REVSTAT – Statistical Journal
Volume 21, Number 3, July 2023, 405–424
https://doi.org/10.57805/revstat.v21i3.174

Estimating the Parameters of Burr Type XII Distribution with Fuzzy Observations

Authors: Abbas Abdul Hussein
– Department of Mathematics, University of Thi-Qar, Thi-Qar, Iraq
abbas1989.aa1989@gmail.com

Riyadh Al-Mosawi ✉
– Department of Mathematics, University of Thi-Qar, Thi-Qar, Iraq
riyadhrm@gmail.com ; riyadhrm@utq.edu.iq

Received: February 2021 Revised: January 2022 Accepted: January 2022

Abstract:

- In this article, the classical as well as the Bayesian estimation problems of two-parameter Burr type XII distribution based on fuzzy data are considered. The maximum likelihood estimators via two methods, namely, Newton-Raphson and Expectation-Maximization algorithms are computed. The standard errors of the estimates are computed using the observed information matrix. For computing the Bayes estimators, three methods viz Lindley's approximation, Tierney–Kadane approximation and highest posterior density method are obtained. Monte-Carlo simulation experiments are conducted to investigate the performance of the proposed methods. Finally, the proposed methods are illustrated by using three different real data sets.

Keywords:

- *Bayesian estimation; Burr type XII distribution; expectation-maximization algorithm; fuzzy observations; Lindley's approximation; maximum likelihood estimation; Tierney–Kadane approximation.*

AMS Subject Classification:

- 62N02, 62N86.

✉ Corresponding author.

1. INTRODUCTION

The Burr type XII distribution was first introduced in the literature by Burr ([5]). It has gained special attention in the last two decades and applied in different fields including the area of reliability, failure time modeling and acceptance sampling plan and so on. The two-parameter Burr type XII distribution has the following probability density function

$$f(x;\alpha,\beta) = \alpha\beta x^{\alpha-1}(1+x^{\alpha})^{-(\beta+1)},\ x>0, \alpha>0, \beta>0, \tag{1.1}$$

where α and β represent the shape parameters. It is easy to see that when $\alpha = 1$, the Burr type XII reduces to the log-logistic distribution. Maximum likelihood and Bayesian inferential issues for the unknown parameters of Burr type XII distribution with different types of data were considered by several authors. See, for example, Wang *et al.* ([35]), Moore and Papadopoulos ([21]), Ghitany and Al-Awadhi ([11]), Mousa and Jaheen ([22]), Wahed ([34]), Li *et al.* ([17]), Jaheen and Okasha ([13]), Panahi and Asadi ([29]), Al-Baldawi *et al.* ([1]), Rao *et al.* ([31]), Belaghi *et al.* ([3]) and Hakim *et al.* ([12]).

All the earlier works on the estimation of the parameters of the Burr type XII distribution have been done under the assumption of precise data. In the classical estimation theory, we consider only one source of uncertainty available, namely randomness. However, in many practical situations, in addition to the randomness, we may face other source of uncertainties, namely, vague uncertainty. Vagueness occurs as a result of imprecisely recording or measuring the observations due to, for example, machine errors, human errors, etc. For instance, the lifetime of a specific electric device may be recorded as vague statements like "about 3 years", "approximately less than 2 years", "approximately 3 years", "approximately between 3 and 4 years" and so on.

In recent years, many papers extended the statistical methods to analysis of fuzzy data for different distributions. Among others, Denœux ([8]), for a general parametric statistical model, showed that the EM algorithm may be used for analyzing statistical problems involving fuzzy data. Pak *et al.* ([27]) investigated different classical and Bayesian methods for estimating the parameters of Weibull distribution when the available data are in the form of fuzzy numbers. Pak *et al.* ([28]) discussed different procedures for estimating the parameter of Rayleigh distribution under doubly type II censoring when the available observations are described by means of fuzzy information. They computed the maximum likelihood, highest posterior density and method of moments estimators. Makhdoom *et al.* ([20]) estimated the parameter of exponential distribution on the basis of type II censoring scheme when the available data are in the form of fuzzy numbers. The Bayes estimate of the unknown parameter was also obtained under the assumption of gamma prior. Khoolenjani and Shahsanaie ([15]) derived the maximum likelihood estimator of the mean of exponential distribution under type II censoring scheme when the lifetime observations are in the form of fuzzy numbers. They also obtained the estimate, via Bayesian method, of the unknown parameter. Pak ([23]) obtained the maximum likelihood estimation and Bayesian estimation for Lindley distribution when the available observations are reported in the form of fuzzy data. The classical and Bayesian inferences for the Pareto distribution of life time fuzzy observations was studied by Shafiq ([32]). Chaturvedi *et al.* ([6]) presented procedures of parameter estimation of the Rayleigh distribution based on type II progressively hybrid censored fuzzy lifetime data.

Classical as well as the Bayesian procedures for the estimation of unknown parameters were investigated. Pak and Mahmoudi ([26]) estimated the parameters of Lomax distribution when the available observations are described by means of fuzzy information. They computed the maximum likelihood and the Bayesian estimators. Basharat *et al.* ([2]) derived the distribution of a linear combination of two independent exponential random variables. The parameter estimates of the proposed distribution were obtained by using the maximum likelihood estimation method and the method of moments from fuzzy data. Finally, Pak *et al.* ([25]) provided Bayesian inference for the parameters of the generalized exponential model under asymmetric and symmetric loss functions when the observations are described in terms of fuzzy numbers.

To the best of our knowledge, there are no studies focused on the analysis of fuzzy data on the parameter estimation of two-parameter Burr type XII distribution. The main purpose of this paper is to investigate the inferential procedures for the distribution of the two parameters of Burr type XII, where the available data is in the form of fuzzy data. In Section 2, we review the basic notations and definitions of fuzzy set theory. In Section 3, we address the estimation of the unknown parameters of the maximum likelihood estimates using the Newton–Raphson and expectation-maximization (EM) algorithm. In Section 4, the Bayes estimates of the unknown parameters are obtained via Lindley's approximation, Tierney–Kadane approximation and highest posterior distribution estimation method under the assumption of Gamma priors. A Monte Carlo simulation study is conducted in Section 5, to assess the performance of the proposed estimators. For illustration, analyses of three datasets are provided. Finally, some conclusions are provided in Section 6.

2. BASIC DEFINITION OF FUZZY SETS

In this section, we review some basic definitions and notations of fuzzy sets and fuzzy probability theory used in this paper. Suppose a random experiment with a probability space $(\mathbb{R}^m, \mathcal{B}^m, P_\theta)$, where $\mathbb{R}^m$ is an m-dimensional Euclidean space, $\mathcal{B}^m$ is the smallest Borel σ-field defined on $\mathbb{R}^m$ and P_θ, $\theta \in \Theta$, is a probability measure defined on $\mathcal{B}^m$. In many applications, we have a situation that the outcome of the experiment cannot be observed exactly and only partial information is available. For example, the lifetime of a specific electric device may be recorded as "about 3 years", "approximately less than 2 years", "approximately 3 years", "approximately between 3 and 4 years" and so on. These lifetimes can be modeled and described in the form of fuzzy subset. A fuzzy set $\tilde{A}$ in $\mathbb{R}^m$ is characterized by a membership function $\mu_{\tilde{A}} : \mathbb{R}^m \to [0,1]$, where $\mu_{\tilde{A}}(x)$, $x \in \mathbb{R}^m$, represents the degree of membership of x in $\tilde{A}$. A fuzzy event is a fuzzy set whose membership function is Borel measurable function. According to Zadeh ([36]) the probability of a fuzzy event $\tilde{A}$ is computed by

$$P(\tilde{A}) = \int \mu_{\tilde{A}}(x) dP_\theta. \tag{2.1}$$

The most common fuzzy subsets that are frequently encountered in fuzzy statistical analysis are the fuzzy numbers and among them, the triangular fuzzy numbers are the most common type. A triangular fuzzy number, written as $\tilde{x} = (a, b, c)$, has the following membership function

$$\mu_{\tilde{x}}(x) = \begin{cases} \dfrac{x-a}{b-a}, & \text{if } a \le x \le b, \\ \dfrac{c-x}{c-b}, & \text{if } b \le x \le c, \\ 0, & \text{otherwise.} \end{cases}$$

In particular, assume X to be a random variable with a probability density function (p.d.f.) $g(x;\theta)$ that is absolutely continuous with probability measure P_θ. The conditional probability of a crisp (non-fuzzy) set A given a fuzzy set $\tilde{B}$ is given by (see Denœux ([8]))

$$P(A|\tilde{B}) = \frac{\int_A \mu_{\tilde{B}}(x) g(x;\theta)dx}{\int \mu_{\tilde{B}}(x) g(x;\theta)dx}.$$

Consequently, the conditional density of X given $\tilde{B}$ can thus be computed by

$$g(x|\tilde{B}) = \frac{\mu_{\tilde{B}}(x) g(x;\theta)}{\int \mu_{\tilde{B}}(x) g(x;\theta)dx}.$$

3. MAXIMUM LIKELIHOOD ESTIMATION

Let $X_1, X_2, ..., X_n$ denote a random sample of size n from Burr type XII distribution with p.d.f. given in (1.1). Let $\mathbf{X} = (X_1, X_2, ..., X_n)$ denote the corresponding random vector. If a realization $\mathbf{x}$ of $\mathbf{X}$ was exactly observed, the likelihood function can be written as

$$L(\alpha, \beta|\mathbf{x}) = (\alpha\beta)^n \prod_{i=1}^{n} x_i^{\alpha-1}(1+x_i^\alpha)^{-\beta-1}. \tag{3.1}$$

Suppose now $\mathbf{x}$ is not observed precisely, and only partial information about $\mathbf{x}$ is available in form of fuzzy observation $\tilde{\mathbf{x}} = (\tilde{x_1}, ..., \tilde{x_n})$ with Borel measurable membership function $\mu_{\tilde{\mathbf{x}}}(\mathbf{x}) = (\mu_{\tilde{x_1}}(x), ..., \mu_{\tilde{x_n}}(x))$. Then, based on fuzzy observation $\tilde{\mathbf{x}}$, the log-likelihood function reduces to

$$\begin{aligned} l(\alpha, \beta|\tilde{\mathbf{x}}) &= n\log\alpha + n\log\beta + \sum_{i=1}^{n} \log \int x^{\alpha-1}(1+x^\alpha)^{-\beta-1}\mu_{\tilde{x}_i}(x)dx \\ &= n\log\alpha + n\log\beta + \sum_{i=1}^{n} \log \int A(x)\mu_{\tilde{x}_i}(x)dx, \end{aligned} \tag{3.2}$$

where

$$A(x) = x^{\alpha-1}(1+x^\alpha)^{-\beta-1}. \tag{3.3}$$

The maximum likelihood estimate of the parameters α and β can be obtained by maximizing the log-likelihood $l(\alpha, \beta|\tilde{\mathbf{x}})$ with respect to α and β. First we need to prove the following result.

Theorem 3.1. *The MLEs of α and β for $\alpha > 0$ and $\beta > 0$ exist and unique.*

Proof: The detailed proof of the theorem is deferred in the Appendix. □

By taking the partial derivatives of the log-likelihood $l(\alpha,\beta|\tilde{\mathbf{x}})$ with respect to α and β and equating the resulted equations to zero, we get the following two normal equations

$$\frac{\partial l(\alpha,\beta|\tilde{\mathbf{x}})}{\partial \alpha} \equiv l_\alpha = \frac{n}{\alpha} + \sum_{i=1}^{n} \frac{\int A_\alpha(x)\mu_{\tilde{x}_i}(x)dx}{\int A(x)\mu_{\tilde{x}_i}(x)dx} = 0 \tag{3.4}$$

and

$$\frac{\partial l(\alpha,\beta|\tilde{\mathbf{x}})}{\partial \beta} \equiv l_\beta = \frac{n}{\beta} + \sum_{i=1}^{n} \frac{\int A_\beta(x)\mu_{\tilde{x}_i}(x)dx}{\int A(x)\mu_{\tilde{x}_i}(x)dx} = 0, \tag{3.5}$$

where

$$A_\alpha(x) \equiv \frac{\partial A(x)}{\partial \alpha} = (1+x^\alpha)^{-\beta-2}x^{\alpha-1}\log(x)[1-\beta x^\alpha],$$
$$A_\beta(x) \equiv \frac{\partial A(x)}{\partial \beta} = -x^{\alpha-1}(1+x^\alpha)^{-\beta-1}\log(1+x^\alpha).$$

Since there are no closed forms to the normal equations (3.4) and (3.5), iterative numerical methods can be used to obtain the MLEs. In this section, we propose two methods to compute the MLEs of α and β, namely; Newton–Raphson method and EM method.

3.1. Newton–Raphson algorithm

The Newton–Raphson (NR) method is a numerical approach that is commonly used to compute MLEs of the unknown parameters. In this method, the solution of the likelihood function is obtained through an iterative procedure. First, we obtain the second-order derivatives of the log-likelihood with respect to α and β in order to implement the NR method:

$$l_{\alpha\alpha} = \frac{-n}{\alpha^2} + \sum_{i=1}^{n} \frac{\int A(x)\mu_{\tilde{x}_i}(x)dx \int A_{\alpha\alpha}(x)\mu_{\tilde{x}_i}(x)dx - (\int A_\alpha(x)\mu_{\tilde{x}_i}(x)dx)^2}{(\int A(x)\mu_{\tilde{x}_i}(x)dx)^2}, \tag{3.6}$$

$$l_{\beta\beta} = \frac{-n}{\beta^2} + \sum_{i=1}^{n} \frac{\int A(x)\mu_{\tilde{x}_i}(x)dx \int A_{\beta\beta}(x)\mu_{\tilde{x}_i}(x)dx - (\int A_\beta(x)\mu_{\tilde{x}_i}(x)dx)^2}{(\int A(x)\mu_{\tilde{x}_i}(x)dx)^2}, \tag{3.7}$$

$$l_{\alpha\beta} = \sum_{i=1}^{n} \frac{\int A(x)\mu_{\tilde{x}_i}(x)dx \int A_{\alpha\beta}(x)\mu_{\tilde{x}_i}(x)dx - \int A_\alpha(x)\mu_{\tilde{x}_i}(x)dx \int A_\beta(x)\mu_{\tilde{x}_i}(x)dx}{(\int A(x)\mu_{\tilde{x}_i}(x)dx)^2}, \tag{3.8}$$

where

$$A_{\alpha\alpha}(x) = x^{\alpha-1}(\log(x))^2(1+x^\alpha)^{-\beta-3}\Big[x^{2\alpha}(\beta+1)(\beta+2) - 3x^\alpha(\beta+1)(1+x^\alpha) + (1+x^\alpha)^2\Big],$$
$$A_{\beta\beta}(x) = x^{\alpha-1}(1+x^\alpha)^{-\beta-1}(\log(1+x^\alpha))^2$$
$$A_{\alpha\beta}(x) = x^{2\alpha-2}(1+x^\alpha)^{-\beta-2}\log(x)[(\beta+1)x\log(1+x^\alpha) - (1+x^\alpha)\log(1+x^\alpha) - x].$$

Assume $\alpha^{(k)}$ and $\beta^{(k)}$ are the values of α and β at the k-th iteration. Then at $(k+1)$-th iteration, the updated values of α and β are obtained as

$$\begin{pmatrix} \alpha^{(k+1)} \\ \beta^{(k+1)} \end{pmatrix} = \begin{pmatrix} \alpha^{(k)} \\ \beta^{(k)} \end{pmatrix} - \begin{pmatrix} l_{\alpha\alpha} & l_{\alpha\beta} \\ l_{\alpha\beta} & l_{\beta\beta} \end{pmatrix}^{-1}_{\alpha=\alpha^{(k)},\beta=\beta^{(k)}} \begin{pmatrix} l_{\alpha} \\ l_{\beta} \end{pmatrix}_{\alpha=\alpha^{(k)},\beta=\beta^{(k)}},$$

which is equivalent to

$$\alpha^{(k+1)} = \alpha^{(k)} - \frac{l_\alpha l_{\beta\beta} - l_\beta l_{\alpha\beta}}{l_{\alpha\alpha} l_{\beta\beta} - l^2_{\alpha\beta}}\Big|_{\alpha=\alpha^{(k)},\beta=\beta^{(k)}}, \tag{3.9}$$

$$\beta^{(k+1)} = \beta^{(k)} - \frac{l_\beta l_{\alpha\alpha} - l_\alpha l_{\alpha\beta}}{l_{\alpha\alpha} l_{\beta\beta} - l^2_{\alpha\beta}}\Big|_{\alpha=\alpha^{(k)},\beta=\beta^{(k)}}. \tag{3.10}$$

The iteration process then continues until convergence, i.e., $|\alpha^{(k+1)} - \alpha^{(k)}| + |\beta^{(k+1)} - \beta^{(k)}| < \varepsilon$, for some pre-specified $\varepsilon > 0$.

To estimate the standard error of maximum likelihood estimators, $\hat{\alpha}$ and $\hat{\beta}$, we use the observed information matrix method. The variance-covariance matrix of the MLEs of α and β is defined as

$$\Sigma = \begin{bmatrix} \mathrm{var}(\hat{\alpha}) & \mathrm{cov}(\hat{\alpha}, \hat{\beta}) \\ \mathrm{cov}(\hat{\alpha}, \hat{\beta}) & \mathrm{var}(\hat{\beta}) \end{bmatrix},$$

and can be estimated by using the inverse of the observed information matrix

$$I(\hat{\alpha}, \hat{\beta}) = \begin{pmatrix} -l_{\alpha\alpha} & -l_{\alpha\beta} \\ -l_{\alpha\beta} & -l_{\beta\beta} \end{pmatrix}_{\alpha=\hat{\alpha},\beta=\hat{\beta}}, \tag{3.11}$$

where $l_{\alpha\alpha}, l_{\beta\beta}$ and $l_{\alpha\beta}$ are given in (3.6),(3.7) and (3.8), respectively. Then the $100(1-\gamma)\%$ Wald confidence intervals of α and β using the observed information matrix can be constructed, respectively, as

$$\hat{\alpha} \pm z_{\gamma/2}\sqrt{\mathrm{var}(\hat{\alpha})} \text{ and } \hat{\beta} \pm z_{\gamma/2}\sqrt{\mathrm{var}(\hat{\beta})},$$

where z_p is the upper p-th percentile of the standard normal distribution.

It is known that Newton–Raphson method is very sensitive to the initial values of parameters. In addition, the calculation of the second-order derivatives of the log-likelihood based on fuzzy data sometimes can be rather tedious. So we propose to use an alternative method to the Newton–Raphson method which is the EM algorithm.

3.2. EM Algorithm

In this subsection, we propose to use the EM algorithm to calculate the MLEs of the unknown parameters.

The EM algorithm, proposed by Dempster *et al.* ([7]), is a very powerful technique used in parameter estimation based on incomplete or missing information data. As stated by Pradhan and Kundu ([30]), the EM algorithm is an iterative method and each iteration consists of two main steps; Expectation(E)-step and Maximization(M)-step. In E-step, we form the "pseudo-likelihood" function by replacing the incomplete or missing observations in the likelihood function with their corresponding expected values. In the M-step, we maximize the "pseudo-likelihood" function with respect to the parameters. Let us denote the observed data set by $\tilde{\mathbf{X}} = (\tilde{X}_1, ..., \tilde{X}_n)$ and let the complete data denoted by $\mathbf{X} = (X_1, ..., X_n)$. Define $\mathbf{Z} = (Z_1, ..., Z_n)$ where Z_i represents the conditional expectation of the complete observation X_i given the corresponding fuzzy observation $\tilde{X}_i$ with membership function $\mu_{\tilde{x}_i}(x)$. Observe that

$$Z_i = E(X_i|\tilde{X}_i) = \frac{\int x f(x;\alpha,\beta)\mu_{\tilde{x}_i}(x)dx}{\int f(x;\alpha,\beta)\mu_{\tilde{x}_i}(x)dx}, \qquad i = 1, ..., n. \tag{3.12}$$

Then the pseudo likelihood function takes the form

$$L^c(\alpha,\beta|\mathbf{z}) = (\alpha\beta)^n \prod_{i=1}^{n} z_i^{\alpha-1}(1+z_i^{\alpha})^{-\beta-1}, \tag{3.13}$$

with pseudo log-likelihood function

$$l^c(\alpha,\beta|\mathbf{z}) = n\log\alpha + n\log\beta + (1-\alpha)\sum_{i=1}^{n}\log(z_i) - (\beta+1)\sum_{i=1}^{n}\log(1+z_i^{\alpha}). \tag{3.14}$$

By taking the partial derivatives of l^c with respect to α and β, respectively, and equating the resulted equations to zero we obtain the following equations:

$$\frac{n}{\alpha} + \sum_{i=1}^{n}\log(z_i) - (\beta+1)\sum_{i=1}^{n}\frac{z_i^{\alpha}\log(z_i)}{(1+z_i^{\alpha})} = 0, \tag{3.15}$$

$$\frac{n}{\beta} - \sum_{i=1}^{n}\log(1+z_i^{\alpha}) = 0. \tag{3.16}$$

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

Step 1. Given starting values of α and β, say $\alpha^{(0)}$ and $\beta^{(0)}$, and take k=0.

Step 2. At the $(k+1)$-th iteration,

Step 2.1. E-step. Evaluate $\mathbf{Z} = (Z_1, ..., Z_n)$, where $Z_i \equiv Z_i(\alpha^{(k)}, \alpha^{(k)})$ is computed using the expression (3.12) with α and β are replaced by $\alpha^{(k)}$ and $\beta^{(k)}$, respectively.

Step 2.2. M-step. Solve the equations (3.15) and (3.16) and obtain the next values $\alpha^{(k+1)}$ and $\beta^{(k+1)}$ of α and β, respectively.

Step 3. If $|\alpha^{(k+1)} - \alpha^{(k)}| + |\beta^{(k+1)} - \beta^{(k)}| < \varepsilon$, for some pre-specified value $\varepsilon > 0$, then set $\alpha^{(k+1)}$ and $\beta^{(k+1)}$ as the maximum likelihood estimators of α and β, otherwise, set $k = k+1$ and go to **Step 2**.

Estimating the standard errors and constructing the confidence intervals in this section are similar to that given in Section 2 with NR estimates are replaced by EM estimates.

4. BAYESIAN ESTIMATION

In this section, we estimate the unknown parameters of Burr type XII distribution using Bayesian method under squared error loss function. The Bayes estimators are obtained using three different methods; Lindley's approximation, Tierney–Kadane approximation and highest posterior density methods. Assume that the parameters α and β have independent gamma priors such that $\alpha \sim \pi_1(\alpha) = Gamma(a, b)$ and $\beta \sim \pi_2(\beta) = Gamma(c, d)$. Based on the above priors, the joint posterior density function of α and β given the data can be written as follows

$$\pi(\alpha, \beta|\tilde{\mathbf{x}}) = \frac{\alpha^{n+a-1}\beta^{n+c-1}e^{-b\alpha-d\beta}\prod_{i=1}^{n}\int_0^\infty x^{\alpha-1}(1+x^\alpha)^{-\beta-1}\mu_{\tilde{x}_i}(x)dx}{\int_0^\infty\int_0^\infty \alpha^{n+a-1}\beta^{n+c-1}e^{-b\alpha-d\beta}\prod_{i=1}^{n}\int_0^\infty x^{\alpha-1}(1+x^\alpha)^{-\beta-1}\mu_{\tilde{x}_i}(x)dxd\alpha d\beta}. \tag{4.1}$$

Then, under a squared error loss function, the Bayes estimate of any function of α and β, say $g(\alpha, \beta)$, is given by

$$E(g(\alpha, \beta)|\tilde{\mathbf{x}}) = \int_0^\infty \int_0^\infty g(\alpha, \beta)\pi(\alpha, \beta|\tilde{\mathbf{x}})d\alpha d\beta. \tag{4.2}$$

Note that Equation (4.2) cannot be obtained analytically; therefore, in the following, we propose to use three methods, namely; Lindley's approximation and Tierney–Kadane approximation and highest posterior density methods to solve it and compute the Bayes estimators.

4.1. Lindley's Approximation

Lindley's approximation was proposed by Lindley ([18]) to approximate the integrals involved in Bayes estimator. Lindley proposed a ratio of integrals of the form

$$E(g(\alpha, \beta)|\tilde{\mathbf{x}}) = \frac{\int_0^\infty \int_0^\infty g(\alpha, \beta)e^{Q(\alpha,\beta)}d\alpha d\beta}{\int_0^\infty \int_0^\infty e^{Q(\alpha,\beta)}d\alpha d\beta} \tag{4.3}$$

that can be approximated by

$$\begin{aligned} \hat{g}(\alpha, \beta) &= g(\hat{\alpha}, \hat{\beta}) + \frac{1}{2}\Big[(\hat{g}_{\alpha\alpha} + 2\hat{g}_\alpha\hat{\rho}_\alpha)\hat{\sigma}_{\alpha\alpha} + (\hat{g}_{\alpha\beta} + 2\hat{g}_\beta\hat{\rho}_\alpha)\hat{\sigma}_{\alpha\beta} + (\hat{g}_{\alpha\beta} + 2\hat{g}_\alpha\hat{\rho}_\beta)\hat{\sigma}_{\alpha\beta} \\ &+ (\hat{g}_{\beta\beta} + 2\hat{g}_\beta\hat{\rho}_\beta)\hat{\sigma}_{\beta\beta}\Big] + \frac{1}{2}\Big[(\hat{g}_\alpha\hat{\sigma}_{\alpha\alpha} + \hat{g}_\beta\hat{\sigma}_{\alpha\beta})(l_{\alpha\alpha\alpha}\hat{\sigma}_{\alpha\alpha} + 2\hat{l}_{\alpha\alpha\beta}\hat{\sigma}_{\alpha\beta} + \hat{l}_{\alpha\beta\beta}\hat{\sigma}_{\beta\beta}) \\ &+ (\hat{g}_\alpha\hat{\sigma}_{\alpha\beta} + \hat{g}_\beta\hat{\sigma}_{\beta\beta})(\hat{l}_{\alpha\alpha\beta}\hat{\sigma}_{\alpha\alpha} + 2\hat{l}_{\alpha\beta\beta}\hat{\sigma}_{\alpha\beta} + \hat{l}_{\beta\beta\beta}\hat{\sigma}_{\beta\beta})\Big], \end{aligned} \tag{4.4}$$

where

$$Q(\alpha, \beta) = \log[\pi_1(\alpha)\pi_2(\beta)] + \log L(\alpha, \beta|\tilde{\mathbf{x}}) \equiv \rho(\alpha, \beta) + \ell(\alpha, \beta|\tilde{\mathbf{x}}).$$

The expressions $\hat{l}, \hat{g}, \hat{\rho}$ and $\hat{\sigma}$ denote, respectively, the functions l, g, ρ and σ evaluated at $\hat{\alpha}$ and $\hat{\beta}$, the MLEs of α and β. Here, the expressions $\hat{g}_\alpha$, $\hat{g}_\beta$, $\hat{g}_{\alpha\alpha}$, $\hat{g}_{\alpha\beta}$ and $\hat{g}_{\beta\beta}$ denote the first and the second order partial derivatives of g with respect α and β evaluated at the MLEs of α and β. First note that, the expressions of $l_\alpha, l_\beta, l_{\alpha\alpha}, l_{\beta\beta}$ and $l_{\alpha\beta}$ are given in (3.4), (3.5), (3.6), (3.7) and (3.8), respectively. The third order of partial derivatives of the log-likelihood function with respect to α and β are given by

$$\begin{aligned}
l_{\alpha\alpha\alpha} &= \frac{2n}{\alpha^3} + \sum_{i=1}^{n} \frac{C_i^2 C_{i,\alpha\alpha\alpha} - 3C_i C_{i,\alpha} C_{i,\alpha\alpha} + 2C_{i,\alpha}^3}{C_i^3}, \\
l_{\beta\beta\beta} &= \frac{2n}{\beta^3} + \sum_{i=1}^{n} \frac{C_i^2 C_{i,\beta\beta\beta} - 3C_i C_{i,\beta} C_{i,\beta\beta} + 2C_{i,\beta}^3}{C_i^3}, \\
l_{\alpha\beta\beta} &= \sum_{i=1}^{n} \frac{C_i^2 C_{i,\alpha\beta\beta} - 2C_i C_{i,\beta} C_{i,\alpha\beta} - C_i C_{i,\alpha} C_{i,\beta\beta} + 2C_{i,\alpha} C_{i,\beta}^2}{C_i^3}, \\
l_{\alpha\alpha\beta} &= \sum_{i=1}^{n} \frac{C_i^2 C_{i,\alpha\alpha\beta} - 2C_i C_{i,\alpha} C_{i,\alpha\beta} - C_i C_{i,\alpha\alpha} C_{i,\beta} + 2C_{i,\alpha}^2 C_{i,\beta}}{C_i^3},
\end{aligned}$$

where

$$\begin{aligned}
C_i &= \int A(x)\mu_{\tilde{x}_i}(x)dx, \\
C_{i,\alpha} &= \int A_\alpha(x)\mu_{\tilde{x}_i}(x)dx, C_{i,\alpha\alpha} = \int A_{\alpha\alpha}(x)\mu_{\tilde{x}_i}(x)dx, C_{i,\alpha\alpha\alpha} = \int A_{\alpha\alpha\alpha}(x)\mu_{\tilde{x}_i}(x)dx, \\
C_{i,\beta} &= \int A_\beta(x)\mu_{\tilde{x}_i}(x)dx, C_{i,\beta\beta} = \int A_{\beta\beta}(x)\mu_{\tilde{x}_i}(x)dx, C_{i,\beta\beta\beta} = \int A_{\beta\beta\beta}(x)\mu_{\tilde{x}_i}(x)dx, \\
C_{i,\alpha\beta} &= \int A_{\alpha\beta}(x)\mu_{\tilde{x}_i}(x)dx, C_{i,\alpha\alpha\beta} = \int A_{\alpha\alpha\beta}(x)\mu_{\tilde{x}_i}(x)dx, C_{\alpha\beta\beta} = \int A_{\alpha\beta\beta}(x)\mu_{\tilde{x}_i}(x)dx,
\end{aligned}$$

and

$$\begin{aligned}
A_{\alpha\alpha\alpha}(x) &= x^{2\alpha-1}(\beta+1)(\log(x))^3(1+x^\alpha)^{-\beta-4}\Big[-x^{2\alpha}(\beta+2)(\beta+3) \\
&+ 6x^\alpha(1+x^\alpha)(\beta+2) - 7(1+x^\alpha)^2\Big] + x^{\alpha-1}(\log(x))^3(1+x^\alpha)^{-\beta-1}, \\
A_{\beta\beta\beta}(x) &= -x^{\alpha-1}(\log(1+x^\alpha))^3(1+x^\alpha)^{-\beta-1}, \\
A_{\alpha\beta\beta}(x) &= x^{\alpha-1}\log(1+x^\alpha)\log(x)(1+x^\alpha)^{-\beta-2}\Big[-x^\alpha(\beta+1)\log(1+x^\alpha), \\
&+ 2x^\alpha + \log(1+x^\alpha)(1+x^\alpha)\Big], \\
A_{\alpha\alpha\beta}(x) &= (\beta+1)(\log(x))^2 x^{2\alpha-1}(1+x^\alpha)^{-\beta-3}\Big[-x^\alpha(\beta+2)\log(1+x^\alpha) + x^\alpha \\
&+ 3(1+x^\alpha)\log(1+x^\alpha)\Big] + (\log(x))^2 x^{2\alpha-1}(1+x^\alpha)^{-\beta-3}\Big[x^\alpha(\beta+2) \\
&- 3(1+x^\alpha)\Big] - (\log(x))^2 x^{\alpha-1}(1+x^\alpha)^{-\beta-1}\log(1+x^\alpha).
\end{aligned}$$

The function ρ given by

$$\rho(\alpha,\beta) = (a-1)\log(\alpha) - b\alpha + (c-1)\log(\beta) - d\beta$$

has the following partial derivatives:

$$\begin{aligned}
\rho_\alpha &= \frac{\partial\rho(\alpha,\beta)}{\partial\alpha} = \frac{a-1}{\alpha} - b, \\
\rho_\beta &= \frac{\partial\rho(\alpha,\beta)}{\partial\beta} = \frac{c-1}{\beta} - d.
\end{aligned}$$

In addition

$$\begin{pmatrix} \sigma_{\alpha\alpha} & \sigma_{\alpha\beta} \\ \sigma_{\alpha\beta} & \sigma_{\beta\beta} \end{pmatrix} = \begin{pmatrix} -l_{\alpha\alpha} & -l_{\alpha\beta} \\ -l_{\alpha\beta} & -l_{\beta\beta} \end{pmatrix}^{-1}.$$

If $g(\alpha,\beta)=\alpha$, we obtain $g_\alpha=1$ and $g_{\alpha\alpha}=g_\beta=g_{\beta\beta}=g_{\alpha\beta}=0$. Thus the Bayes estimator using Lindley's approximation is given by

$$\begin{aligned}\hat{\alpha} &= \hat{\alpha}_{MLE}+\hat{\rho}_\alpha\hat{\sigma}_{\alpha\alpha}+\hat{\rho}_\beta\hat{\sigma}_{\beta\alpha}+\frac{1}{2}\Big[\hat{\sigma}_{\alpha\alpha}(\hat{l}_{\alpha\alpha\alpha}\hat{\sigma}_{\alpha\alpha}+\hat{l}_{\alpha\alpha\beta}\hat{\sigma}_{\alpha\beta}+\hat{l}_{\alpha\alpha\beta}\hat{\sigma}_{\beta\alpha}+\hat{l}_{\alpha\beta\beta}\hat{\sigma}_{\beta\beta})\\ &+ (\hat{\sigma}_{\beta\alpha})(\hat{l}_{\alpha\alpha\beta}\hat{\sigma}_{\alpha\alpha}+\hat{l}_{\alpha\beta\beta}\hat{\sigma}_{\alpha\beta}+\hat{l}_{\alpha\beta\beta}\sigma\hat{_{\beta\alpha}}+\hat{l}_{\beta\beta\beta}\hat{\sigma}_{\beta\beta})\Big].\end{aligned}$$

If $g(\alpha,\beta)$=β, we obtain $g_\beta=1$ and $g_{\alpha\alpha}=g_\alpha=g_{\beta\beta}=g_{\alpha\beta}=0$. Then the Bayes estimates of β is given by

$$\begin{aligned}\hat{\beta} &= \hat{\beta}_{MLE}+\hat{\rho}_\alpha\hat{\sigma}_{\beta\alpha}+\hat{\rho}_\beta\hat{\sigma}_{\beta\beta}+\frac{1}{2}\Big[\hat{\sigma}_{\alpha\beta}(\hat{l}_{\alpha\alpha\alpha}\hat{\sigma}_{\alpha\alpha}+\hat{l}_{\alpha\alpha\beta}\hat{\sigma}_{\alpha\beta}+\hat{l}_{\alpha\alpha\beta}\hat{\sigma}_{\beta\alpha}+\hat{l}_{\alpha\beta\beta}\hat{\sigma}_{\beta\beta})\\ &+ (\hat{\sigma}_{\beta\beta})(\hat{l}_{\alpha\alpha\beta}\hat{\sigma}_{\alpha\alpha}+\hat{l}_{\alpha\beta\beta}\hat{\sigma}_{\alpha\beta}+\hat{l}_{\alpha\beta\beta}\hat{\sigma}_{\beta\alpha}+\hat{l}_{\beta\beta\beta}\hat{\sigma}_{\beta\beta})\Big].\end{aligned}$$

4.2. Tierney–Kadane approximation

In this subsection, we utilize another approximation of the integral (4.2) to compute the Bayes estimators. Using Laplace transformation, Tierney and Kadane [33] proposed an alternative method to approximate the ratio of integrals. The advantage of using Tierney–Kadane method is that it requires only the first and the second derivatives of the posterior density. The posterior expectation of a $g(\alpha,\beta)$ can be written as

$$E(g(\alpha,\beta|\tilde{\tilde{x}})) = \frac{\int_0^\infty\int_0^\infty e^{nH^*(\alpha,\beta)}d\alpha d\beta}{\int_0^\infty\int_0^\infty e^{nH(\alpha,\beta)}d\alpha d\beta}, \tag{4.5}$$

where

$$\begin{aligned} H(\alpha,\beta) &= \frac{1}{n}\Big[(a-1)\log(\alpha)-b\alpha+(c-1)\log(\beta)-d\beta+l(\alpha,\beta|\tilde{\mathbf{x}})\Big],\\ H^*(\alpha,\beta) &= H(\alpha,\beta)+\frac{1}{n}\log(g(\alpha,\beta)).\end{aligned}$$

Then the integral given in Equation (4.5) can be approximated by

$$\hat{g}(\alpha,\beta) = \Big(\frac{\det\sum^*}{\det\sum}\Big)^{\frac{1}{2}}\exp\{n[H^*(\bar{\alpha}^*,\bar{\beta}^*)-H(\bar{\alpha},\bar{\beta})]\}, \tag{4.6}$$

where $(\bar{\alpha}^*,\bar{\beta}^*)$ and $(\bar{\alpha},\bar{\beta})$ maximize H^* and H, respectively, $\sum^*$ and $\sum$ are the negatives of the inverse Hessian matrix of H^* and H evaluated at $(\bar{\alpha}^*,\bar{\beta}^*)$ and $(\bar{\alpha},\bar{\beta})$, respectively. Therefore $(\bar{\alpha},\bar{\beta})$ can be obtained by solving the following two equations

$$\begin{aligned} H_\alpha &= \frac{\partial H(\alpha,\beta)}{\partial\alpha} = \frac{a-1}{\alpha}-b+l_\alpha(\alpha,\beta|\tilde{\mathbf{x}}) = 0,\\ H_\beta &= \frac{\partial H(\alpha,\beta)}{\partial\beta} = \frac{c-1}{\beta}-d+l_\beta(\alpha,\beta|\tilde{\mathbf{x}}) = 0,\end{aligned}$$

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

$$\det \sum = \left(\bar{H}_{\alpha\alpha}\bar{H}_{\beta\beta} - \bar{H}^2_{\alpha\beta}\right)^{-1},$$

where

$$\bar{H}_{\alpha\alpha} \equiv \frac{\partial \bar{H}_\alpha}{\partial \alpha} = -\frac{a-1}{\bar{\alpha}^2} + l_{\alpha\alpha}(\bar{\alpha}, \bar{\beta}|\tilde{\mathbf{x}}),$$

$$\bar{H}_{\beta\beta} \equiv \frac{\partial \bar{H}_\beta}{\partial \beta} = -\frac{a-1}{\bar{\beta}^2} + l_{\beta\beta}(\bar{\alpha}, \bar{\beta}|\tilde{\mathbf{x}}),$$

$$\bar{H}_{\alpha\beta} \equiv \frac{\partial \bar{H}_\alpha}{\partial \beta} = l_{\alpha\beta}(\bar{\alpha}, \bar{\beta}|\tilde{\mathbf{x}}).$$

Similarly, for the function $H^*(\alpha, \beta)$, the determinant of the negative of the inverse Hessian of $H^*(\alpha, \beta)$ evaluated at $(\bar{\alpha^*}, \bar{\beta^*})$ is given by

$$\det \sum{}^* = (\bar{H}^*_{\alpha\alpha}\bar{H}^*_{\beta\beta} - \bar{H}^{*2}_{\alpha\beta})^{-1}.$$

For g$(\alpha, \beta) = \alpha$, we get

$$H^*_\alpha(\alpha, \beta) = H(\alpha, \beta) + \frac{1}{n}\log(\alpha)$$

and consequently, we have

$$H^*_{\alpha,\alpha} = \frac{\partial H^*(\alpha, \beta)}{\partial \alpha} = H_\alpha + \frac{1}{n\alpha},$$

$$H^*_{\alpha,\beta} = \frac{\partial H^*(\alpha, \beta)}{\partial \beta} = H_\beta,$$

$$H^*_{\alpha,\alpha\beta} = \frac{\partial H^*(\alpha, \beta)}{\partial \alpha\beta} = H_{\alpha\beta},$$

$$H^*_{\alpha,\alpha\alpha} = \frac{\partial H^*_1}{\partial \alpha} = H_{\alpha\alpha} - \frac{1}{n\alpha^2},$$

$$H^*_{\alpha,\beta\beta} = \frac{\partial H^*_2}{\partial \beta} = H_{\beta\beta}.$$

For $g(\alpha, \beta) = \beta$, we have

$$H^*_\beta(\alpha, \beta) = \frac{1}{n}\log(\beta) + H(\alpha, \beta)$$

and

$$H^*_{\beta,\alpha} = \frac{\partial H^*(\alpha, \beta)}{\partial \alpha} = H_\alpha,$$

$$H^*_{\beta,\beta} = \frac{\partial H^*(\alpha, \beta)}{\partial \beta} = H_\beta + \frac{1}{n\beta},$$

$$H^*_{\beta,\alpha\beta} = \frac{\partial H^*(\alpha, \beta)}{\partial \alpha\beta} = H_{\alpha\beta},$$

$$H^*_{\beta,\alpha\alpha} = \frac{\partial D^*_1}{\partial \alpha} = H_{\alpha\alpha},$$

$$H^*_{\beta,\alpha\alpha} = \frac{\partial D^*_2}{\partial \beta} = H_{\beta\beta} - \frac{1}{n\beta^2}.$$

Finally, substituting the above expressions in (4.6), we obtain the Bayes estimates of α and β.

4.3. Highest posterior density estimation

The highest posterior density estimation is another popular method used to compute the Bayes estimates. The highest posterior density (HPD) estimate represents the mode of the posterior density. The Bayes estimates using HPD method can be obtained by solving the equations

$$\frac{\partial \pi(\alpha,\beta|\tilde{\mathbf{x}})}{\partial \alpha} = \frac{n+a-1}{\alpha} - b + \frac{\int A_\alpha(x)\mu_{\tilde{x}_i}(x)dx}{\int A(x)\mu_{\tilde{x}_i}(x)dx} = 0, \tag{4.7}$$

$$\frac{\partial \pi(\alpha,\beta|\tilde{\mathbf{x}})}{\partial \beta} = \frac{n+c-1}{\beta} - d + \frac{\int A_\beta(x)\mu_{\tilde{x}_i}(x)dx}{\int A(x)\mu_{\tilde{x}_i}(x)dx} = 0. \tag{4.8}$$

It can be seen that, the solutions of the above two equation cannot be obtained explicitly and, similar to the maximum likelihood method, numerical methods like Newton–Raphson can be used to solve them.

5. SIMULATION EXPERIMENTS

In this section, we conduct Monte-Carlo simulation experiments to show how the various approaches work with different sample sizes. The performance of the proposed approaches was compared on the basis of their expected biases, root mean square error, average of standard errors and of 95% confidence intervals. The true values of the parameters (α, β) are assumed to be $(1.25, 1.5), (1.5, 0.5)$ and $(0.5, 0.75)$, respectively. The sample sizes are chosen as $n = 25, 50$ and 100 to represent small, moderate and large samples, respectively. Each observation from Burr type XII, x_i, was then fuzzified with the corresponding membership function $\mu_{\tilde{x}_i}(x)$, where

$$\mu_{\tilde{x}_i}(x) = \begin{cases} \dfrac{x-(x_i-a_i)}{a_i}, & \text{if } x_i - a_i \le x \le x_i, \\ \dfrac{(x_i+a_i)-x}{a_i}, & \text{if } x_i \le x \le x_i + a_i, \\ 0, & \text{otherwise,} \end{cases} \tag{5.1}$$

and $a_i = 0.05x_i$ (see, for example, Pak and Chatrabgoun ([24]), Pak *et al.* ([27]), Chaturvedi ([6])). That is the observer is unable to provide exact value of observation and an interval of plausible values $[x_i - a_i, x_i + a_i]$ is provided. For example the triangular fuzzy number $(0.1805, 0.1995)$ represents the observed value 0.19 i.e. the interval of plausible values of 0.19 is $[0.1805, 0.1995]$. Then, we compute the MLEs of α and β for the fuzzy sample via Newton–Raphson (NR) and Expectation-Maximization (EM) algorithm. The process is replicated 1000 times. In each replication, we compute the average of biases (Bias), sample standard error (SSE) and the root mean squared error (RMSE) using the expressions

$$\text{Bias}(\theta) = \frac{1}{k}\sum_{i=1}^{k}(\theta_i - \theta_0),$$

$$\text{SSE}(\theta) = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(\theta_i - \bar{\theta})^2}$$

and

$$\text{RMSE}(\theta) = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(\theta_i - \theta_0)^2},$$

where θ represents α or β, θ_0 is the true value of θ, $\bar{\theta}$ is the mean of the estimates of θ and k is the number of replications. Moreover, to compute the estimated standard error (ESE) for the MLEs, we use the observed information matrix given in (3.11). Approximated 95% confidence intervals for the MLE are constructed using the observed information matrix. Moreover, in each iteration, we compute the Bayes estimators using Lindley's approximation, Tierney–Kadane approximation and highest posterior density (HP) methods. At the end, we compute the averages of the absolute biases, sample standard deviation, estimated standard deviation, root mean squared error and 95% confidence intervals. For computing Bayes estimators, we consider gamma priors for α and β with hyperparameters (a, b) and (c, d) , respectively. To make the comparison meaningful, it is assumed that the priors are non-informative $a = b = c = d = 0$ but these priors are improper priors hence we have tried $a = b = c = d = 0.001$ to get proper priors. However, these results are same as those obtained for improper priors. The simulation results of the MLEs and Bayes estimators are reported in Tables 1–2. We have utilized R-4.0.3 software to compute the proposed estimators. The stopping criteria for the algorithms is based on the sum of the absolute differences between two consecutive values of parameters estimates less than 10^{-4}.

From Table 1, we observe that the biases for all estimators, in general, are reasonably small which indicate that the estimated values are close to the true parameter values. As expected, the biases of all estimators become better when the sample size increases. The values of sample standard error (SSE) of the MLEs are approximately close to estimated standard error (ESE) for all the cases and hence the estimated standard error can be used to estimate the standard error of the estimators. In addition, the Bias, SSE, ESE, RMSE and the length of 95% confidence inetrvals of all MLEs are decreasing when sample sizes increasing for all the cases. The estimated coverage probabilities of 95% confidence intervals (CP) are very close to the nominal level for all the cases. Hence, the performance of the MLEs are satisfactory in terms of the biases, standard errors and coverage probabilities of the estimates. Moreover, the Bias of the computed MLEs estimators using EM algorithm for most of the cases are slightly higher than that of the MLEs computed using EM-algorithm. In addition, the central processing time CPU required for NR per iteration is shorter than that of EM algorithm. Figure 1 demonstrates the histograms for the MLEs of α and β when $n = 100$ for the three sets of values. The histograms show approximately normal distribution of the MLEs of α and β.

From Table 2, the biases of the Bayesian estimates of all three methods are also reasonably small. It is clear that the Bias and RMSE are decreasing for increasing values of sample sizes. Moreover, the Bias and RMSE of the Bayes estimates obtained under highest posterior density (HP) are smaller than that of Lindley's method (LN) and Tierney–Kadane approximation (TK). Hence we recommend to use HP method for computing Bayes estimator. From the above results, we conclude that the estimation methods proposed in the article to compute the MLEs and Bayes estimators perform very well.

Table 1: Simulation results for MLEs of α and β.

n			Bias	RMSE	ESE	SSE	95% CI	Length	CP
25	$\alpha = 1.25$	NR	0.068	0.258	0.218	0.249	$(0.95, 1.82)$	0.87	93.6
		EM	0.069	0.255	0.218	0.245	$(0.93, 1.81)$	0.88	93.8
	$\beta = 1.50$	NR	0.072	0.353	0.321	0.345	$(1.05, 2.35)$	1.30	94.1
		EM	0.075	0.350	0.320	0.342	$(1.01, 2.30)$	1.29	94.8
50	$\alpha = 1.25$	NR	0.029	0.162	0.150	0.159	$(1.02, 1.61)$	0.59	94.5
		EM	0.031	0.160	0.152	0.160	$(1.04, 1.60)$	0.56	94.6
	$\beta = 1.50$	NR	0.027	0.226	0.220	0.224	$(1.15, 2.03)$	0.88	95.0
		EM	0.029	0.225	0.218	0.222	$(1.11, 1.98)$	0.87	95.0
100	$\alpha = 1.25$	NR	0.013	0.106	0.104	0.105	$(1.07, 1.48)$	0.41	94.9
		EM	0.015	0.108	0.103	0.105	$(1.08, 1.46)$	0.38	94.8
	$\beta = 1.50$	NR	0.012	0.153	0.154	0.153	$(1.24, 1.85)$	0.61	95.0
		EM	0.012	0.155	0.156	0.155	$(1.24, 1.83)$	0.59	95.1
25	$\alpha = 1.50$	NR	0.155	0.511	0.415	0.484	$(1.01, 2.72)$	1.71	93.8
		EM	0.157	0.515	0.413	0.486	$(1.02, 2.72)$	1.70	93.6
	$\beta = 0.50$	NR	0.007	0.143	0.135	0.146	$(0.29, 0.87)$	0.58	94.8
		EM	0.007	0.144	0.138	0.144	$(0.30, 0.87)$	0.57	94.3
50	$\alpha = 1.50$	NR	0.062	0.290	0.270	0.291	$(1.12, 2.19)$	1.07	94.7
		EM	0.067	0.297	0.269	0.290	$(1.12, 2.19)$	1.07	94.5
	$\beta = 0.50$	NR	0.002	0.098	0.098	0.097	$(0.34, 0.73)$	0.39	95.0
		EM	0.003	0.097	0.097	0.097	$(0.34, 0.73)$	0.39	95.0
100	$\alpha = 1.50$	NR	0.026	0.188	0.182	0.186	$(1.21, 1.93)$	0.72	95.2
		EM	0.027	0.187	0.180	0.185	$(1.20, 1.93)$	0.73	95.4
	$\beta = 0.50$	NR	0.002	0.069	0.069	0.070	$(0.38, 0.66)$	0.28	94.9
		EM	0.001	0.071	0.072	0.071	$(0.38, 0.65)$	0.27	94.9
25	$\alpha = 0.50$	NR	0.080	0.280	0.211	0.267	$(0.51, 1.37)$	0.86	93.4
		EM	0.082	0.270	0.210	0.265	$(0.51, 1.38)$	0.87	93.8
	$\beta = 0.75$	NR	0.006	0.140	0.140	0.143	$(0.30, 0.86)$	0.56	94.4
		EM	0.007	0.145	0.138	0.145	$(0.30, 0.86)$	0.56	94.2
50	$\alpha = 0.50$	NR	0.033	0.149	0.134	0.145	$(0.56, 1.10)$	0.54	94.5
		EM	0.034	0.150	0.132	0.143	$(0.56, 1.10)$	0.56	94.6
	$\beta = 0.75$	NR	0.002	0.096	0.095	0.095	$(0.34, 0.73)$	0.39	95.0
		EM	0.002	0.097	0.097	0.097	$(0.34.0.73)$	0.39	95.2
100	$\alpha = 0.50$	NR	0.013	0.094	0.096	0.093	$(0.60, 0.97)$	0.37	95.2
		EM	0.013	0.092	0.091	0.094	$(0.60, 0.96)$	0.36	95.5
	$\beta = 0.75$	NR	0.003	0.069	0.070	0.069	$(0.38, 0.66)$	0.28	94.9
		EM	0.002	0.070	0.069	0.070	$(0.38, 0.65)$	0.27	94.6

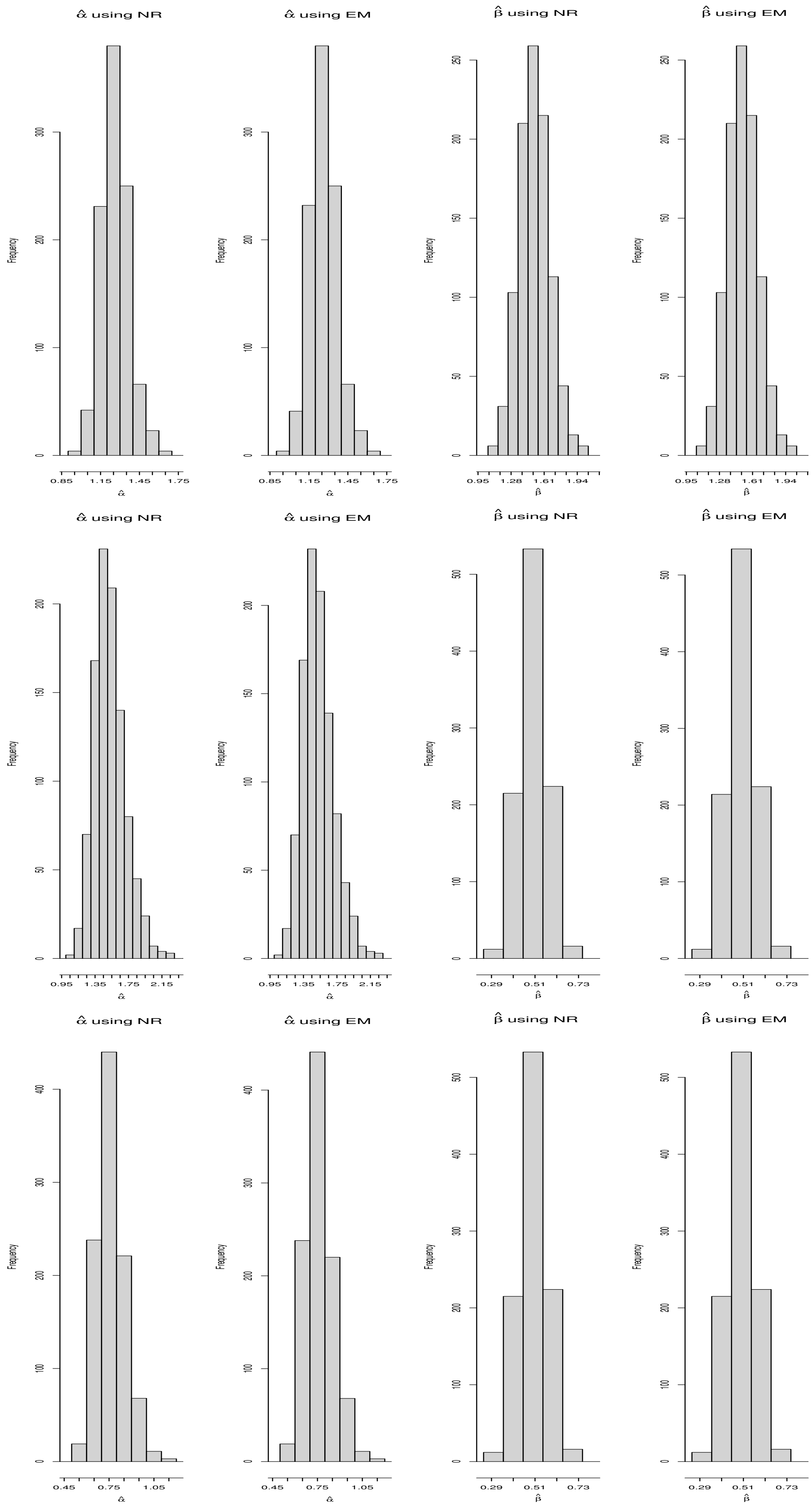

Figure 1: Histograms of the estimated values of the MLEs, $\hat{\alpha}$ and $\hat{\beta}$, for $n = 100$. The first line for ($\alpha = 1.25$, $\beta = 1.5$), the second line for ($\alpha = 1.5$, $\beta = 0.5$) and the third line for ($\alpha = 0.5$, $\beta = 0.75$).

Table 2: Simulation results for Bayesian estimates of α and β.

n		LN	TK	HPD	LN	TK	HPD
		$\alpha = 1.25$			$\beta = 1.5$		
25	Bias	0.068	0.065	0.040	0.068	0.067	0.014
	RMSE	0.259	0.264	0.247	0.348	0.345	0.328
50	Bias	0.029	0.029	0.016	0.025	0.022	0.003
	RMSE	0.162	0.160	0.158	0.224	0.220	0.218
100	Bias	0.013	0.012	0.006	0.011	0.009	−0.001
	RMSE	0.106	0.105	0.105	0.153	0.152	0.151
		$\alpha = 1.5$			$\beta = 0.5$		
25	Bias	0.191	0.199	0.140	0.015	0.015	−0.009
	RMSE	0.529	0.556	0.504	0.141	0.143	0.140
50	Bias	0.084	0.080	0.058	0.006	0.004	−0.006
	RMSE	0.309	0.309	0.295	0.096	0.096	0.095
100	Bias	0.030	0.034	0.022	0.007	0.004	−0.002
	RMSE	0.191	0.190	0.187	0.069	0.071	0.068
		$\alpha = 0.5$			$\beta = 0.75$		
25	Bias	0.103	0.105	0.075	0.014	0.015	−0.009
	RMSE	0.298	0.306	0.276	0.142	0.144	0.141
50	Bias	0.042	0.040	0.029	0.005	0.006	−0.006
	RMSE	0.152	0.154	0.147	0.096	0.092	0.095
100	Bias	0.018	0.017	0.011	0.004	0.002	−0.002
	RMSE	0.095	0.092	0.093	0.069	0.070	0.068

6. APPLICATION EXAMPLES

In this section, we analyze three real data sets to explain how the proposed approaches can be applied in real data analysis. We are assuming that each observation in any of these datasets, x_i, is reported as a fuzzy numbers with membership function given in (5.1). For computing Bayes estimators in this section, we assume gamma priors with hyperparameters $a = b = c = d = 0.001$. This choice of hyperparameters will make the priors proper. However, we have tried to consider different values of hyperparameters, for example, we have considered the cases $a = b = c = d = 1$, and $a = 2$, $b = 1$, $c = 2$, $d = 1$ and the results are not much different than that we have obtained from that case, and are not reported due to the space.

Example 1. The first data set was considered and analyzed by Zimmer *et al.* ([37]) and Lio *et al.* ([19]). The dataset contains the 19 times in minutes to oil breakdown of an insulating fluid under high test voltage (34 kV). The data set is listed as follows: 0.19, 0.78, 0.96, 0.31, 2.78, 3.16, 4.15, 4.67, 4.85, 6.50, 7.35, 8.01, 8.27, 12.06, 31.75, 32.52, 33.91, 36.71, 72.89. Lio *et al.* ([19]) showed that the two-parameters Burr type XII fits the data set very well. The MLEs of (α, β) using Newton–Raphson method are $(1.440, 0.354)$ with standard errors $(0.435, 0.126)$ and 95% confidence intervals $(0.588, 2.292)$ and $(0.106, 0.601)$, respectively, and MLEs using EM algorithm are $(1.436, 0.357)$ with estimated standard error $(0.431, 0.127)$ and 95% confidence intervals (0.590, 2.281) and $(0.108, 0.606)$. In addition, the Bayes estimates of (α, β) are $(1.427, 0.338)$ using Lindley's approximation, $(1.507, 0.364)$ using Tierney–Kadane approximation and $(1.427, 0.338)$ using highest posterior density method.

Example 2. Lawless ([16]) reported the time between failure of air conditioning equipment in a particular type of aircraft. These observations are:

$$0.500,\ 0.875,\ 1.083,\ 1.125,\ 1.208,\ 1.208,\ 2.00,\ 2.375,$$
$$2.458,\ 2.917,\ 3.083,\ 6.375,\ 13.583,\ 16.083,\ 20.917.$$

Kayal *et al.* ([14]) concluded that Burr type XII model fits the data set quite good. The MLEs of (α, β) using Newton–Raphson method are $(3.571, 0.275)$ with standard errors $(1.488, 0.127)$ and 95% confidence intervals $(0.654, 6.487)$ and $(0.026, 0.524)$, respectively, and MLEs using EM algorithm are $(3.500, 0.284)$ with estimated standard error $(1.434, 0.129)$ and 95% confidence intervals $(0.690, 6.311)$ and $(0.031, 0.537)$, respectively. In addition, the Bayes estimates of (α, β) are $(3.519, 0.260)$ using Lindley's approximation, $(3.921, 0.289)$ using Tierney–Kadane approximation and $(3.519, 0.260)$ using highest posterior density method.

Example 3. In this example, we analyze a dataset that represents the survival time of animals observed due to different dosage of poison administered (see Box and Cox ([4])). The observations are listed as:

$$0.18,\ 0.21,\ 0.22,\ 0.22,\ 0.23,\ 0.23,\ 0.23,\ 0.24,\ 0.25,\ 0.29,\ 0.29,\ 0.30,$$
$$0.30,\ 0.31,\ 0.31,\ 0.31,\ 0.33,\ 0.35,\ 0.36,\ 0.36,\ 0.37,\ 0.38,\ 0.38,\ 0.40,$$
$$0.40,\ 0.43,\ 0.43,\ 0.44,\ 0.45,\ 0.45,\ 0.45,\ 0.46,\ 0.49,\ 0.56,\ 0.61,\ 0.62,$$
$$0.63,\ 0.66,\ 0.71,\ 0.71,\ 0.72,\ 0.76,\ 0.82,\ 0.88,\ 0.92,\ 1.02,\ 1.10,\ 1.24.$$

Kayal *et al.* ([14]) analyzed the above data and they concluded that the data might have come from a two-parameter Burr type XII distribution. The MLEs of (α, β) using Newton–Raphson method are $(2.346, 4.938)$ with standard errors $(0.231, 0.822)$ and 95% confidence intervals $(1.893, 2.798)$ and $(1.887, 2.785)$, respectively, and MLEs using EM algorithm are $(2.336, 5.075)$ with estimated standard error $(0.229, 0.850)$ and 95% confidence intervals $(3.326, 6.550)$ and $(3.408, 6.742)$, respectively. In addition, the Bayes estimates of (α, β) are $(2.373,\ 4.928)$ using Lindley's approximation, $(2.338, 4.923)$ using Tierney–Kadane approximation and $(2.304, 4.761)$ using highest posterior density method.

7. CONCLUSION

In this article, we have considered both classical and Bayesian analysis of fuzzy survival time observations when the lifetime of the items follows two-parameter Burr type XII distribution. The MLEs do not have explicit forms. Thus, Newton–Raphson and Expectation-Maximization algorithms have been used to compute the MLEs and both of them work quite well. The Bayes estimates under the squared error loss function also do not exist in explicit form. In this case, we have proposed to use Lindley's approximation, Tierney–Kadane approximation and highest posterior density method to compute the Bayes estimates when the two unknown parameters have independent gamma priors. However, we have considered gamma priors, but a more general prior, namely a prior which has the log-concave p.d.f. may be used, and the method can be easily incorporated in that case. Moreover, in Bayesian estimation, we proposed to use a very well-known symmetric loss function which is the squared-error loss function. However, we may extend the results of the paper by adopting other loss function like LINEX. Another direction for extension is to consider censored fuzzy observations like type II progressively censored fuzzy observations.

A. Proof of Theorem 3.1

Recall that, the log-likelihood function of α and β is given by

$$l(\alpha,\beta|\tilde{\mathbf{x}}) = n\log\alpha + n\log\beta + \sum_{i=1}^{n}\log\int A(x)\mu_{\tilde{x}_i}(x)dx,$$

where

$$A(x) = x^{\alpha-1}(1+x^{\alpha})^{-\beta-1}. \tag{A.1}$$

Observe that, for fixed $\beta > 0$, we have

$$\lim_{\alpha\to 0} l(\alpha,\beta|\tilde{\mathbf{x}}) = \lim_{\alpha\to\infty} l(\alpha,\beta|\tilde{\mathbf{x}}) = -\infty$$

and, for fixed $\alpha > 0$, we have

$$\lim_{\beta\to 0} l(\alpha,\beta|\tilde{\mathbf{x}}) = \lim_{\beta\to\infty} l(\alpha,\beta|\tilde{\mathbf{x}}) = -\infty.$$

We can see that

$$\frac{\partial^2\log(A(x))}{\partial\alpha^2} = -\frac{(\beta+1)(\log(x))^2x^{\alpha}}{(1+x^{\alpha})^2} < 0$$

for fixed $\beta > 0$, i.e., $A(x)$ is strictly log-concave in α for fixed $\beta > 0$. Similarly, we can prove that $A(x)$ is log-concave in β for fixed $\alpha > 0$. By Prekopa–Leindler inequality (see Gardner [10]) we obtain that $\int A(x)\mu_{\tilde{x}_i}(x)dx$ is strictly log-concave in α (or β) for fixed $\beta > 0$ (or $\alpha > 0$). Therefore, for fixed α (or β), $l(\alpha,\beta|\tilde{\mathbf{x}})$ is strictly concave and unimodal function with respect to β (or α). Moreover,

$$\lim_{\substack{\alpha\to 0\\ \beta\to 0}} l(\alpha,\beta|\tilde{\mathbf{x}}) = \lim_{\substack{\alpha\to 0\\ \beta\to\infty}} l(\alpha,\beta|\tilde{\mathbf{x}}) = \lim_{\substack{\alpha\to\infty\\ \beta\to 0}} l(\alpha,\beta|\tilde{\mathbf{x}}) = \lim_{\substack{\alpha\to\infty\\ \beta\to\infty}} l(\alpha,\beta|\tilde{\mathbf{x}}) = -\infty.$$

The rest of the proof is the same as that of Dey *et al.* ([9]).

ACKNOWLEDGMENTS

The authors are grateful to the anonymous referees and the associate editor for making many helpful comments and suggestions on an earlier version of this manuscript which resulted in this improved version.

REFERENCES

[1] AL-Baldawi, T.H.K.; Rasheed, H.A. and Jaseim, S.H. (2015). Using generalized square loss function to estimate the shape parameter of the Burr type XII distribution, *International Journal of Advanced Research*, **3**(5), 393–398.

[2] Basharat, H.; Mustafa, S.; Mahmood, S. and Jun, Y.B. (2019). Inference for the linear combination of two independent exponential random variables based on fuzzy data, *Hacettepe Journal of Mathematics and Statistics*, **48**(6), 1859–1869.

[3] Belaghi, R.A.; Noori Asl, M.N. and Singh, S. (2017). On estimating the parameters of the Burr XII model under progressive type-I interval censoring, *Journal of Statistical Computation and Simulation*, **87**(16), 3132–3151.

[4] Box, G.E. and Cox, D.R. (1964). An analysis of transformations, *Journal of the Royal Statistical Society: Series B (Methodological)*, **26**(2), 211–243.

[5] Burr, I.W. (1942). Cumulative frequency functions, *The Annals of Mathematical Statistics*, **13**(2), 215–232.

[6] Chaturvedi, A.; Singh, S.K. and Singh, U. (2018). Statistical inferences of type II progressively hybrid censored fuzzy data with Rayleigh distribution, *Austrian Journal of Statistics*, **47**(3), 40–62.

[7] Dempster, A.P.; Laird, N.M. and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society: Series B (Methodological)*, **39**(1), 1–22.

[8] Denœux, T. (2011). Maximum likelihood estimation from fuzzy data using the EM algorithm, *Fuzzy Sets and Systems*, **183**(1), 72–91.

[9] Dey, T.; Dey, S. and Kundu, D. (2016). On progressively type-II censored two-parameter Rayleigh distribution, *Communications in Statistics – Simulation and Computation*, **45**(2), 438–455.

[10] Gardner, R. (2002). The Brunn-Minkowski inequality, *Bulletin of the American Mathematical Society*, **39**(3), 355–405.

[11] Ghitany, M. and Al-Awadhi, S. (2002). Maximum likelihood estimation of Burr XII distribution parameters under random censoring, *Journal of Applied Statistics*, **29**(7), 955–965.

[12] Hakim, A.R.; Novita, M. and Fithriani, I. (2019). Using Jeffrey prior information to estimate the shape parameter of Burr distribution, *Journal of Physics: Conference Series, IOP Publishing*, **1218**(1).

[13] Jaheen, Z.F. and Okasha, H.M. (2011). E-Bayesian estimation for the Burr type XII model based on type II censoring, *Applied Mathematical Modelling*, **35**(10), 4730–4737.

[14] Kayal, T.; Tripathi, Y.M.; Rastogi, M.K. and Asgharzadeh, A. (2017). Inference for Burr XII distribution under type-I progressive hybrid censoring, *Communications in Statistics – Simulation and Computation*, **46**(9), 7447–7465.

[15] Khoolenjani, N.B. and Shahsanaie, F. (2016). Estimating the parameter of exponential distribution under type II censoring from fuzzy data, *Journal of Statistical Theory and Applications*, **15**(2), 181–195.

[16] Lawless, J.F. (2011). *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons.

[17] Li, X.; Shi, Y.; Wei, J. and Chai, J. (2007). Empirical Bayes estimators of reliability performances using LINEX loss under progressively type II censored samples, *Mathematics and Computers in Simulation*, **73**(5), 320–326.

[18] Lindley, D.V. (1980). Approximate Bayesian methods, *Trabajos de Estadística y de Investigación Operativa*, **31**(1), 223–245.

[19] Lio, Y.; Tsai, T.R. and Wu, S.J. (2010). Acceptance sampling plans from truncated life tests based on the Burr type XII percentiles, *Journal of the Chinese institute of Industrial Engineers*, **27**(4), 270–280.

[20] Makhdoom, I.; Nasiri, P. and Pak, A. (2016). Estimating the parameter of exponential distribution under type II censoring from fuzzy data, *Journal of Modern Applied Statistical Methods*, **15**(2) 495–509.

[21] Moore, D. and Papadopoulos, A.S. (2000). The Burr type XII distribution as a failure model under various loss functions, *Microelectronics Reliability*, **40**(12), 2117–2122.

[22] Mousa, M.A. and Jaheen, Z. (2002). Statistical inference for the Burr model based on progressively censored data, *Computers & Mathematics with Applications*, **43**(10), 1441–1449.

[23] Pak, A. (2017). Statistical inference for the parameter of Lindley distribution based on fuzzy data, *Brazilian Journal of Probability and Statistics*, **31**(3), 502–515.

[24] Pak, A. and Chatrabgoun, O. (2016). Inference for exponential parameter under progressive type II censoring from imprecise lifetime, *Electronic Journal of Applied Statistical Analysis*, **9**(1), 227–245.

[25] Pak, A.; Khoolenjani, N.B.; Alamatsaz, M.H. and Mahmoudi, M.R. (2020). Bayesian method for the generalized exponential model using fuzzy data, *International Journal of Fuzzy Systems*, **22**(4), 1243-1260.

[26] Pak, A. and Mahmoudi, M.R. (2018). Estimating the parameters of Lomax distribution from imprecise information, *Journal of Statistical Theory and Applications*, **17**(1), 122–135.

[27] Pak, A.; Parham, G.A. and Saraj, M. (2013a). Inference for the Weibull distribution based on fuzzy data, *Revista Colombiana de Estadistica*, **36**(2), 337–356.

[28] Pak, A.; Parham, G.A. and Saraj, M. (2013b). On estimation of Rayleigh scale parameter under doubly type II censoring from imprecise data, *Journal of Data Science*, **11**(2), 305–322.

[29] Panahi, H. and Asadi, S. (2011). Analysis of the type II hybrid censored Burr type XII distribution under LINEX loss function, *Applied Mathematical Sciences*, **5**, 3929–3942.

[30] Pradhan, B. and Kundu, D. (2014). Analysis of interval-censored data with Weibull lifetime distribution, *Sankhya B*, **76**(1), 120–139.

[31] Rao, G.S.; Aslam, M. and Kundu, D. (2015). Burr-XII distribution parametric estimation and estimation of reliability of multicomponent stress-strength, *Communications in Statistics – Theory and Methods*, **44**(23), 4953–4961.

[32] Shafiq, M. (2017). Classical and Baysian inference of Pareto distribution and fuzzy life times, *Pakistan Journal of Statistics*, **33**(1), 15–25.

[33] Tierney, L. and Kadane, J.B. (1986). Accurate approximations for posterior moments and marginal densities, *Journal of the American Statistical Association*, **81**(393), 82–86.

[34] Wahed, A.S. (2006). Bayesian inference using Burr model under asymmetric loss function: an application to carcinoma survival data, *Journal of Statistical Research*, **40**(1), 45–57.

[35] Wang, F.; Keats, J.B. and Zimmer, W.J. (1996). Maximum likelihood estimation of the Burr XII parameters with censored and uncensored data, *Microelectronics Reliability*, **36**(3), 359–362.

[36] Zadeh, L.A. (1968). Probability measures of fuzzy events, *Journal of Mathematical Analysis and Applications*, **23**(2), 421–427.

[37] Zimmer, W.J.; Keats, J.B. and Wang, F. (1998). The Burr XII distribution in reliability analysis, *Journal of Quality Technology*, **30**(4), 386–394.