}) h(\theta)\, d\theta = \frac{\int_0^\infty e^{nF^*(\theta)}\, d\theta}{\int_0^\infty e^{nF(\theta)}\, d\theta}, \tag{3.5}$$

where

$$F(\theta) = \frac{1}{n} \ln \pi(\text{data}, \theta)$$

and

$$F^*(\theta) = F(\theta) + \frac{1}{n} \ln h(\theta)$$

Tierney and Kadane (1986) applied Laplace's method to produce an approximation of (3.5) as follows:

$$\hat{h}_{\text{BT}}(\theta) = \left[\frac{\phi^*}{\phi}\right]^{1/2} \exp\{n[F^*(\bar{\theta}^*) - F(\bar{\theta})]\}, \tag{3.6}$$

in which $\bar{\theta}^*$ and $\bar{\theta}$ maximize $F^*(\theta)$ and $F(\theta)$, respectively, and ϕ^* and ϕ are minus the inverse of the second derivatives of $F^*(\theta)$ and $F(\theta)$ at $\bar{\theta}^*$ and $\bar{\theta}$, respectively.

We apply this approximation to obtain the Bayes estimate of the parameter θ. Setting $h(\theta)=\theta$, we have

$$\begin{aligned} F(\theta) = \frac{1}{n}\Big\{ &(2n+a-1)\log\theta - n\log(1+\theta) - \theta b \\ &+ \sum_{i=1}^{n} \log \int (1+x)\exp(-\theta x)\mu_{\tilde{x}_i}(x)\,dx \Big\}, \end{aligned} \tag{3.7}$$

and

$$\begin{aligned} F^*(\theta) = \frac{1}{n}\Big\{ &(2n+a)\log\theta - n\log(1+\theta) - \theta b \\ &+ \sum_{i=1}^{n} \log \int (1+x)\exp(-\theta x)\mu_{\tilde{x}_i}(x)\,dx \Big\}. \end{aligned} \tag{3.8}$$

Substituting from (3.7) and (3.8) in (3.6), the Bayes estimate of θ under squared error loss takes the form

$$\begin{aligned} \hat{\theta}_{\mathrm{BT}} = &\left(\frac{\phi^*}{\phi}\right)^{1/2} \left(\frac{\bar{\theta}^{*2n+a}}{\bar{\theta}^{2n+a-1}}\right) \left(\frac{1+\bar{\theta}}{1+\bar{\theta}^*}\right)^{n} \exp\big(b(\bar{\theta}-\bar{\theta}^*)\big) \\ &\times \prod_{i=1}^{n} \frac{\int (1+x)\exp(-\bar{\theta}^* x)\mu_{\tilde{x}_i}(x)\,dx}{\int (1+x)\exp(-\bar{\theta} x)\mu_{\tilde{x}_i}(x)\,dx}. \end{aligned} \tag{3.9}$$

3.2 MCMC and HPD credible interval

In this subsection, we first draw random samples from the posterior density function (3.3). Then, we compute the Bayes estimate of θ and also construct its HPD credible interval. Since the density function $\pi(\theta|\text{data})$ can not be computed explicitly, we use a Metropolis–Hastings algorithm to generate samples from posterior density of θ as follows:

Step (1) Start with an initial guess θ_0 and set $j=1$.

Step (2) Generate θ_j from $\pi(\theta|\text{data})$ by the following steps:

(a) Given initial value of θ, say $\rho=\theta_{j-1}$.

(b) Generate τ from the normal proposal distribution $g(\theta)\equiv I(\theta>0)N(\hat{\theta},1)$, where $I(\cdot)$ is the indicator function and $\hat{\theta}$ is the MLE of θ.

(c) Let $p(\rho,\tau)=\min\{1, \frac{\pi(\tau|\text{data})q(\rho)}{\pi(\rho|\text{data})q(\tau)}\}$.

(d) Generate a random number u from uniform $(0,1)$. Retain τ if $u\le p(\rho,\tau)$, else accept ρ.

Step (3) Repeat Step (2), M times and obtain θ_j for $j=1,\dots,M$.

The retained sample values, say $\theta_1, \dots, \theta_M$, are a random sample from the posterior density $\pi(\theta|\text{data})$. Now, by using Monte Carlo integration technique Rubinstein and Kroese (2006), the Bayes estimate of θ under squared error loss function can be obtained as

$$\hat{\theta}_{\text{BM}} = \frac{1}{M}\sum_{t=1}^{M}\theta_j.$$

For constructing HPD credible interval of θ, we can use the method proposed by Chen and Shao (1999) as follows.

Let $\theta_{(1)} < \cdots < \theta_{(M)}$ be the ordered values of θ_j for $j = 1, \dots, M$. Then, consider the following $100(1-\alpha)\%$ credible intervals of θ:

$$(\theta_{(1)}, \theta_{((1-\alpha)M)}), \dots, (\theta_{(\alpha M)}, \theta_{(M)}).$$

The HPD credible interval of θ can be derived by choosing the interval which has the shortest length.

4 Simulation study and comparisons

In this section, simulation studies are conducted to compare the performances of the different estimators and also different confidence/credible intervals. We mainly compare the performances of the MLE and Bayes estimates of the unknown parameter, in terms of their average values and mean squared errors. We also compare the average lengths of the asymptotic confidence intervals to the HPD credible intervals and their coverage percentages. All the computations are performed on R 2.11.1.

For simulation purposes, we have taken $\theta = 1$ and different choices of sample sizes, namely $n = 15, 20, 30, 50, 70$. For each n, we have generated random sample from the Lindley distribution. Then, using the method proposed by Pak, Parham and Saraj (2013), each realization of the generated samples was fuzzified by employing fuzzy information system $\{\tilde{x}_1, \dots, \tilde{x}_8\}$ corresponding to the membership functions

$$\mu_{\tilde{x}_1}(x) = \begin{cases} 1, & x \le 0.05, \\ \dfrac{0.25 - x}{0.2}, & 0.05 \le x \le 0.25, \\ 0, & \text{otherwise}, \end{cases}$$

$$\mu_{\tilde{x}_2}(x) = \begin{cases} \dfrac{x - 0.05}{0.2}, & 0.05 \le x \le 0.25, \\ \dfrac{0.5 - x}{0.25}, & 0.25 \le x \le 0.5, \\ 0, & \text{otherwise}, \end{cases}$$

$$\mu_{\tilde{x}_3}(x) = \begin{cases} \dfrac{x-0.25}{0.25}, & 0.25 \le x \le 0.5, \\ \dfrac{0.75-x}{0.25}, & 0.5 \le x \le 0.75, \\ 0, & \text{otherwise}, \end{cases}$$

$$\mu_{\tilde{x}_4}(x) = \begin{cases} \dfrac{x-0.5}{0.25}, & 0.5 \le x \le 0.75, \\ \dfrac{1-x}{0.25}, & 0.75 \le x \le 1, \\ 0, & \text{otherwise}, \end{cases}$$

$$\mu_{\tilde{x}_5}(x) = \begin{cases} \dfrac{x-0.75}{0.25}, & 0.75 \le x \le 1, \\ \dfrac{1.5-x}{0.5}, & 1 \le x \le 1.5, \\ 0, & \text{otherwise}, \end{cases}$$

$$\mu_{\tilde{x}_6}(x) = \begin{cases} \dfrac{x-1}{0.5}, & 1 \le x \le 1.5, \\ \dfrac{2-x}{0.5}, & 1.5 \le x \le 2, \\ 0, & \text{otherwise}, \end{cases}$$

$$\mu_{\tilde{x}_7}(x) = \begin{cases} \dfrac{x-1.5}{0.5}, & 1.5 \le x \le 2, \\ 3-x, & 2 \le x \le 3, \\ 0, & \text{otherwise}, \end{cases}$$

$$\mu_{\tilde{x}_8}(x) = \begin{cases} x-2, & 2 \le x \le 3, \\ 1, & x \ge 3, \\ 0, & \text{otherwise}. \end{cases}$$

The estimate of the parameter θ for the fuzzy sample were computed using the maximum likelihood method and Bayesian procedure. In computing the MLE, we have used the true value of θ as the initial guess value of θ. For computing the Bayes estimate, we have assumed that θ has gamma prior, including the non-informative gamma prior, that is, $a = b = 0$, and informative gamma prior, that is, $a = b = 2$. We replicate the process 10,000 times and report the average values (AV) and mean squared errors (MSE) of the estimates in Tables 1–3.

We have also computed approximate 95% confidence interval and also the HPD credible interval of the unknown parameter. Criteria appropriate to the evaluation of the two methods under scrutiny include: closeness of the coverage probability to its nominal value and expected interval width. For each simulated sample, we have computed confidence/credible intervals and checked whether the true value of the parameter lay within the intervals and recorded the length of the intervals.

Table 1 *Averages values and mean squared errors of the ML estimates of θ, coverage probabilities and expected width of 95% confidence interval for different sample sizes*

	$\hat{\theta}$		Confidence interval	
n	AV	MSE	Coverage	Length
15	1.0314	0.0446	0.9357	0.7593
20	1.0279	0.0407	0.9430	0.7285
30	1.0233	0.0355	0.9485	0.6221
50	1.0164	0.0264	0.9502	0.4770
70	1.0115	0.0221	0.9514	0.3854
100	1.0068	0.0144	0.9533	0.1935

Table 2 *Averages values and mean squared errors of the Bayes estimates of θ, coverage probabilities and expected width of the credible interval for $a = b = 0$*

	$\hat{\theta}_{BT}$		$\hat{\theta}_{BM}$		Credible interval	
n	AV	MSE	AV	MSE	Coverage	Length
15	1.0319	0.0448	1.0322	0.0450	0.9376	0.7327
20	1.0286	0.0413	1.0290	0.0415	0.9438	0.7046
30	1.0235	0.0356	1.0236	0.0358	0.9489	0.6153
50	1.0168	0.0266	1.0172	0.0269	0.9506	0.4718
70	1.0119	0.0223	1.0121	0.0224	0.9519	0.3819
100	1.0073	0.0149	1.0074	0.0149	0.9537	0.1926

Table 3 *Averages values and mean squared errors of the Bayes estimates of θ, coverage probabilities and expected width of the credible interval for $a = b = 2$*

	$\hat{\theta}_{BT}$		$\hat{\theta}_{BM}$		Credible interval	
n	AV	MSE	AV	MSE	Coverage	Length
15	1.0292	0.0411	1.0295	0.0412	0.9421	0.7109
20	1.0240	0.0385	1.0247	0.0387	0.9470	0.6856
30	1.0185	0.0317	1.0188	0.0319	0.9497	0.5732
50	1.0109	0.0235	1.0109	0.0235	0.9513	0.4478
70	1.0085	0.0196	1.0086	0.0196	0.9528	0.3625
100	1.0048	0.0131	1.0051	0.0132	0.9561	0.1877

The estimated coverage probability was computed as the number of intervals that covered the true value divided by 10,000 while the estimated expected width of the intervals was computed as the sum of the lengths for all intervals divided by 10,000. The coverage probabilities and the expected widths for different sample sizes are presented in Tables 1–3.

From Tables 1–3, some of the point are quite clear. The MSE of the estimators decrease significantly as the sample size n increases, as one would expected. The performances of the Bayes estimates with non-informative prior assumption and the maximum likelihood estimates are identical in terms of AVs and MSEs; however, it is observed that the Bayes estimates with informative prior are uniformly better. It is also seen that the Bayes estimates obtained by Tierney and Kadane's approximation and MCMC method behave in a similar manner. It should be mentioned that although the MCMC techniques are computationally expensive, but in turn we can use them to construct HPD credible interval.

Now considering the confidence and credible intervals, it is observed that the asymptotic results of the MLE work quite well. It can maintain the coverage percentages in most of the cases even when the sample size is small. The widths of the confidence/credible intervals narrow down with an increase in the sample size n. The performances of the credible intervals are satisfactory and their coverage percentages are close to the corresponding nominal level. Moreover, in most of the cases, the average lengths of the credible intervals are slightly shorter than the confidence intervals. It can be further observed that an informative prior distribution improves the performance of the Bayesian credible interval compared to the one using non-informative prior.

5 Conclusions

In this paper, we have considered the classical and Bayesian inference procedures for the Lindley distribution parameter when the available data are described in terms of fuzzy numbers. We have obtained the maximum likelihood estimate of the unknown parameter using EM algorithm. For computing the Bayes estimate, we have used Tierney and Kadane's approximation as well as MCMC technique, with different types of prior information. We have further constructed approximate confidence interval and HPD credible interval of the unknown parameter. The performances of the different methods have been compared by Monte Carlo simulations. Based on the results of the simulation study, we see clearly that, the Bayes estimates based on non-informative prior and maximum likelihood estimates give similar estimation results; however, the Bayes estimates with informative prior have smaller MSE, showing that additional prior information about the parameter θ provides an improvement in the estimates.

Acknowledgements

The authors would like to thank Editors and the referees for their constructive comments and suggestions which improved and enriched the presentation of the paper.

References

Chen, M. H. and Shao, Q. M. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. *Journal of Computational and Graphical Statistics* **8**, 69–92. MR1705909

Gil, M. A., López-Diaz, M. and Ralescu, D. A. (2006). Overview on the development of fuzzy random variables. *Fuzzy Sets and Systems* **157**, 2546–2557. MR2328381

Ghitany, M. E., Al-Mutairi, D. K. and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation* **79**, 279–287. MR2477530

Gupta, P. K. and Singh, B. (2012). Parameter estimation of Lindley distribution with hybrid censored data. *International Journal of System Assurance Engineering and Management*. doi:10.1007/s13198-012-0120-y.

Gupta, P. K., Singh, B. Krishna, H., and Kumar, K. (2011). Reliability estimation in Lindley distribution with progressively type II right censored sample. *Mathematics and Computers in Simulation* **82**, 281–294. MR2846043

Lindley, D. V. (1958). Fudicial distributions and Bayesian theorem. *Journal of the Royal Statistical Society B* **20**, 102–107. MR0095550

Mazucheli, J. and Achcarb, J. A. (2011). The Lindley distribution applied to competing risks lifetime data. *Computer Methods and Programs in Biomedicine* **104**, 188–192.

Pak, A., Parham, G. H. and Saraj, M. (2013). Inferences on the competing risk reliability problem for exponential distribution based on fuzzy data. *IEEE Transactions on Reliability* **63**, 2–13.

Rubinstein, R. Y. and Kroese, D. P. (2006). *Simulation and the Monte Carlo Method*, 2nd ed. New Jersey: Wiley. MR2365210

Shafiq, M. and Viertl, R. (2014). Maximum likelihood estimation for Weibull distribution in case of censored fuzzy life time data, 1–17. Available at http://www.statistik.tuwien.ac.at/forschung/SM/SM-2014-2complete.pdf.

Tierney, L. and Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association* **81**, 82–86. MR0830567

Viertl, R. (2006). Univariate statistical analysis with fuzzy data. *Computational Statistics & Data Analysis* **55**, 133–147. MR2297592

Wu, H. C. (2004). Fuzzy Bayesian estimation on lifetime data. *Computational Statistics* **19**, 613–633. MR2108493

Zadeh, L. A. (1968). Probability measures of fuzzy events. *Journal of Mathematical Analysis and Applications* **10**, 421–427. MR0230569

Zarei, R., Amini, M., Taheri, S. M. and Rezaei, A. H. (2012). Bayesian estimation based on vague lifetime data. *Soft Computing* **16**, 165–174.

Department of Computer Sciences
Faculty of Mathematical Sciences
Shahrekord University
P.O. Box 115, Shahrekord
Iran
E-mail: abbas.pak1982@gmail.com