Estimating the parameters of Lomax distribution from imprecise information

Abbas Pak*

Department of Computer Sciences, Faculty of Mathematical Sciences, Shahrekord University, P. O. Box 115, Shahrekord, Iran.
pak@sci.sku.ac.ir

Mohammad Reza Mahmoudi

Department of Statistics, Fasa University, Fasa, Iran
mahmoudi.m.r@fasau.ac.ir

Traditional statistical approaches for estimating the parameters of Lomax distribution have dealt with precise information. However, in real world situations, some information about an underlying system might be imprecise and are represented in the form of fuzzy information. In this paper, we consider the problem of estimating the parameters of Lomax distribution when the available observations are described by means of fuzzy information. We obtain the maximum likelihood estimate of the parameters by using the Newton-Raphson as well as the EM algorithm. We also provide an approximation namely, Tierney and Kadane's approximation, to compute the Bayes estimates of the unknown parameters. The estimation procedures are discussed in details and compared via Monte Carlo simulations in terms of their estimated biases and mean squared errors. Finally, analysis of one data set is provided for the purpose of illustration.

Keywords: Imprecise data; Fuzzy information; Lomax distribution; Maximum likelihood estimation; Bayesian estimation.

2000 Mathematics Subject Classification: 62N02, 62N86

1. Introduction

A random variable X is said to have Lomax distribution, if its probability density function (pdf) and cumulative distribution function are given, respectively, by

$$f(x) = \alpha\lambda \left(1 + \lambda x\right)^{-(\alpha+1)}, \quad x > 0 \tag{1.1}$$

and

$$F(x) = 1 - (1 + \lambda x)^{-\alpha}, \quad x > 0, \tag{1.2}$$

---

*Corresponding author

where $\alpha > 0$ and $\lambda > 0$ are the shape and scale parameters, respectively. From now on the Lomax distribution with parameters α and λ will be denoted by $LOM(\alpha, \lambda)$. This distribution is also known as pareto distribution of the second type since it is a special case of the generalized Pareto (GP) model with pdf

$$f(x;\xi,\mu,\sigma) = \frac{1}{\sigma}\left(1 + \xi\frac{x-\mu}{\sigma}\right)^{-\left(\frac{1}{\xi}+1\right)}, \tag{1.3}$$

where $\mu, \xi \in R$, $\sigma \in (0, +\infty)$ and for $\xi > 0$ the range of x is $x > 0$ and for $\xi < 0$ the range of x is $0 < x < \sigma/|\xi|$. The limit case $\xi = 0$ corresponds to the exponential distribution. Note that, by defining $\xi/\sigma \equiv \lambda$, $1/\xi \equiv \alpha$ and $\mu \equiv 0$, the pdf of Lomax distribution in relation (1.1) is obtained. There is a large amount of literature about the estimation of GP distribution parameters using different approaches; see for example, Singh and Guo [30], Kremer [18], Shi et al. [28], Lin and Wang [20], Oztekun [22], Hsu et al. [16], Castillo and Daoudi [9] and Bermudez and Kotz [6]. For more details about GP model and its properties, see Arnold [2].

Lomax distribution provides a very good alternative to the common lifetime distributions such as exponential, Weibull, or gamma when the experimenter presumes that the population distribution may be heavy-tailed (see Bryson [7]). Also, it has been shown that its utilities for modeling and analyzing lifetime data in medical and biological sciences, engineering, etc. So, it has been received greatest attention from theoretical and statisticians primarily due to its use in lifetime testing studies. The Lomax distribution has been used in the literature in a number of ways. For example, it has been extensively used for reliability modelling and life testing; see, for example, Balkema and de Haan [5]. Ahsanullah [1] studied the record values of Lomax distribution. Balakrishnan and Ahsanullah [4] introduced some recurrence relations between the moments of record values from Lomax distribution.

Several authors have addressed inferential issues for the Lomax distribution based on complete and censored samples. The order statistics from nonidentical right-truncated Lomax random variables have been studied by Childs et al. [10]. Howlader and Hossain [15] considered Bayesian estimation of the survival function of the Lomax distribution. Ghitany et al. [13] considered Marshall-Olkin approach and extended Lomax distribution. Elfattah et al. [12] derived the Bayesian and the non-Baysian estimators for the same sample size from the Lomax distribution based on the progressive type-I censoring. Raqab et al. [26] discussed different predictors of failure times such as best linear unbiased predictors, and maximum-likelihood and approximate maximum-likelihood predictors based on multistage progressive censoring from Pareto distribution, and recently Cramer and Schmiedt [11] discussed the optimal censoring scheme for estimating the parameters of the Lomax distribution using progressive type-II censoring. Asgharzadeh and Valiollahi [3] derived the Bayesian estimator of the scale parameter of the Lomax distribution based on progressive type-II censoring.

From the above, it is evident that all the earlier works on the estimation of the parameters of Lomax distribution have been done under the assumption of precise data. However, in real world situations, we deal with experiments whose observation does not provide exact but imprecise information. For example, the time of reaction of a person to a certain stimulus in a psychological experience can not be exactly determined, but the psychologist is able to determine it by means of the following imprecise information: "The time of reaction is approximately 25 to 35 seconds". To deal with the lack of precision of the data, it is necessary to incorporate fuzzy concept to statistical techniques. In recent years, many papers on generalization of classical statistical methods to

analysis of fuzzy data have been published. Among others, Buckley [8] considered fuzzy systems whose performance will depend on fuzzy probability distributions and whose measures of performance can be described by fuzzy numbers. Pak et al. [23] and [24] conducted a series of studies to develop the inferential procedures for the lifetime distributions on the basis of fuzzy numbers. Maximum likelihood estimation of exponential model using type-II fuzzy censored data is considered by Khoolenjani and Shahsanaie [17]. Makhdoom et al. [21] studied Bayesian estimation of the parameter of exponential distribution under type II censoring from fuzzy data.

To our knowledge there are no reports on estimating the parameters of Lomax distribution based on fuzzy data. The main aim of this paper is to obtain the inferential procedures for the Lomax distribution when the available observations are reported by means of fuzzy information. In Section 2, we use the Newton-Raphson and the EM algorithms to determine the maximum likelihood estimates of the unknown parameters. In Section 3, we obtain the Bayes estimates of the parameters α and λ by using the approximation form of Tierney and Kadane [32] under the assumption of gamma priors. A Monte Carlo simulation study is presented in Section 4, which provides a comparison of all estimation procedures developed in this paper and analysis of a data set is provided. Finally, conclusions and recommendations are provided in Section 5.

Let us first review the fundamental notation and basic definitions of fuzzy set theory used in the paper. Consider an experiment characterized by a probability space $\mathbf{X} = (X, \mathscr{B}_X, P_\theta)$, where $(X, \mathscr{B}_X)$ is a measurable space and P_θ belongs to a specified family of probability measures $\{P_\theta, \theta \in \Theta\}$ on $(X, \mathscr{B}_X)$. Assume that the observer cannot distinguish or transmit with exactness the outcome in the performance of $\mathbf{X}$, but that rather the available observation may be described in terms of fuzzy information which is defined as follows:

Definition 1. A fuzzy event $\tilde{x}$ on X, characterized by a Borel measurable membership function $\mu_{\tilde{x}}(x)$ from X to $[0,1]$, where $\mu_{\tilde{x}}(x)$ represents the "grade of membership" of x to $\tilde{x}$, is called *fuzzy information* associated with the experiment $\mathbf{X}$.

Among the various types of fuzzy information, the triangular and trapezoidal fuzzy numbers are most convenient and useful in describing fuzzy data. For triangular membership function, the triangular fuzzy number can be defined as $\tilde{x} = (\sigma_1, \sigma_2, \sigma_3)$ and its membership function is defined by the following expression:

$$\mu_{\tilde{x}}(x) = \begin{cases} \frac{x - \sigma_1}{\sigma_2 - \sigma_1} & \sigma_1 \leq x \leq \sigma_2, \\ \frac{\sigma_3 - x}{\sigma_3 - \sigma_2} & \sigma_2 \leq x \leq \sigma_3, \\ 0 & otherwise. \end{cases}$$

The set consisting of all observable events from the experiment $\mathbf{X}$ determines a fuzzy information system associated with it, which is defined as follows.

Definition 2. (see Tanaka et al. [31]). *A fuzzy information system* (f.i.s.) $\tilde{\mathbf{X}}$ associated with the experiment $\mathbf{X}$ is a fuzzy partition with fuzzy events on X, that is a finite set of fuzzy events on X satisfying the orthogonality condition

$$\sum_{\tilde{x} \in \tilde{\mathbf{X}}} \mu_{\tilde{x}}(x) = 1 \quad \text{for all} x \in X.$$

On the other hand, according to Zadeh [33] given the experiment $\mathbf{X} = (X, \mathscr{B}_X, P_\theta)$, $\theta \in \Theta$, and a f.i.s. $\tilde{\mathbf{X}}$ associated with it, each probability measure P_θ on $(X, \mathscr{B}_X)$ induces a probability measure on $\tilde{\mathbf{X}}$ defined as follows:

Definition 3. The probability distribution on $\tilde{\mathbf{X}}$ induced by P_θ is the mapping P from $\tilde{\mathbf{X}}$ to $[0,1]$ such that

$$\mathrm{P}(\tilde{x}) = \int_X \mu_{\tilde{x}}(x) dP_\theta(x), \quad \tilde{x} \in \tilde{\mathbf{X}}. \tag{1.4}$$

In particular, the conditional density of a continuous random variable $\mathbf{Y}$ with p.d.f. $g(\mathbf{y})$ given the fuzzy event $\tilde{A}$ can be defined as

$$g(\mathbf{y} \big| \tilde{A}) = \frac{\mu_{\tilde{A}}(\mathbf{y})g(\mathbf{y})}{\int \mu_{\tilde{A}}(\mathbf{u})g(\mathbf{u})d\mathbf{u}} \ . \tag{1.5}$$

For more details about the membership functions and probability measures of fuzzy sets, one can refer to Pak et al. [23] and the references therein.

2. Maximum likelihood estimation

Suppose that $X_1,...,X_n$ is a random sample of size n from Lomax distribution with pdf given by (1.1). Let $\mathbf{X} = (X_1,...,X_n)$ denotes the corresponding random vector. If a realization $\mathbf{x}$ of $\mathbf{X}$ was known exactly, we could obtain the complete-data likelihood function as

$$L(\mathbf{x}; \alpha, \lambda) = (\alpha\lambda)^n \prod_{i=1}^n (1 + \lambda x_i)^{-(\alpha+1)}. \tag{2.1}$$

Now consider the problem where $\mathbf{x}$ is not observed precisely, and only partial information about $\mathbf{x}$ is available in the form of fuzzy observation $\tilde{\mathbf{x}} = (\tilde{x}_1,...,\tilde{x}_n)$ with the Borel measurable membership function $\mu_{\tilde{\mathbf{x}}}(\mathbf{x})$. In practice, the grade of membership $\mu_{\tilde{\mathbf{x}}}(\mathbf{x})$ is often regarded as a kind of "probability with which the observer gets the information $\tilde{\mathbf{x}}$ when he really has obtained the exact outcome $\mathbf{x}$ ".

Example 1. A geologist is interested in analyzing the length of the largest axis of boulders in the upper reaches of a particular river. Assume that the geologist does not have a mechanism of measurement sufficiently precise to determine exactly the length of the largest axis of boulders. More precisely, suppose that the lack of roundness of these boulders only allows him to approximate the length of the largest axes by means of the following fuzzy observations: $\tilde{x}_1$ ="approximately lower than 30 inches", $\tilde{x}_2$ ="approximately 35 to 50 inches", $\tilde{x}_3$ ="approximately 55 inches", $\tilde{x}_4$ ="approximately 60 to 65 inches", $\tilde{x}_5$ ="approximately higher than 70 inches", which are characterized by the membership functions in Fig. 1 (Clearly, a f.i.s. $\tilde{\mathbf{X}} = \{\tilde{x}_1,...,\tilde{x}_6\}$ can be immediately constructed by defining $\mu_{\tilde{x}_6} = 1 - \sum_{i=1}^5 \mu_{\tilde{x}_i}$).

Once $\tilde{\mathbf{x}}$ is given, we can obtain the observed data log-likelihood function by using the expression (1.4) as follows:

$$L_O(\tilde{\mathbf{x}}; \alpha, \lambda) = n\log\alpha + n\log\lambda + \sum_{i=1}^n \log \int (1 + \lambda x)^{-(\alpha+1)} \mu_{\tilde{x}_i}(x) dx. \tag{2.2}$$

The maximum likelihood estimate of the parameters α and λ can be obtained by maximizing the log-likelihood $L_O(\tilde{\mathbf{x}}; \alpha, \lambda)$. Equating the partial derivatives of the log-likelihood (2.2) with respect
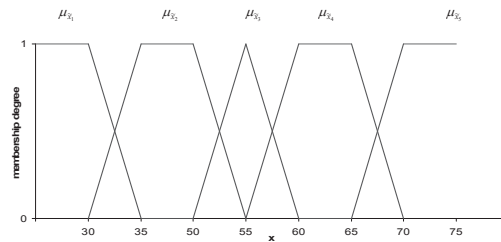
Fig. 1. Membership functions of the fuzzy observations $\tilde{x}_1$, $\tilde{x}_2$, $\tilde{x}_3$, $\tilde{x}_4$ and $\tilde{x}_5$.

to α and λ to zero, the resulting two equations are:

$$\frac{\partial L_O(\tilde{\mathbf{x}};\alpha,\lambda)}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} \frac{\int (1+\lambda x)^{-(\alpha+1)} \log(1+\lambda x) \mu_{\tilde{x}_i}(x)dx}{\int (1+\lambda x)^{-(\alpha+1)} \mu_{\tilde{x}_i}(x)dx} = 0, \qquad (2.3)$$

$$\frac{\partial L_O(\tilde{\mathbf{x}};\alpha,\lambda)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} \frac{\int (\alpha+1)x(1+\lambda x)^{-(\alpha+2)} \mu_{\tilde{x}_i}(x)dx}{\int (1+\lambda x)^{-(\alpha+1)} \mu_{\tilde{x}_i}(x)dx} = 0. \qquad (2.4)$$

Since there are no closed form of the solutions to the likelihood equations (2.3) and (2.4), an iterative numerical search can be used to obtain the MLEs. In the following, we describe the Newton-Raphson method and the EM algorithm to determine the MLEs of the parameters α and λ.

2.1. *Newton-Raphson algorithm*

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

$$\theta^{(h+1)} = \theta^{(h)} - \left[\frac{\partial L_O(\tilde{\mathbf{x}};\theta)}{\partial \theta} \Big|_{\theta=\theta^{(h)}} \right]^T \left[\frac{\partial^2 L_O(\tilde{\mathbf{x}};\theta)}{\partial \theta \partial \theta^{\mathbf{T}}} \Big|_{\theta=\theta^{(h)}} \right]^{-1}, \qquad (2.5)$$

where

$$\frac{\partial L_O(\tilde{\mathbf{x}};\theta)}{\partial \theta} = \begin{pmatrix} \frac{\partial L_O(\tilde{\mathbf{x}};\alpha,\lambda)}{\partial \alpha} \\ \frac{\partial L_O(\tilde{\mathbf{x}};\alpha,\lambda)}{\partial \lambda} \end{pmatrix}$$

and

$$\frac{\partial^2 L_O(\tilde{\mathbf{x}};\theta)}{\partial \theta \partial \theta^{\mathbf{T}}} = \begin{pmatrix} \frac{\partial^2 L_O(\tilde{\mathbf{x}};\alpha,\lambda)}{\partial \alpha^2} & \frac{\partial^2 L_O(\tilde{\mathbf{x}};\alpha,\lambda)}{\partial \alpha \partial \lambda} \\ \frac{\partial^2 L_O(\tilde{\mathbf{x}};\alpha,\lambda)}{\partial \alpha \partial \lambda} & \frac{\partial^2 L_O(\tilde{\mathbf{x}};\alpha,\lambda)}{\partial \lambda^2} \end{pmatrix}.$$

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

$$\frac{\partial^2 L_O(\tilde{\mathbf{x}};\alpha,\lambda)}{\partial \alpha^2} = -\frac{n}{\alpha^2} + \sum_{i=1}^{n} \frac{\int (1+\lambda x)^{-(\alpha+1)} [\log(1+\lambda x)]^2 \mu_{\tilde{x}_i}(x)dx}{\int (1+\lambda x)^{-(\alpha+1)} \mu_{\tilde{x}_i}(x)dx}$$
$$- \sum_{i=1}^{n} \left[\frac{\int (1+\lambda x)^{-(\alpha+1)} [\log(1+\lambda x)]^2 \mu_{\tilde{x}_i}(x)dx}{\int (1+\lambda x)^{-(\alpha+1)} \mu_{\tilde{x}_i}(x)dx} \right]^2,$$

$$\frac{\partial^2 L_O(\tilde{\mathbf{x}};\alpha,\lambda)}{\partial \lambda^2} = -\frac{n}{\lambda^2} + \sum_{i=1}^{n} \frac{\int (\alpha+1)(\alpha+2)x^2(1+\lambda x)^{-(\alpha+3)}\mu_{\tilde{x}_i}(x)dx}{\int (1+\lambda x)^{-(\alpha+1)}\mu_{\tilde{x}_i}(x)dx}$$
$$- \sum_{i=1}^{n} \left[\frac{\int (\alpha+1)x(1+\lambda x)^{-(\alpha+2)}\mu_{\tilde{x}_i}(x)dx}{\int (1+\lambda x)^{-(\alpha+1)}\mu_{\tilde{x}_i}(x)dx}\right]^2,$$

$$\frac{\partial^2 L_O(\tilde{\mathbf{x}};\alpha,\lambda)}{\partial \alpha \partial \lambda} = \sum_{i=1}^{n} \frac{\int x(1+\lambda x)^{-(\alpha+2)}[(\alpha+1)\log(1+\lambda x)-1]\mu_{\tilde{x}_i}(x)dx}{\int (1+\lambda x)^{-(\alpha+1)}\mu_{\tilde{x}_i}(x)dx}$$
$$- \sum_{i=1}^{n} \left(\frac{\int (1+\lambda x)^{-(\alpha+1)}\log(1+\lambda x)\mu_{\tilde{x}_i}(x)dx}{\int (1+\lambda x)^{-(\alpha+1)}\mu_{\tilde{x}_i}(x)dx}\right.$$
$$\left.\times \left[\frac{\int (\alpha+1)x(1+\lambda x)^{-(\alpha+2)}\mu_{\tilde{x}_i}(x)dx}{\int (1+\lambda x)^{-(\alpha+1)}\mu_{\tilde{x}_i}(x)dx}\right]\right).$$

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

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

2.2. *EM algorithm*

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

From (2.1), the log-likelihood function for the complete data vector $\mathbf{x}$ becomes:

$$\log L(\mathbf{x};\alpha,\lambda) = n\log\alpha + n\log\lambda - (\alpha+1)\sum_{i=1}^{n}\log(1+\lambda x_i). \tag{2.6}$$

Taking the derivative with respect to α and λ, respectively, on (2.6), the following likelihood equations are obtained:

$$\frac{n}{\alpha} = \sum_{i=1}^{n}\log(1+\lambda x_i), \tag{2.7}$$

$$\frac{n}{\lambda} = (\alpha+1)\sum_{i=1}^{n}\frac{x_i}{1+\lambda x_i}. \tag{2.8}$$

Therefore the EM algorithm is given by the following iterative process:
1. Given starting values of α and λ, say $\alpha^{(0)}$ and $\lambda^{(0)}$ and set $h = 0$.
2. In the $(h+1)$th iteration,

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

(1.5):

$$E_{\alpha^{(h)},\lambda^{(h)}}(\log(1+\lambda^{(h)}X)\,|\tilde{\mathbf{x}}) = \frac{\int(1+\lambda^{(h)}x)^{-(\alpha^{(h)}+1)}\log(1+\lambda^{(h)}x)\mu_{\tilde{x}_i}(x)dx}{\int(1+\lambda^{(h)}x)^{-(\alpha^{(h)}+1)}\mu_{\tilde{x}_i}(x)dx},$$

$$E_{\alpha^{(h)},\lambda^{(h)}}\left(\frac{X}{1+\lambda^{(h)}X}\,|\tilde{\mathbf{x}}\right) = \frac{\int x(1+\lambda^{(h)}x)^{-(\alpha^{(h)}+2)}\mu_{\tilde{x}_i}(x)dx}{\int(1+\lambda^{(h)}x)^{-(\alpha^{(h)}+1)}\mu_{\tilde{x}_i}(x)dx},$$

and the likelihood equations (2.7) and (2.8) are replaced by

$$\frac{n}{\alpha} = \sum_{i=1}^{n} E_{\alpha^{(h)},\lambda^{(h)}}(\log(1+\lambda^{(h)}X)\,|\tilde{\mathbf{x}})\,, \tag{2.9}$$

$$\frac{n}{\lambda} = (\alpha+1)\sum_{i=1}^{n} E_{\alpha^{(h)},\lambda^{(h)}}\left(\frac{X}{1+\lambda^{(h)}X}\,|\tilde{\mathbf{x}}\right). \tag{2.10}$$

(1) The M-step requires to solve the Eqs. (2.9) and (2.10) and obtain the next values, $\alpha^{(h+1)}$ and $\lambda^{(h+1)}$, of α and λ, respectively, as follows:

$$\alpha^{(h+1)} = \frac{n}{\sum_{i=1}^{n} E_{\alpha^{(h)},\lambda^{(h)}}(\log(1+\lambda^{(h)}X)\,|\tilde{\mathbf{x}})},$$

$$\lambda^{(h+1)} = \frac{n}{(\alpha^{(h+1)}+1)\sum_{i=1}^{n} E_{\alpha^{(h)},\lambda^{(h)}}\left(\frac{X}{1+\lambda^{(h)}X}\,|\tilde{\mathbf{x}}\right)}.$$

3. Checking convergence, if the convergence occurs then the current $\alpha^{(h+1)}$ and $\lambda^{(h+1)}$ are the maximum likelihood estimates of α and λ via EM algorithm; otherwise, set $h = h+1$ and go to Step 2. The MLE of (α,λ) via EM algorithm is thereafter refereed as $(\hat{\alpha}_{EM},\hat{\lambda}_{EM})$ in this paper.

3. Bayesian estimation

In recent decades, the Bayes viewpoint, as a powerful and valid alternative to traditional statistical perspectives, has received frequent attention for statistical inference. In this section, we consider the Bayesian estimation of the unknown parameters by using a squared error loss function. In order to evaluate the behavior of the parameters, it is assumed that α and λ have independent gamma priors with the pdfs

$$\pi_1(\alpha) = \frac{c^d}{\Gamma(c)}\alpha^{c-1}e^{-\lambda d}, \quad \alpha > 0, \tag{3.1}$$

and

$$\pi_2(\lambda) = \frac{b^a}{\Gamma(a)}\lambda^{a-1}e^{-\lambda b}, \quad \lambda > 0. \tag{3.2}$$

Based on the above priors, the joint posterior density function of α and λ given the data can be written as follows:

$$\pi(\alpha,\lambda\,|\tilde{\mathbf{x}}) = \frac{\pi_1(\alpha)\pi_2(\lambda)\ell(\tilde{\mathbf{x}};\alpha,\lambda)}{\int_0^\infty\int_0^\infty \pi_1(\alpha)\pi_2(\lambda)\ell(\tilde{\mathbf{x}};\alpha,\lambda)d\alpha d\lambda}, \tag{3.3}$$

where

$$\ell(\tilde{\mathbf{x}};\alpha,\lambda) = \alpha^n \lambda^n \prod_{i=1}^{n} \int (1+\lambda x_i)^{-(\alpha+1)} \mu_{\tilde{x}_i}(x)dx,$$

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

$$E(g(\alpha,\lambda)|\tilde{\mathbf{x}}) = \frac{\int_0^\infty \int_0^\infty g(\alpha,\lambda)\pi_1(\alpha)\pi_2(\lambda)\ell(\tilde{\mathbf{x}};\alpha,\lambda)d\alpha d\lambda}{\int_0^\infty \int_0^\infty \pi_1(\alpha)\pi_2(\lambda)\ell(\tilde{\mathbf{x}};\alpha,\lambda)d\alpha d\lambda} = \frac{\int_0^\infty \int_0^\infty g(\alpha,\lambda)e^{Q(\alpha,\lambda)}d\alpha d\lambda}{\int_0^\infty \int_0^\infty e^{Q(\alpha,\lambda)}d\alpha d\lambda}, \quad (3.4)$$

where $Q(\alpha,\lambda) = \ln[\pi_1(\alpha)\pi_2(\lambda)] + \ln \ell(\tilde{\mathbf{x}};\alpha,\lambda) \equiv \rho(\alpha,\lambda) + L(\alpha,\lambda)$. Note that Eq. (3.4) cannot be obtained analytically; therefore, in the following, we adopt Tierney and Kadane's approximation for computing the Bayes estimates.

Setting $H(\alpha,\lambda) = Q(\alpha,\lambda)/n$ and $H^*(\alpha,\lambda) = [\ln g(\alpha,\lambda) + Q(\alpha,\lambda)]/n$, the expression in (3.4) can be reexpressed as

$$E(g(\alpha,\lambda)|\tilde{\mathbf{x}}) = \frac{\int_0^\infty \int_0^\infty e^{nH^*(\alpha,\lambda)}d\alpha d\lambda}{\int_0^\infty \int_0^\infty e^{nH(\alpha,\lambda)}d\alpha d\lambda}. \quad (3.5)$$

Following Tierney and Kadane [32], Eq. (3.5) can be approximated as

$$\hat{g}_{BT}(\alpha,\lambda) = \left(\frac{\det \Sigma^*}{\det \Sigma}\right)^{\frac{1}{2}} \exp\left\{n[H^*(\bar{\alpha}^*,\bar{\lambda}^*) - H(\bar{\alpha},\bar{\lambda})]\right\}, \quad (3.6)$$

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

$$H(\alpha,\lambda) = \frac{1}{n}\{k + (n+c-1)\log\alpha + (n+a-1)\log\lambda - \lambda b - \alpha d + \sum_{i=1}^{n}\log\int(1+\lambda x)^{-(\alpha+1)}\mu_{\tilde{x}_i}(x)dx\},$$

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

$$\frac{\partial H(\alpha,\lambda)}{\partial \alpha} = \frac{1}{n}\left\{\frac{n+c-1}{\alpha} - d - \sum_{i=1}^{n}\frac{\int(1+\lambda x)^{-(\alpha+1)}\log(1+\lambda x)\mu_{\tilde{x}_i}(x)dx}{\int(1+\lambda x)^{-(\alpha+1)}\mu_{\tilde{x}_i}(x)dx}\right\} = 0,$$

$$\frac{\partial H(\alpha,\lambda)}{\partial \lambda} = \frac{1}{n}\left\{\frac{n+a-1}{\lambda} - b - \sum_{i=1}^{n}\frac{\int(\alpha+1)x(1+\lambda x)^{-(\alpha+2)}\mu_{\tilde{x}_i}(x)dx}{\int(1+\lambda x)^{-(\alpha+1)}\mu_{\tilde{x}_i}(x)dx}\right\} = 0,$$

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

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

where

$$H_{11} = \frac{1}{n}\left\{-\frac{n+c-1}{\bar{\alpha}^2} + \sum_{i=1}^{n}\left(\frac{\int(1+\bar{\lambda}x)^{-(\bar{\alpha}+1)}[\log(1+\bar{\lambda}x)]^2\mu_{\tilde{x}_i}(x)dx}{\int(1+\bar{\lambda}x)^{-(\bar{\alpha}+1)}\mu_{\tilde{x}_i}(x)dx}\right) - \sum_{i=1}^{n}\left[\frac{\int(1+\bar{\lambda}x)^{-(\bar{\alpha}+1)}\log(1+\bar{\lambda}x)\mu_{\tilde{x}_i}(x)dx}{\int(1+\bar{\lambda}x)^{-(\bar{\alpha}+1)}\mu_{\tilde{x}_i}(x)dx}\right]^2\right\},$$

$$H_{22} = \frac{1}{n}\Big\{-\frac{n+a-1}{\bar{\lambda}^2} + \sum_{i=1}^{n}\Big(\frac{\int(\bar{\alpha}+1)(\bar{\alpha}+2)x^2(1+\bar{\lambda}x)^{-(\bar{\alpha}+3)}\mu_{\tilde{x}_i}(x)dx}{\int(1+\bar{\lambda}x)^{-(\bar{\alpha}+1)}\mu_{\tilde{x}_i}(x)dx}\Big)$$
$$-\sum_{i=1}^{n}\Big[\frac{\int(\bar{\alpha}+1)x(1+\bar{\lambda}x)^{-(\bar{\alpha}+2)}\mu_{\tilde{x}_i}(x)dx}{\int(1+\bar{\lambda}x)^{-(\bar{\alpha}+1)}\mu_{\tilde{x}_i}(x)dx}\Big]^2\Big\},$$

$$H_{12} = \frac{1}{n}\Big\{\sum_{i=1}^{n}\frac{\int x(1+\bar{\lambda}x)^{-(\bar{\alpha}+2)}[(\bar{\alpha}+1)\log(1+\bar{\lambda}x)-1]\mu_{\tilde{x}_i}(x)dx}{\int(1+\bar{\lambda}x)^{-(\bar{\alpha}+1)}\mu_{\tilde{x}_i}(x)dx}$$
$$-\sum_{i=1}^{n}\Big(\frac{\int(\bar{\alpha}+1)x(1+\bar{\lambda}x)^{-(\bar{\alpha}+2)}\mu_{\tilde{x}_i}(x)dx}{\int(1+\bar{\lambda}x)^{-(\bar{\alpha}+1)}\mu_{\tilde{x}_i}(x)dx}$$
$$\times\Big[\frac{\int(1+\bar{\lambda}x)^{-(\bar{\alpha}+1)}\log(1+\bar{\lambda}x)\mu_{\tilde{x}_i}(x)dx}{\int(1+\bar{\lambda}x)^{-(\bar{\alpha}+1)}\mu_{\tilde{x}_i}(x)dx}\Big]\Big)\Big\}$$

Now, following the same argument with $g(\alpha, \lambda) = \alpha$ and λ, respectively, in $H^*(\alpha, \lambda)$, $\hat{\alpha}_{BT}$ and $\hat{\lambda}_{BT}$ in Eq. (3.6) can then be obtained straightforwardly.
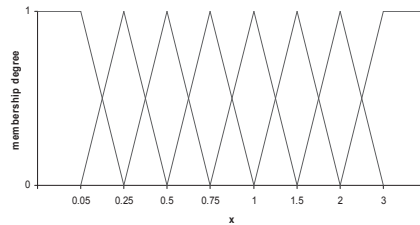

Figure 2: Fuzzy information system used to encode the simulated data

4. Numerical Study

4.1. *Monte Carlo Simulations*

In this section, we present some experimental results, mainly to observe how the different methods behave for different sample sizes. We obtain the estimates of the unknown parameters α and λ by using the methods provided in the preceding sections. The performance of the competitive estimates has been compared on the basis of their estimated biases and mean squared errors. The computations are performed using R 2.14.0 [27], which is a non-commercial, open source software package for statistical computing.

First, for two sets of parameter values namely; ($\alpha = 1, \lambda = 1$) and ($\alpha = 2, \lambda = 1.5$)and various choices of sample size *n*, we have generated i.i.d random samples from the Lomax distribution. Then, using the method proposed by Pak et al. [24], each realization of the generated samples was fuzzified by employing the fuzzy information system shown in Fig.2, and the estimates of parameters for the fuzzy sample were computed using the maximum likelihood method (via Newton-Raphson and EM algorithms) and Bayesian procedure. To make the comparison between ML and Bayes estimates meaningful, two types of priors are suggested, namely non-informative prior and informative gamma prior (Lin et al. [19]). The first one is used when we do not have any prior information about the parameters while the informative priors are used to investigate any improvements in the performances of Bayes estimators. Here, we have considered two different choices of hyperparameters as

Prior I: (non-informative) $a = b = c = d = 0$

Prior II: (informative) $a = b = c = d = 2$.

It must be noted that the non-informative prior I is non-proper also. Press [25] suggested to use very small non-negative values of the hyperparameters in this case, and it will make the priors proper. We have tried $a = b = c = d = 0.0001$. The results are not significantly different than the corresponding results obtained using non-proper priors, and are not reported due to space. From now on, the Bayes estimates of parameters obtained by using the above priors will be denoted by BET1 and BET2, respectively. The estimated biases and mean squared errors of the estimates over 1000 replications are presented in Tables 1-4.

From the experiments, we found that using the NR or EM algorithm for the computation of maximum likelihood estimates of α and λ give similar estimation results. Moreover, with the EM algorithm, there is no need to evaluate the first and second derivatives of the log-likelihood function, which helps save the central processing unit (CPU) time of each iteration; although less iteration is required by the NR method, the CPU time required per iteration is substantially shorter for the EM algorithm. The performances of the estimates are satisfactory in terms of biases and MSEs, even for small sample sizes. It can be further observed that when we do not have any prior information about the parameters, using the Bayes estimates we may not gain much as expected. For all the methods, it is observed that as the sample size increases, the biases and MSEs of the estimates decrease as expected. For small and moderate sample sizes, the Bayesian approach based on informative priors, gives the most precise parameter estimates as shown by ABs and MSEs in Tables 1-4. For large sample sizes ($n = 100, 200$), the performances of the MLEs and Bayes estimates are almost identical.

Table 1. Estimated bias (EB) and mean squared error (MSE) of the ML and Bayes estimates of α for different sample sizes ($\alpha = 1, \lambda = 1$).

n	NR		EM		BET1		BET2	
	EB	MSE	EB	MSE	EB	MSE	EB	MSE
15	0.0567	0.0396	0.0569	0.0397	0.0553	0.0391	0.0535	0.0366
20	0.0509	0.0378	0.0511	0.0380	0.0504	0.0375	0.0477	0.0332
25	0.0482	0.0341	0.0485	0.0343	0.0479	0.0338	0.0439	0.0319
30	0.0417	0.0276	0.0419	0.0277	0.0403	0.0269	0.0381	0.0253
40	0.0362	0.0219	0.0364	0.0220	0.0361	0.0217	0.0345	0.0196
50	0.0319	0.0154	0.0319	0.0154	0.0315	0.0152	0.0307	0.0122
70	0.0254	0.0117	0.0256	0.0118	0.0237	0.0114	0.0239	0.0104
100	0.0213	0.0093	0.0214	0.0093	0.0213	0.0092	0.0208	0.0089
200	0.0173	0.0081	0.0174	0.0082	0.0171	0.0080	0.0168	0.0076

4.2. *Illustrative example*

In this example, we consider a date set that were obtained from a meteorological study by Simpson [29] and further analyzed by Helu et al. [14]. The study was based on the radar-evaluated rainfall from 52 south Florida cumulus clouds, 26 seeded clouds, and 26 control clouds. Since the rainfall data evaluated by radar systems inevitably have some degree of imprecision, it is suggested to report

Table 2. Estimated bias (EB) and mean squared error (MSE) of the ML and Bayes estimates of λ for different sample sizes ($\alpha = 1, \lambda = 1$).

n	NR		EM		BET1		BET2	
	EB	MSE	EB	MSE	EB	MSE	EB	MSE
15	0.0873	0.0542	0.0875	0.0543	0.0856	0.0537	0.0822	0.0517
20	0.0846	0.0507	0.0849	0.0508	0.0840	0.0498	0.0805	0.0492
25	0.0721	0.0440	0.0722	0.0441	0.0713	0.0432	0.0689	0.0420
30	0.0697	0.0416	0.0698	0.0417	0.0685	0.0409	0.0644	0.0397
40	0.0611	0.0357	0.0611	0.0357	0.0605	0.0348	0.0572	0.0308
50	0.0558	0.0313	0.0560	0.0315	0.0557	0.0310	0.0521	0.0245
70	0.0432	0.0271	0.0432	0.0271	0.0429	0.0262	0.0419	0.0228
100	0.0395	0.0249	0.0396	0.0251	0.0395	0.0248	0.0387	0.0244
200	0.0246	0.0177	0.0247	0.0177	0.0242	0.0176	0.0241	0.0169

Table 3. Estimated bias (EB) and mean squared error (MSE) of the ML and Bayes estimates of α for different sample sizes ($\alpha = 2, \lambda = 1.5$).

n	NR		EM		BET1		BET2	
	EB	MSE	EB	MSE	EB	MSE	EB	MSE
15	0.0941	0.0687	0.0942	0.0687	0.0926	0.0675	0.0871	0.0619
20	0.0912	0.0634	0.0914	0.0635	0.0908	0.0631	0.0835	0.0578
25	0.0858	0.0573	0.0860	0.0574	0.0839	0.0559	0.0772	0.0513
30	0.0730	0.0506	0.0733	0.0506	0.0724	0.0502	0.0694	0.0462
40	0.0691	0.0428	0.0692	0.0428	0.0690	0.0427	0.0645	0.0408
50	0.0617	0.0382	0.0617	0.0384	0.0611	0.0380	0.0603	0.0374
70	0.0582	0.0341	0.0582	0.0342	0.0579	0.0338	0.0578	0.0336
100	0.0424	0.0296	0.0426	0.0296	0.0424	0.0294	0.0421	0.0290
200	0.0316	0.0182	0.0318	0.0183	0.0314	0.0178	0.0310	0.0177

the partial information on the rainfalls by means of lower and upper bounds, as well as a point estimate. Assume that the imprecision of the rainfalls is formulated by triangular fuzzy numbers $\tilde{x}_i = (\mu_i, x_i, \eta_i), i = 1, ..., 52$, where $\mu_i = 0.05x_i$ and $\eta_i = 0.03x_i$ with the membership functions

$$\mu_{\tilde{x}}(x) = \begin{cases} \frac{x - (x_i - \mu_i)}{\mu_i} & x_i - \mu_i \leq x \leq x_i, \\ \frac{x_i + \eta_i - x}{\eta_i} & x_i \leq x \leq x_i + \eta_i. \end{cases}$$

Here, for the fuzzy sample $\tilde{x}_1, ..., \tilde{x}_{52}$, we employ NR and EM algorithms to compute the ML estimates of the parameters. The stopping criterion is based on the difference between the two consecutive iterates, with a tolerance value $\varepsilon = 10^{-6}$. The final MLEs are ($\hat{\alpha}_{NR} = 0.7843, \hat{\lambda}_{NR} = 0.0054$) and ($\hat{\alpha}_{EM} = 0.7844, \hat{\lambda}_{EM} = 0.0052$). For computing the Bayes estimates, two different choices of hyperparameters are considered as $(a, b, c, d) = (0.0001, 0.0001, 0.0001, 0.0001), (2, 2, 2, 2)$. Then, using the Tierney and Kadane's approximation, the Bayes estimate of the parameters become ($\hat{\alpha}_{BT} = 0.7824, \hat{\lambda}_{BT} = 0.0057$) and ($\hat{\alpha}_{BT} = 0.7761, \hat{\lambda}_{BT} = 0.0063$) for the above non-informative and informative priors, respectively. It is observed that the ML estimates of the parameters from

Table 4. Estimated bias (EB) and mean squared error (MSE) of the ML and Bayes estimates of λ for different sample sizes ($\alpha = 2, \lambda = 1.5$).

n	NR		EM		BET1		BET2	
	EB	MSE	EB	MSE	EB	MSE	EB	MSE
15	0.1456	0.1075	0.1457	0.1076	0.1408	0.1055	0.1289	0.0966
20	0.1391	0.0928	0.1393	0.0930	0.1375	0.0918	0.1321	0.0874
25	0.1328	0.0861	0.1328	0.0861	0.1327	0.0854	0.1296	0.0815
30	0.1279	0.0827	0.1282	0.0828	0.1276	0.0823	0.1208	0.0792
40	0.1044	0.0695	0.1045	0.0697	0.1032	0.0692	0.0983	0.0648
50	0.0953	0.0613	0.0954	0.0613	0.0941	0.0610	0.0877	0.0580
70	0.0765	0.0489	0.0767	0.0489	0.0762	0.0488	0.0714	0.0456
100	0.0492	0.0325	0.0494	0.0327	0.0490	0.0324	0.0478	0.0313
200	0.0378	0.0219	0.0379	0.0219	0.0377	0.0215	0.0364	0.0209

Newton-Raphson and EM algorithms are about the same. Moreover, the Bayes estimates evaluated based on non-informative priors, are very close to the corresponding MLEs.

5. Conclusions

In the literature, there are well-developed estimation techniques for the parameters of Lomax distribution based on complete and censored data. But, traditionally it is assumed that the available data are obtained as exact numbers. However, in real world situations, the results of the experimental performance can not always be recorded or measured precisely, but each observable event may only be identified with a fuzzy subset of the sample space. Therefore, we need suitable statistical methodology to handle these data as well. In this paper, we have discussed different estimation procedures for the Lomax distribution when the obtained data are reported in terms of fuzzy information. They include the maximum likelihood method (via Newton-Raphson and EM algorithms) and Bayesian procedure. We have then carried out a simulation study to assess the performance of these procedures. Based on the results of the simulation study, we see clearly that, the maximum likelihood estimates based on Newton-Raphson and EM algorithms behave in a very similar manner, but the EM algorithm is computationally slower. Among the two estimation procedures developed in the paper, the Bayesian procedure with informative priors gives smaller biases and MSEs compared to the maximum likelihood method. Also, the MSE values of the estimates are close to each other as the sample size increases.

References

[1] M., Ahsanullah, Record values of Lomax distribution. *Statistica Nederlandica*, **41**(1.1), (1991), 21-29.

[2] B., Arnold, *Pareto Distributions, Second Edition*. Chapman & Hall/ CRC Press, London, (2015).

[3] A., Asgharzadeh, and R., Valiollahi, Estimation of the scale parameter of the Lomax distribution under progressive censoring. *International Journal of Statistics and Economics*, **6**, (2011), 37-48.

[4] N., Balakrishnan, and M., Ahsanullah, Relations for single and product moments of record values from Lomax distribution. *Sankhya B*, **56**(1.2), (1994), 140-146.

[5] A., Balkema, and L., De Haan, Residual life time at great age. *Annals of Probability*, **2**(1.5), (1974), 792-804.

[6] P.D., Bermudez, and S., Kotz, Parameter estimation of the generalized Pareto distribution-Part I. *Journal of Statistical Planning and Inference*. **140**, (2010), 1353-1373.

[7] M.C., Bryson, Heavy-tailed distributions: properties and tests. *Technometrics*, **16**(1.1), (1974), 61-68.

[8] J.J., Buckley, Fuzzy system. *Soft Computing*, **9**, (2005), 757-760.

[9] J.D., Castillo, and J., Daoudi, Estimation of the generalized Pareto distribution. *Statistics and Probability Letters*, **79**, (2009), 684-688.

[10] A., Childs, N., Balakrishnan, and M., Moshref, Order statistics from non-identical right-truncated Lomax random variables with applications. *Statistical Papers*, **42**(1.2), (2001), 187-206.

[11] E., Cramer, and A.B., Schmiedt, Progressively type-II censored competing risks data from Lomax distributions. *Computational Statistics and Data Analysis*, **55**(1.3), (2011), 1285-1303.

[12] A., Elfattah, F., Alaboud, and A., Alharby, On sample size estimation for Lomax distribution. *Australian Journal of Basic and Applied Sciences*, **4**, (2007), 373-378.

[13] M.E., Ghitany, F.A., Al-Awadhi, and L.A., Alkhalfan, Marshall-Olkin extended lomax distribution and its application to censored data. *Communications in Statistics*, **36**(2.5), (2007), 1855-1866.

[14] A., Helu, H., Samawi, and M.Z., Raqab, Estimation on Lomax progressive censoring using the EM algorithm. *Journal of Statistical Computation and Simulation*, (2013), **85**(1.5), 1035-1052.

[15] H.A., Howlader, and A.M., Hossain, Bayesian survival estimation of Pareto distribution of the second kind based on failure censored data. *Computational Statistics and Data Analysis*, **8**(1.3), (2002), 301-314.

[16] B.M., Hsu, B.S., Chen, and M.H., Shu, Evaluating lifetime performance for the Pareto model with censored information. *Second International Conference on Innovative Computing, Information and Control*, (2007), DOI: 10.1109/ICICIC.2007.293.

[17] N.B., Khoolenjani, and F., Shahsanaie, Estimating the parameter of Exponential distribution under Type-II censoring from fuzzy data. *Journal of Statistical Theory and Applications*, **15**(1.2), (2016), 181-195.

[18] E., Kremer, A characterization of the generalized Pareto distribution with an application to reinsurance. *Blatter Der DGVFM*, **23**(1.1), (1997), 17-19.

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

[20] C.T., Lin, and W.Y., Wang, Estimation for the generalized Pareto distribution with censored data. *Communications in Statistics-Simulation and Computation*, **29**(1.4), (2000), 1183-1213.

[21] I., Makhdoom, P., Nasisri, and A., Pak, Estimating the Parameter of Exponential Distribution under Type II Censoring From Fuzzy Data. *Journal of Modern Applied Statistical Methods*, **15**(1.2), (2016), 495-509.

[22] T., Oztekun, Comparison of parameter estimation methods for the three-parameter generalized Pareto distribution. *Turkish Journal of Agriculture & Forestry*, **29**, (2005), 419-428.

[23] A., Pak, G.H., Parham, M., and Saraj, Inference for the Weibull Distribution Based on Fuzzy Data. *Revista Colombiana de Estadistica* , **36**(1.2), (2013), 339-358.

[24] A., Pak, G.H., Parham, M., and Saraj, M., Inferences on the Competing Risk Reliability Problem for Exponential Distribution Based on Fuzzy Data. *IEEE Transactions on Reliability*, **63**(1.1), (2014), 1-10.

[25] S.J., Press, *The Subjectivity of Scientists and the Bayesian Approach*. New York, NY, USA: Wiley, (2001).

[26] M.Z., Raqab, A., Asgharzadeh, and R., Valiollahi, Prediction for Pareto distribution based on progressively Type-II censored samples. *Computational Statistics and Data Analysis*. **54**, (2010), 1732-1743.

[27] R Development Core Team, *A Language and Environment for Statistical Computing: R Foundation for Statistical Computing*, Vienna, Austria, (2011).

[28] G., Shi, H.V., Atkinson, C.M., Sellars, and C.W., Anderson, Application of the generalized Pareto distribution to the estimation of the size of the maximum inclusion in clesn steels. *Acta Materialia*, **47**(1.5), (1999), 1455-1468.

[29] J., Simpson, Use of the gamma distribution in single-cloud rainfall analysis. *Monthly Weather Review*, **100**, (1972), 309-312.

[30] V.P., Singh, and H., Guo, Parameter estimation for 3-parameter generalized pareto distribution by the principle of maximum entropy. *Hydrological Sciences Journal*, 40(2), (1995), 165-181.

[31] H., Tanaka, T., Okuda, and K., Asai, Fuzzy information and decision in statistical model. In: *Advances in Fuzzy Sets Theory and Applications. North-Holland, Amsterdam*, (1979), 303-320.

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

[33] L.A., Zadeh, Probability measures of fuzzy events. *Journal of Mathematical Analysis and Application*, **10**, (1968), 421-427.