STATISTICS IN TRANSITION new series, March 2024
Vol. 25, No. 1, pp. 125–143, https://doi.org/10.59170/stattrans-2024-008
Received – 02.11.2022; accepted – 08.01.2024

On Bayesian inference of reliability parameter in Burr-type XII model based on imprecise data: a survey on fuzzy modelling

Iman Makhdoom[1], **Abbas Pak**[2]

Abstract

There are always two major sources of uncertainty in measurements related to lifetime surveys: *variation among the observations* and *imprecision of individual observation* called *fuzziness*. The typical statistical analysis is based on variation among the observations and does not consider the imprecision due to individual observation. However, ignoring the imprecision of individual observations may cause losing information and getting misleading results. It is mandatory to analyse such data, to extend the real numbers classically and Bayesian estimation methods to fuzzy numbers. Inference on the Burr-type (BT) XII model, based on precise measurements, is carried out by researchers, yet the problem of estimating parameters, in the presence of fuzzy data, remains unresolved. We are estimating the BT XII distribution parameters and their corresponding reliability when the available data are in the fuzzy numbers. The maximum likelihood estimation (MLE), the Bayesian method and the method of moments are used for estimating parameters. Finally, these estimators are compared via a Monte-Carlo simulation study.

Key words: Bayesian estimation, Burr-type XII distribution, Maximum likelihood estimates, Markov chain Monte Carlo, EM algorithm, Fuzzy data analysis.

1. Introduction

One of the important application of statistics is to analyse the lifetime data. Various distributions are suggested to model the existing real lifetime data. In this regard, Burr (1942), in his original paper, presented a system of distributions that contains twelve different types of distribution functions useful in lifetime studies, which yield a variety of density shapes. The two-parameter Burr-type (BT) XII model has the cumulative distribution function (CDF) and the probability density function (PDF) as follows:

$$F(x;c,k) = 1-(x^c+1)^{-k}, \quad f(x;c,k) = \frac{kcx^{c-1}}{(1+x^c)^{k+1}}, \quad x>0 \tag{1}$$

respectively, where $c>0$ and $k>0$ are the shape parameters. The corresponding reliability function (RF) and failure rate function (FR) are also given by

$$R(t) = (1+t^c)^{-k}, \qquad \gamma(t) = \frac{ckt^{c-1}}{1+t^c} \quad t>0. \tag{2}$$

[1]Department of Statistics, Payame Noor University(PNU), P.O.Box,19395-4697 Tehran, Iran. E-mail: makhdoom@pnu.ac.ir. ORCID: https://orcid.org/0000-0002-2768-1024.
[2]Department of Computer Sciences, Faculty of Mathematical Sciences, Shahrekord University, P.O.Box 115, Shahrekord, Iran. E-mail: abbas.pak1982@gmail.com. ORCID: https://orcid.org/0000-0003-2388-3523.

respectively. Figure 1 represents the PDFs of BT XII distribution for various quantities of c and k.

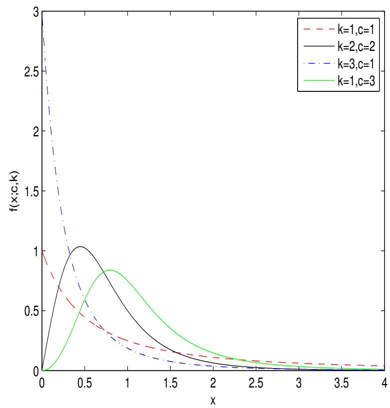

Figure 1: Probability density function for different values of c and k.

This family of distributions, especially types III, X, and XII, have been considered and frequently studied in recent years. Investigations on the BT XII distribution, from point of view its flexibilities, have been carried out by many authors (Hate (1949), Burr (1973), Rodriguez (1977) and Singh, Singh and Kumar (2016)). Wingo (1983) has presented the maximum likelihood methods for fitting of the BT XII model to life test. Wingo (1993) also extended his work and obtained the estimating parameters of this distribution for the progressively censored scheme in life test data. See also Wu and Yu (2005), Xiuchun, Yimin, Jieqiong and Jian (2007), Soliman (2005), Moore and Papadopoulos (2000), Mousa, and Jaheen (2002), for a nice account of it.

The above inference methods, to estimate the parameters of the BT XII distribution, are limited to crisp data. But a matter of concern for the statisticians has been the exact measurement of continuous real variables. For this, numerous methods are considered to measure continuous variables precisely, yet the problem of precise measurement is unresolved, and the numbers solely are approximated.

So, there are always situations that data sometimes cannot be measured and recorded precisely due to machine errors, human error, or unexpected situations and it always remained a problem for the researchers. Note that the problem here is different from censoring and our interest is not the imprecision arising from inspection times, but it is the result of random experiment reported from the observer and its limited perception or recollection of the precise numerical value.

Here, we mention a few examples for fuzzy data. The measurement of the depth of a river because of its water level fluctuation is an imprecise quantity. It may be said that its depth is approximately 40 meters. It cannot be measured precisely at blood pressure, differentiation between a high or low blood pressure. The measurement of temperature is fuzziness. High or low temperature is imprecise quantity due to lack of differentiated between high and low temperature. The effective or ineffective teacher is also one example of fuzziness. Now, we present one case for fuzzy lifetime data as follows. The lifetime of some shafts may be stated as vague values such as: “about 1000 hours",“approximately 1400 hours", “almost between 1000 and 1200 hours", “essentially less than 1200 hours" and so on. Hence, we can conclude that there are always two types of uncertainty in measurements: *variation among the observations* and *imprecision of individual observations* (see Viertl (2011)).

The conventional statistical analysis is only related to variation among the observations and does not consider the vagueness of individual observations. Ignoring this imprecision may be the reason why we lose some information and get false results. The vagueness of such data can be characterized using fuzzy sets that were first introduced by Zadeh (1965). Realizing the importance of fuzziness in recent years, several authors get deep concentration on the fuzzy sets to estimation theory; but still, in most of the publications fuzziness is ignored. Gertner and Zhu (1996) considered Bayesian approximation in the forest studies when samples or prior knowledges are fuzzy. Wu (2004) obtained the Bayesian estimates on lifetime data for fuzzy environments. Gil, López-Diaz and Ralescu (2006) indicated a backward analysis on interpretation, modelling, and impact of the meaning of the fuzzy random variable. Huang, Zuo and Sun (2006) proposed a new method to determine the membership function of the estimates of the parameters and the reliability function of multiparameter lifetime distributions. Coppi, Gilb and Kiersc (2006) presented some applications of fuzzy techniques in statistical analysis. Viertl (2006) discussed a generalization of classical statistical inference methods for univariate fuzzy numbers. Akbari and Rezaei (2007) proposed a new method for uniformly minimum variance unbiased fuzzy point estimation. Zarei, Amini, Taheri and Rezaei (2012) considered the Bayesian estimation of failure rate and the mean time to failure based on vague set theory in the case of complete and censored data sets. Pak, Parham and Saraj (2013, 2014) carried out a series of studies to develop the inferential procedures for the lifetime distributions based on vague information and Shafiq and Atif (2015) obtained the survival models that deal with imprecise lifetime measurements. Very recently, Pak (2016) has investigated some inferences for the Lindley distribution based on fuzzy data.

To our knowledge, there exist no interpretation on estimating the parameters of BT XII distribution from fuzzy data. Since the classical statistical estimation procedures are not suitable for the fuzzy sets; we have to extend the conventional methods to estimate the parameters of BT XII distribution in the new situations. Therefore our main object is to develop the inferential procedures for BT XII parameters when the available data are fuzzy numbers.

The rest of the paper is set up as follows. In Section 2, we consider a review on the original understandings and basic definition of fuzzy set theory. A generalized likelihood function based on fuzzy data is introduced in Section 3. We also present the common method of maximum likelihood for estimating the parameters c and k by taking advantage of the Newton-Raphson (NR) and Expectation Maximization (EM) algorithms in this section. In Section 4, we carry out the estimating parameters of c and k using the moment method. In Section 5, we apply a Bayesian approach for estimating of the unknown parameters using the approximation forms of Tirney and Kadane (1986) and Markov Chain Monte

Carlo (MCMC) technique. Extensive numerical experiments are performed to compare the accuracy of the various proposed methods in Section 6. Finally, Section 7 concludes this research.

2. Preliminary concepts of fuzzy sets theory

Let us first review the fundamental notation and basic definitions of fuzzy set theory used in the paper. In the following, we explaine some special concepts of fuzzy sets theory from Viertl, (2011) and Zadeh (1965, 1968).

Notice an experiment determined by a probability space $X = (X, \mathscr{B}_X, P_\theta)$, where a measurable space is as $(X, \mathscr{B}_X)$ and P_θ belongs to a certain family of probability measures $\{P_\theta, \theta \in \Theta\}$ on $(X, \mathscr{B}_X)$. Consider that the observer cannot distinguish or transmit with exactness the outcome in the performance of 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, determined 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 X.

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

Definition 2 (see Tanaka, Okuda and Asai (1979)). *A fuzzy information system* (f.i.s.) $\tilde{X}$ associated with the experiment X is a fuzzy partition $\{\tilde{x}_1, ..., \tilde{x}_k\}$, i.e., a set of fuzzy events on X satisfying the orthogonality condition $\sum_{i=1}^{k} \mu_{\tilde{x}_i}(x) = 1 \quad$ for all $x \in X$.

On the other hand, according to Zadeh (1968) given the experiment $X = (X, \mathscr{B}_X, P_\theta)$, $\theta \in \Theta$, and a f.i.s. $\tilde{\boldsymbol{X}}$ associated with it, each probability measure P_θ on $(X, \mathscr{B}_X)$ induces a probability measure on $\tilde{\boldsymbol{X}}$ defined as follows:

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

$$\mathrm{P}(\tilde{x}) = \int_X \mu_{\tilde{x}}(x) dP_\theta(x), \qquad \tilde{x} \in \tilde{X}. \tag{3}$$

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

$$g(\mathbf{y}|\tilde{A}) = \frac{\mu_{\tilde{A}}(\mathbf{y}) g(\mathbf{y})}{\int \mu_{\tilde{A}}(\mathbf{u}) g(\mathbf{u}) d\mathbf{u}}. \tag{4}$$

Definition 4 (see Shafiq and Viertl (2014)): A subset $\tilde{x}$ of the set of real numbers (denoted by $\mathbb{R}$) is named *fuzzy number* and is characterized by the so-called membership function $\mu_{\tilde{x}}(.)$. A fuzzy number must fulfill $\mu_{\tilde{x}} : \mathbb{R} \longrightarrow [0,1]$ is Borel-measurable; $\exists x_0 \in \mathbb{R} : \mu_{\tilde{x}}(x_0) = 1$; and the so-called $\lambda-$cuts $(0 < \lambda \leq 1)$, defined as $B_\lambda(\tilde{x}) = \{x \in \mathbb{R} : \mu_{\tilde{x}}(x) \geq \lambda\}$, are all closed interval, i.e., $B_\lambda(\tilde{x}) = [a_\lambda, b_\lambda]$.

The conventional membership functions for analysing of fuzzy lifetime data are called as *triangular and trapezoidal* fuzzy numbers. A triangular fuzzy number is described as $\tilde{x} = (\nu, \omega, \delta)$ and the trapezoidal fuzzy number can also be characterized as $\tilde{x} = (\delta, \nu, \omega, \eta)$

with the corresponding membership functions

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

respectively. For a detailed study on the fuzzy sets, membership functions and triangular and trapezoidal fuzzy numbers one can refer to Singpurwalla and Booker (2004) and Pak, Parham and Saraj (2013).

3. Maximum likelihood estimation

Let $X_1,...,X_n$ be a random sample of size n from the BT XII distribution with PDF given by (1). Let $\mathbf{X}=(X_1,...,X_n)$ denotes the corresponding random vector. If a realization $\mathbf{x}=(x_1,...,x_n)$ of $\mathbf{X}$ is known exactly, we can obtain the complete data likelihood function as

$$L(c,k;\mathbf{x})=(kc)^n\prod_{i=1}^{n}\frac{x_i^{c-1}}{(1+x_i^c)^{k+1}} \tag{5}$$

Now, consider the problem where the results of an experimental performance are not recorded or measured precisely, but that rather the available data are identified by means 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 $\mathbf{x}$. Once $\tilde{\mathbf{x}}$ is given, and assuming the joint membership function $\mu_{\tilde{\mathbf{x}}}(\mathbf{x})$ to be decomposable as $\mu_{\tilde{\mathbf{x}}}(\mathbf{x})=\mu_{\tilde{x}_1}(x_1)\times...\times\mu_{\tilde{x}_n}(x_n)$, its probability can be computed based on Zadeh's definition (see Zadeh (1968)) of the probability of a fuzzy event. By using the expression (3), the observed-data likelihood function based on the fuzzy sample $\tilde{\mathbf{x}}$ can then be obtained as

$$Lo(c,k;\tilde{\mathbf{x}})=P(\tilde{\mathbf{x}};c,k)=\int f(\mathbf{x};c,k)\mu_{\tilde{\mathbf{x}}}(\mathbf{x})d\mathbf{x}. \tag{6}$$

Since the data vector $\mathbf{x}$ is a realization of an independent identically distributed (i.i.d.) random vector $\mathbf{X}$, the likelihood function (6) can be written as:

$$Lo^*(c,k;\tilde{\mathbf{x}})=(kc)^n\prod_{i=1}^{n}\int\frac{x^{c-1}}{(1+x^c)^{k+1}}\mu_{\tilde{x}_i}(x)dx \tag{7}$$

Then, the observed data log-likelihood function is as follows:

$$L^{**}(c,k,\tilde{\mathbf{x}})=n\ln(kc)+\sum_{i=1}^{n}\ln\left(\int\frac{x^{c-1}}{(1+x^c)^{k+1}}\mu_{\tilde{x}_i}(x)dx\right). \tag{8}$$

The maximum likelihood estimate (MLE) of parameters c and k can be obtained by maximizing the log-likelihood $L^{**}(c,k,\tilde{x})$. Equating the partial derivatives of the log-likelihood

(8) with respect to c and k to zero, the resulting two equations are:

$$\frac{\partial}{\partial k}L^{**}(c,k,\tilde{x}) = \frac{n}{k} - \sum_{i=1}^{n} \frac{\int x^{c-1}(1+x^c)^{-k-1}\ln(1+x^c)\mu_{\tilde{x}_i(x)}dx}{\int x^{c-1}(1+x^c)^{-k-1}\mu_{\tilde{x}_i(x)}dx} = 0 \tag{9}$$

and

$$\frac{\partial}{\partial c}L^{**}(c,k,\tilde{x}) = \frac{n}{c} + \sum_{i=1}^{n} \frac{\int x^{c-1}(1+x^c)^{-k-1}(1-(k+1)(1+x^c)^{-1}x^c)\ln(x)\mu_{\tilde{x}_i(x)}dx}{\int x^{c-1}(1+x^c)^{-k-1}\mu_{\tilde{x}_i(x)}dx} = 0. \tag{10}$$

There is no closed form solution for the likelihood equation, therefore an iterative numerical search is used to obtain the MLEs. In next section the Newton-Raphson method and EM algorithm are used to obtain the MLE of the c and k parameters.

3.1. Newton-Raphson Algorithm

In this method, the solution of the likelihood equations is obtained through an iterative procedure. Let $\theta = (k,c)^T$ be the parameter vector. Then, at $(h+1)$th step of iteration process, the updated parameter is computed as

$$\theta^{(h+1)} = \theta^{(h)} - \left[\frac{\partial^2 L^{**}(\theta;\tilde{\mathbf{x}})}{\partial\theta\partial\theta^T}|_{\theta=\theta^{(h)}}\right]^{-1}\left[\frac{\partial L^{**}(\theta;\tilde{\mathbf{x}})}{\partial\theta}|_{\theta=\theta^{(h)}}\right] \tag{11}$$

in which

$$\frac{\partial L^{**}(\theta;\tilde{\mathbf{x}})}{\partial\theta} = \begin{pmatrix} \frac{\partial L^{**}(c,k;\tilde{\mathbf{x}})}{\partial c} \\ \frac{\partial L^{**}(c,k;\tilde{\mathbf{x}})}{\partial k} \end{pmatrix}, \quad \frac{\partial^2 L^{**}(\theta;\tilde{\mathbf{x}})}{\partial\theta\partial\theta^T} = \begin{pmatrix} \frac{\partial L^{**}(c,k;\tilde{\mathbf{x}})}{\partial c^2} & \frac{\partial L^{**}(c,k;\tilde{\mathbf{x}})}{\partial c\partial k} \\ \frac{\partial L^{**}(c,k;\tilde{\mathbf{x}})}{\partial k\partial c} & \frac{\partial L^{**}(c,k;\tilde{\mathbf{x}})}{\partial k^2} \end{pmatrix}.$$

For proceeding with the NR method, we need the second-order derivatives of the log-likelihood with respect to the parameters that are obtained as follows.

$$\begin{aligned} \frac{\partial^2}{\partial c^2}L^{**}(c,k,\tilde{\mathbf{x}}) &= \frac{-n}{c^2} \\ &+ \sum_{i=1}^{n} \frac{\left\{\int x^{c-1}(1+x^c)^{-k-1}mdx - \int(k+1)(1+x^c)^{-k-2}x^{2c-1}mdx\right\}B}{B^2} \\ &+ \sum_{i=1}^{n} \frac{\left\{\int(k+1)(1+x^c)^{-k-3}x^{2c-1}\ln^2 x\mu_{\tilde{x}_i}(x)dx - \int x^{2c-1}(k+1)(1+x^c)^{-k-2}\ln^2 x\mu_{\tilde{x}_i}(x)dx\right\}B}{B^2} \\ &- \sum_{i=1}^{n} \frac{\left\{\int x^{c-1}(1+x^c)^{-k-1}\ln x\mu_{\tilde{x}_i}(x)dx - \int(k+1)(1+x^c)^{-k-2}x^{2c-1}\ln x\mu_{\tilde{x}_i}(x)dx\right\}A}{B^2}, \end{aligned} \tag{12}$$

$$\begin{aligned} \frac{\partial^2}{\partial k^2}L^{**}(c,k,\tilde{\mathbf{x}}) &= \frac{-n}{k^2} \\ &+ \sum_{i=1}^{n} \frac{\left(\int x^{c-1}(1+x^c)^{-k-1}\ln^2(1+x^c)\mu_{\tilde{x}_i(x)}dx\right)B}{B^2} - \sum_{i=1}^{n} \frac{\left(\int x^{c-1}(1+x^c)^{-k-1}\ln(1+x^c)\mu_{\tilde{x}_i(x)}dx\right)^2}{B^2}, \end{aligned} \tag{13}$$

$$\begin{aligned} \frac{\partial^2}{\partial c\partial k}L^{**}(c,k,\tilde{\mathbf{x}}) &= \\ &\sum_{i=1}^{n} \frac{\left(-\int x^{c-1}(1+x^c)^{-k-1}(1-(k+1)(1+x^c)^{-1}x^c)\ln(1+x^c)\ln x\mu_{\tilde{x}_i}(x)dx - \int(1+x^c)^{-k-2}x^{2c-1}\ln x\mu_{\tilde{x}_i}(x)dx\right)B}{B^2} \\ &+ \sum_{i=1}^{n} \frac{\left(\int x^{c-1}\ln(1+x^c)(1+x^c)^{-k-1}\mu_{\tilde{x}_i}(x)dx\right)\left(\int x^{c-1}(1+x^c)^{-k-1}(1-(k+1)(1+x^c)^{-1}x^c)\ln(x)\mu_{\tilde{x}_i(x)}dx\right)}{B^2}, \end{aligned} \tag{14}$$

in which, $m = (1-(k+1)(1+x^c)^{-1}x^c)\ln^2(x)\mu_{\tilde{x}_i}(x)$, $A = \int x^{c-1}(1+x^c)^{-k-1}(1-(k+1)(1+x^c)^{-1}x^c)\ln(x)\mu_{\tilde{x}_i(x)}dx$, $B = \int x^{c-1}(1+x^c)^{-k-1}\mu_{\tilde{x}_i(x)}dx$. The iteration process continues until convergence, i.e., until $\|\theta^{(h+1)} - \theta^{(h)}\| < \varepsilon$ for some pre-fixed $\varepsilon > 0$.

Note that the second-order derivatives of the log-likelihood are needed at every iteration in this method. The calculation of the derivatives based on fuzzy data can be some tedious in most of the time. To solve this problem, an EM algorithm will be present in the following section.

3.2. EM Algorithm

The EM algorithm is a convenient method for incomplete data problems. Since the observed fuzzy data $\tilde{\mathbf{x}}$ can be considered an incomplete data vector $\mathbf{x}$, therefore an EM algorithm is used to obtain the MLE of the unknown parameters c and k, (see Denoeux (2011)).
From the Eq. (5), the log-likelihood function for the complete data vector $\mathbf{x}$ is given by,

$$\ln L(c,k,\mathbf{x}) = n\ln k + n\ln c + (c-1)\sum_{i=1}^{n}\ln x_i - (k+1)\sum_{i=1}^{n}\ln(1+x_i^c) \tag{15}$$

Taking the derivative with respect to c and k, respectively, on (15), the following likelihood equations are obtained:

$$\frac{n}{c} = (k+1)\sum_{i=1}^{n}\frac{x_i^c\ln x_i}{1+x_i^c} - \sum_{i=1}^{n}\ln x_i, \qquad \frac{n}{k} = \sum_{i=1}^{n}\ln(1+x_i^c). \tag{16}$$

So, the EM algorithm is given by the following iterative process:

- Given starting values of c and k say $c^{(0)}$ and $k^{(0)}$ and set $h = 0$. In the $(h+1)$-th iteration, the E-step requires to compute the following conditional expectations using the expression (4):

$$E_{1i} = E_{c^{(h)},k^{(h)}}(\ln X|\tilde{x}_i) = \frac{\int x^{c^{(h)}-1}(\ln x)(1+x^{c^{(h)}})^{-k^{(h)}-1}\mu_{\tilde{x}_i}(x)dx}{\int x^{c^{(h)}-1}(1+x^{c^{(h)}})^{-k^{(h)}-1}\mu_{\tilde{x}_i}(x)dx} \tag{17}$$

$$E_{2i} = E_{c^{(h)},k^{(h)}}(\frac{X^c\ln X}{1+X^c}|\tilde{x}_i) = \frac{\int x^{2c^{(h)}-1}(\ln x)(1+x^{c^{(h)}})^{-k^{(h)}-2}\mu_{\tilde{x}_i}(x)dx}{\int x^{c^{(h)}-1}(1+x^{c^{(h)}})^{-k^{(h)}-1}\mu_{\tilde{x}_i}(x)dx} \tag{18}$$

$$E_{3i} = E_{c^{(h)},k^{(h)}}(\ln(1+X^c)|\tilde{x}_i) = \frac{\int x^{c^{(h)}-1}(\ln(1+x^{c^{(h)}})(1+x^{c^{(h)}})^{-k^{(h)}-1}\mu_{\tilde{x}_i}(x)dx}{\int x^{c^{(h)}-1}(1+x^{c^{(h)}})^{-k^{(h)}-1}\mu_{\tilde{x}_i}(x)dx}. \tag{19}$$

  The likelihood equations (16) are replaced by

$$\frac{n}{c} = (k+1)\sum_{i=1}^{n}E_{2i} - \sum_{i=1}^{n}E_{1i}, \qquad \frac{n}{k} = \sum_{i=1}^{n}E_{3i}. \tag{20}$$

  The M-step requires to solve the Eqs. in (20), and obtain the next values, $c^{(h+1)}$ and

$k^{(h+1)}$, of c and k, respectively, as follows:

$$c^{(h+1)} = \frac{n}{(k^{(h+1)}+1)\sum_{i=1}^{n} E_{2i} - \sum_{i=1}^{n} E_{1i}}, \qquad k^{(h+1)} = \frac{n}{\sum_{i=1}^{n} E_{3i}}. \tag{21}$$

- Checking convergence, if the convergence occurs then the current $c^{(h+1)}$ and $k^{(h+1)}$ are the maximum likelihood estimates of c and k using the EM algorithm; otherwise, set $h = h+1$ and go to previous step.

The maximum likelihood estimate of (c,k) by applying the EM algorithm is thereafter refereed to as "$(\hat{c}_{EM}, \hat{k}_{EM})$" in this article.

4. Method of moment

The rth moment of the BT XII distribution (see Rodriguez (1977)) is given by

$$E(X^r) = kB\left(\frac{r}{c}+1, k-\frac{r}{c}\right) = \frac{\Gamma(\frac{r}{c}+1)\Gamma(k-\frac{r}{c})}{\Gamma(k)} \tag{22}$$

in which, $B(.)$ and $\Gamma(.)$ are the beta and the complete gamma functions, respectively.

By equating the first and the second sample moments to the corresponding population moments, to obtain the estimates of moments approach, the following equations are used:

$$\frac{\Gamma(\frac{1}{c}+1)\Gamma(k-\frac{1}{c})}{\Gamma(k)} = \frac{1}{n}\sum_{i=1}^{n} E_{c,k}(X|\tilde{x}_i), \qquad \frac{\Gamma(\frac{2}{c}+1)\Gamma(k-\frac{2}{c})}{\Gamma(k)} = \frac{1}{n}\sum_{i=1}^{n} E_{c,k}(X^2|\tilde{x}_i). \tag{23}$$

Since the closed form of the solutions to Eqs. in (23) could not be obtained, to achieve the parameter estimates we use the following iterative numerical process. Let the initial estimates of c and k, say $c^{(0)}$ and $k^{(0)}$ with $h = 0$. In the $(h+1)$ th iteration, we first compute

$$E_{c^{(h)},k^{(h)}}(X^r|\tilde{x}_i) = \frac{\int x^{c^{(h)}+r-1}(1+x^{c^{(h)}})^{-k^{(h)}-1}\mu_{\tilde{x}_i}(x)dx}{\int x^{c^{(h)}-1}(1+x^{c^{(h)}})^{-k^{(h)}-1}\mu_{\tilde{x}_i}(x)dx}, \quad r = 1,2.$$

We have to solve the system of two equations and two unknowns in the Eqs. (23). These equations are the complex nonlinear equations. Consequently, we may need to use an iterative numerical method to handle the finding of the roots c and k of equations in (23).

5. Bayesian approach

A robust and valid alternative to traditional statistical perspectives has been called Bayesian inference in recent decades. It received frequent attention for statistical inference. In this section the Bayesian estimates under the assumptions that c and k have independent gamma priors are obtained with the pdfs respectively,

$$\pi_1(c) = \frac{b_1^{a_1}}{\Gamma(a_1)} c^{a_1-1} \exp(-cb_1), \quad c > 0 \tag{24}$$

and

$$\pi_2(k) = \frac{b_2^{a_2}}{\Gamma(a_2)} k^{a_2-1} \exp(-kb_2), \quad k > 0 \tag{25}$$

with the parameters $c \sim Gamma(a_1, b_1)$ and $k \sim Gamma(a_2, b_2)$ (see Singh, Singh and Kumar (2016)). Due to the likelihood function in Eq. (7) and the prior distributions mentioned in (24) and (25) the complete form of the posterior function is as follows:

$$\pi(\theta|\tilde{\mathbf{x}}) \propto L^{**}(\theta, \tilde{\mathbf{x}})(c^{a_1-1} \exp(-cb_1))(k^{a_2-1} \exp(-kb_2)). \tag{26}$$

Hence, the joint posterior density function of c and k given the data can be written as follows:

$$\pi(c,k|\tilde{\mathbf{x}}) = \frac{\pi_1(c)\pi_2(k)Lo^*(c,k;\tilde{\mathbf{x}})}{\int_0^\infty \int_0^\infty \pi_1(c)\pi_2(k)Lo^*(c,k;\tilde{\mathbf{x}})dcdk}. \tag{27}$$

Therefore, the Bayes estimate of any function of c and k, say $g(c,k)$, under a squared error loss function is

$$E(g(c,k)|\tilde{\mathbf{x}}) = \frac{\int_0^\infty \int_0^\infty g(c,k)\pi_1(c)\pi_2(k)Lo^*(c,k;\tilde{\mathbf{x}})dcdk}{\int_0^\infty \int_0^\infty \pi_1(c)\pi_2(k)Lo^*(c,k;\tilde{\mathbf{x}})dcdk}. \tag{28}$$

But, we cannot evaluate these estimates explicitly. Hence, we suggest Tierney and Kadane's procedure and MCMC method to approximate them.

5.1. Tierney and Kadane's method

The Eq. (28) can be re-written as follows:

$$E(g(c,k)|\tilde{\mathbf{x}}) = \frac{\int_0^\infty \int_0^\infty g(c,k)e^{Q(c,k)}dcdk}{\int_0^\infty \int_0^\infty e^{Q(c,k)}dcdk} \tag{29}$$

in which, $Q(c,k) = \ln[\pi_1(c)\pi_2(k)] + \ln Lo^*(c,k;\tilde{\mathbf{x}}) \equiv \rho(c,k) + L^{**}(c,k)$. Note that Eq. (29) cannot be obtained analytically. Using this approximation can be useful to solve this issue.

Setting $H(c,k) = \frac{Q(c,k)}{n}$ and $H^*(c,k) = \frac{[\ln g(c,k)+Q(c,k)]}{n}$, the expression in (29) can be reexpressed as

$$E(g(c,k)|\tilde{\mathbf{x}}) = \frac{\int_0^\infty \int_0^\infty e^{nH^*(c,k)}dcdk}{\int_0^\infty \int_0^\infty e^{nH(c,k)}dcdk}. \tag{30}$$

Following the Tierney & Kadane method, which is based on Laplace's method (Tierney and Kadane (1986)), Eq. (30) can be computed as follows:

$$\hat{g}_{Bayes}(c,k) = \left[\frac{det\sum^*}{det\sum}\right]^{\frac{1}{2}} \exp\left\{n\left[H^*(\overline{c}^*,\overline{k}^*) - H(\overline{c},\overline{k})\right]\right\} \tag{31}$$

in which$(\overline{c}^*, \overline{k}^*)$ and $(\overline{c}, \overline{k})$ maximize $H^*(c,k)$ and $H(c,k)$, respectively. Also, $\sum^*$ and $\sum$ are the negatives of the inverse Hessians of $H^*(c,k)$ and $H(c,k)$ at $(\overline{c}^*, \overline{k}^*)$ and $(\overline{c}, \overline{k})$, respectively.

For our case, we have

$$H(c,k) = \frac{1}{n}\{K^* + (a_1 - 1 + n)\ln c + (a_2 - 1 + n)\ln k - b_1 c - b_2 k\} \\ + \frac{1}{n}\left\{\sum_{i=1}^{n} \ln\left(\int \frac{x^{c-1}}{(1+x^c)^{k+1}} \mu_{\tilde{x}_i}(x)dx\right)\right\} \tag{32}$$

in which K^* is a constant. So, we can obtain $(\overline{c}, \overline{k})$ by solving the following two equations:

$$\frac{\partial}{\partial c}H(c,k) = \frac{1}{n}\left\{\frac{a_1 - 1 + n}{c} - b_1\right\} \\ + \frac{1}{n}\left\{\sum_{i=1}^{n} \frac{\int x^{c-1}(1+x^c)^{-k-1}(1-(k+1)(1+x^c)^{-1}x^c)\ln(x)\mu_{\tilde{x}_i(x)}dx}{\int x^{c-1}(1+x^c)^{-k-1}\mu_{\tilde{x}_i(x)}dx}\right\} \tag{33}$$

$$\frac{\partial}{\partial k}H(c,k) = \frac{1}{n}\left\{\frac{a_2 - 1 + n}{k} - b_2\right\} \\ - \frac{1}{n}\left\{\sum_{i=1}^{n} \frac{\int x^{c-1}(1+x^c)^{-k-1}\ln(1+x^c)\mu_{\tilde{x}_i(x)}dx}{\int x^{c-1}(1+x^c)^{-k-1}\mu_{\tilde{x}_i(x)}dx}\right\}. \tag{34}$$

The determinant of the negative of the inverse Hessian of $H(c,k)$ at $(\overline{c}, \overline{k})$ is as follows:

$$det\sum = \left(H_{11}H_{22} - H_{12}^2\right)^{-1}, \tag{35}$$

in which,

$$H_{11} = \tfrac{1}{n}\tfrac{-(a_1-1+n)}{\overline{c}^2} \\ + \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \frac{\left\{\int x^{\overline{c}-1}(1+x^{\overline{c}})^{-\overline{k}-1}\overline{m}dx - \int(\overline{k}+1)(1+x^{\overline{c}})^{-\overline{k}-2}x^{2\overline{c}-1}\overline{m}dx\right\}\overline{B}}{\overline{B}^2} \\ + \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \frac{\left\{\int(\overline{k}+1)(1+x^{\overline{c}})^{-\overline{k}-3}x^{2\overline{c}-1}\ln^2 x\mu_{\tilde{x}_i}(x)dx - \int x^{2\overline{c}-1}(\overline{k}+1)(1+x^{\overline{c}})^{-\overline{k}-2}\ln^2 x\mu_{\tilde{x}_i}(x)dx\right\}\overline{B}}{\overline{B}^2} \\ - \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \frac{\left\{\int x^{\overline{c}-1}(1+x^{\overline{c}})^{-\overline{k}-1}\ln x\mu_{\tilde{x}_i}(x)dx - \int(k+1)(1+x^{\overline{c}})^{-\overline{k}-2}x^{2\overline{c}-1}\ln x\mu_{\tilde{x}_i}(x)dx\right\}\overline{A}}{\overline{B}^2}, \tag{36}$$

$$H_{22} = \tfrac{1}{n}\tfrac{-(a_2-1+n)}{\overline{k}^2} \\ - \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \frac{\left(\int x^{\overline{c}-1}(1+x^{\overline{c}})^{-\overline{k}-1}\ln^2(1+x^{\overline{c}})\mu_{\tilde{x}_i(x)}dx\right)B}{B^2} + \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \frac{\left(\int x^{\overline{c}-1}(1+x^{\overline{c}})^{-\overline{k}-1}\ln(1+x^{\overline{c}})\mu_{\tilde{x}_i(x)}dx\right)^2}{\overline{B}^2}, \tag{37}$$

$$H_{12} = \\ - \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \frac{\left(\int x^{\overline{c}-1}(1+x^{\overline{c}})^{-\overline{k}-1}(1-(\overline{k}+1)(1+x^{\overline{c}})^{-1}x^{\overline{c}})\ln(1+x^{\overline{c}})\ln x\mu_{\tilde{x}_i}(x)dx + \int(1+x^{\overline{c}})^{-\overline{k}-2}x^{2\overline{c}-1}\ln x\mu_{\tilde{x}_i}(x)dx\right)\overline{B}}{\overline{B}^2} \\ + \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \frac{\left(\int x^{\overline{c}-1}\ln(1+x^{\overline{c}})(1+x^{\overline{c}})^{-\overline{k}-1}\mu_{\tilde{x}_i}(x)dx\right)\left(\int x^{\overline{c}-1}(1+x^{\overline{c}})^{-\overline{k}-1}(1-(\overline{k}+1)(1+x^{\overline{c}})^{-1}x^{\overline{c}})\ln(x)\mu_{\tilde{x}_i(x)}dx\right)}{\overline{B}^2}, \tag{38}$$

in which $\overline{m} = (1-(\overline{k}+1)(1+x^{\overline{c}})^{-1}x^{\overline{c}})\ln^2(x)\mu_{\tilde{x}_i}(x)$, $\overline{A} = \int x^{\overline{c}-1}(1+x^{\overline{c}})^{-\overline{k}-1}(1-(\overline{k}+1)(1+x^{\overline{c}})^{-1}x^{\overline{c}})\ln(x)\mu_{\tilde{x}_i(x)}dx$, $\overline{B} = \int x^{\overline{c}-1}(1+x^{\overline{c}})^{-\overline{k}-1}\mu_{\tilde{x}_i(x)}dx$. Now, following the same arguments with $g(c,k) = c$ and k, respectively, in $H^*(c,k)$, $\hat{c}_{Bayes}$ and $\hat{k}_{Bayes}$ in Eq. (31) can then be obtained in a straightforward manner.

5.2. MCMC Method

MCMC methods use the computer simulation procedure to get a Markov sequence with ergodic properties in such a way that they have a limiting distribution. We know if the loss function is squared error, then the Bayes estimates of the parameters $\theta = (c,k)$ are their respective posteriors mean. But due to the complexity of the extraction of samples from the posterior function, we have to apply the well-known "**Gibbs sampling** " technique. Gibbs sampling defines a broad class of MCMC methods that is used in Bayesian analysis. It is also a special example of a general approach referred to as Metropolis-Hasting (MH) algorithm (Hanagl and Ahmadi (2009)).Thus, a multivariate version of MH algorithm is Gibbs sampling. In the following, the Gibbs sampler method is appropriated to compute the Bayes estimates numerically.

The posterior PDFs of c and k are given by,

$$\pi_1^*(c|k,\tilde{\mathbf{x}}) \propto c^{a_1-1+n}\exp(-cb_1)\prod_{i=1}^{n}\int \frac{x^{c-1}}{(1+x^c)^{k+1}}\mu_{\tilde{x}_i}(x)dx \tag{39}$$

and

$$\pi_2^*(k|c,\tilde{\mathbf{x}}) \propto k^{a_2-1+n}\exp(-kb_2)\prod_{i=1}^{n}\int \frac{x^{c-1}}{(1+x^c)^{k+1}}\mu_{\tilde{x}_i}(x)dx \tag{40}$$

Note that the posterior PDFs of c and k in (39) and (40) respectively, are unknown. Hence, we use the MH method to generate a random sample from these distributions. We use the normal distribution as the proposal distribution for the method. So, the Gibbs sampling algorithm is as follows:

Step 1: Start with an initial value $(c^{(0)}, k^{(0)})$ and fix $t = 1$.

Step 2: Generate $c^{(0)}$ from (39) by using of the MH with the $N(c^{(t-1)}, 1)$ proposal distribution and generate $k^{(0)}$ from (40) by using of the MH with the $N(k^{(t-1)}, 1)$ proposal distribution. Fix $t = t+1$.

Step 3: Repeat Step 2, T times.

For running the algorithm above, we need to perform the MH algorithm in Step 2. These algorithms are given by,

I : Fix t=1.

II : Let $v_1 = c_i^{(t-1)}$ and $v_2 = k_i^{(t-1)}$. Generate w_1 and w_2 from the proposal distributions $q \sim N(c_i^{(t-1)}, 1)$ and $q \sim N(k_i^{(t-1)}, 1)$, respectively. Let $p_1^*(v_1, w_1) = \min\left\{1, \frac{\pi_1^*(w_1|\tilde{x})q(v_1)}{\pi_1^*(v_1|\tilde{x})q(w_1)}\right\}$ and $p_2(v_2, w_2) = \min\left\{1, \frac{\pi_2^*(w_2|\tilde{x})q(v_2)}{\pi_2^*(v_2|\tilde{x})q(w_2)}\right\}$. Generate u from $Uniform(0,1)$. If $u <$

$p_1(v_1, w_1)$ we accept w_1 and else accept v_1 and also if $u < p_2(v_2, w_2)$ we accept w_2 and else accept v_2. Fix $t = t + 1$.

III : Repeat Step II , T times.

So, the retained sample values, say $c_1, \dots, c_{T-M_1}$, and $k_1, \dots, k_{T-M_2}$ are random samples from the posterior densities in the equations of (39) and (40), respectively. Now, by using Monte Carlo integration technique (Rodriguez (1977)), the Bayes estimates of c and k under squared error loss function are given by,

$$\hat{c}_{Bayes} = \frac{1}{T - M_1} \sum_{i=M_1+1}^{T} c_i^{(i)}, \hat{k}_{Bayes} = \frac{1}{T - M_2} \sum_{i=M_2+1}^{T} k_i^{(i)},$$

where M_1 and M_2 are the burn-in periods in generating $c_i^{(i)}$ and $k_i^{(i)}$,$(i = 1, \dots, n)$ respectively. We can also conduct the highest posterior density (HPD) confidence interval of parameter $\theta = (c, k)$. First order $c_1^{(i)}, \dots, c_{M_1}^{(i)}$ as $c_{(1)}^{(i)} < \dots < c_{(M_1)}^{(i)}$, then construct all the $(100(1-\eta)\%)$ confidence intervals of c are given by

$$\left(c_{(1)}^{(i)}, c_{([M_1(1-\eta)])}^{(i)}\right), \dots, \left(c_{([M_1\eta])}^{(i)}, c_{([M_1])}^{(i)}\right), \tag{41}$$

where $[M]$ is symbolized as the largest integer less than or equal to M. So, the HPD confidence interval of c is the shortest length interval. Similarly, we can make a $100(1-\eta)\%$ HPD confidence interval of k as follows:

$$\left(k_{(1)}^{(i)}, k_{([M_2(1-\eta)])}^{(i)}\right), \dots, \left(k_{([M_2\eta])}^{(i)}, k_{([M_2])}^{(i)}\right). \tag{42}$$

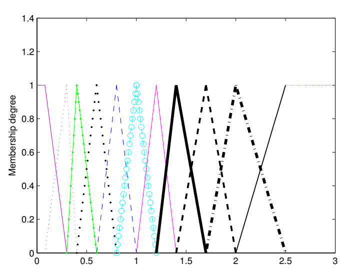

Figure 2: Fuzzy information system used to encode the simulated data.

6. Numerical Study

In the present section, some of the simulation studies are done to compare the performance of Bayesian estimation methods. All numerical computations are made using *MATLAB* software.

6.1. Monte Carlo simulations

In this section, some numerical results via Monte Carlo simulations are used to see how the different methods behave for various sample sizes. We evaluate the estimate of unknown parameters c and k using the methods provided in the preceding sections. For this reason the $i.i.d.$ random samples, say $\mathbf{x}$ of the BT XII distribution for parameter values, namely, $(c,k) = (1,1),(2,3),(3,1)$ and various choices of $n = 10,20,30,40,50,70,100$ are generated. Each realization of $\mathbf{x}$ was produced using the method proposed by Pak, Parham and Saraj (2014). By employing fuzzy information system (f.i.s.) $\{\tilde{x}_1,\ldots,\tilde{x}_{11}\}$ shown in Fig. 2, the corresponding membership functions are given by

$$\mu_{\tilde{x}_1}(x) = \begin{cases} 1 & x \le 0.08, \\ \frac{0.3-x}{0.22} & 0.08 \le x \le 0.3, \\ 0 & otherwise, \end{cases} \qquad \mu_{\tilde{x}_2}(x) = \begin{cases} \frac{x-0.08}{0.22} & 0.08 \le x \le 0.3, \\ \frac{0.4-x}{0.1} & 0.3 \le x \le 0.4, \\ 0 & otherwise, \end{cases}$$

$$\mu_{\tilde{x}_3}(x) = \begin{cases} \frac{x-0.3}{0.1} & 0.3 \le x \le 0.4, \\ \frac{0.6-x}{0.2} & 0.4 \le x \le 0.6, \\ 0 & otherwise, \end{cases} \qquad \mu_{\tilde{x}_4}(x) = \begin{cases} \frac{x-0.4}{0.2} & 0.4 \le x \le 0.6, \\ \frac{0.8-x}{0.2} & 0.6 \le x \le 0.8, \\ 0 & otherwise, \end{cases}$$

$$\mu_{\tilde{x}_5}(x) = \begin{cases} \frac{x-0.6}{0.2} & 0.6 \le x \le 0.8, \\ \frac{1-x}{0.2} & 0.8 \le x \le 1, \\ 0 & otherwise, \end{cases} \qquad \mu_{\tilde{x}_6}(x) = \begin{cases} \frac{x-0.8}{0.2} & 0.8 \le x \le 1, \\ \frac{1.2-x}{0.2} & 1 \le x \le 1.2, \\ 0 & otherwise, \end{cases}$$

$$\mu_{\tilde{x}_7}(x) = \begin{cases} \frac{x-1}{0.2} & 1 \le x \le 1.2, \\ \frac{1.4-x}{0.2} & 1.2 \le x \le 1.4, \\ 0 & otherwise, \end{cases} \qquad \mu_{\tilde{x}_8}(x) = \begin{cases} \frac{x-1.2}{0.2} & 1.2 \le x \le 1.4, \\ \frac{1.7-x}{0.3} & 1.4 \le x \le 1.7, \\ 0 & otherwise, \end{cases}$$

$$\mu_{\tilde{x}_9}(x) = \begin{cases} \frac{x-1.4}{0.3} & 1.4 \le x \le 1.7, \\ \frac{2-x}{0.3} & 1.7 \le x \le 2, \\ 0 & otherwise, \end{cases} \qquad \mu_{\tilde{x}_{10}}(x) = \begin{cases} \frac{x-1.7}{0.3} & 1.7 \le x \le 2, \\ \frac{2.5-x}{0.5} & 2 \le x \le 2.5, \\ 0 & otherwise, \end{cases}$$

$$\mu_{\tilde{x}_{11}}(x) = \begin{cases} \frac{x-2}{0.5} & 2 \le x \le 2.5 \\ 1 & x \ge 2.5, \\ 0 & otherwise, \end{cases}$$

6.2. Implementation

Since the MCMC method will stabilize asymptotically, it needs to examine the reliability of the chain outcome. *Burn-in* is a significant problem that is necessary to be considered. It means that discarding the number of iterations is essential. Some diagnostic tests that indicate the convergence problem can be found in the literature. One of them is trace plot in which the history of the chain is exhibited (see Fig. 3). Plots in Fig. 3, after discarding the initial 3000 iterates, show that the sequences have a stationary pattern. The estimates of the parameters c, k, and R for the fuzzy sample were obtained using the Bayesian approach. For simulation purpose, we have assumed that c, k have gamma priors, including the *noninformative* prior (**Prior I**), i.e. $a_1 = b_1 = a_2 = b_2 = 0$, *less informative* prior (**Prior II**), i.e.

$a_1 = b_1 = a_2 = b_2 = 0.01$, and *most informative* prior (**Prior III**), i.e. $a_1 = b_1 = a_2 = b_2 = 4$. We replicate the process 15000 times and use 12000 iterates after discarding the initial 3000 iterates as *Burn-in* to make the inference. We have also reported the average values (AV) and mean squared errors (MSE) of the estimates through Tables 1-3.

Table 1: The average values (AV) and the mean squared errors (MSE) of the estimate of parameters $\theta = (c,k) = (3,2)$, and $R = 0.7901$.

n			*priorI*			
	AV(c)	MSE(c)	AV(k)	MSE(k)	AV(R)	MSE(R)
10	3.6086	0.3704	3.5707	2.4673	0.7417	0.0031
20	3.3724	0.1387	2.4171	0.3905	0.7909	0.0023
30	3.0210	0.0864	2.6249	0.2004	0.7337	0.0017
40	3.0133	0.0166	2.4477	0.1739	0.7484	0.0014
50	2.7060	0.0102	1.9781	0.0007	0.7518	0.0004
70	2.8709	0.0004	2.0273	0.0004	0.7692	0.0002
100	2.8987	0.0001	2.0015	0.0000	0.7759	0.0001
n			*priorII*			
10	3.6193	0.3836	3.5889	2.5247	0.7427	0.0032
20	3.3789	0.1435	2.4162	0.3839	0.7920	0.0022
30	3.0121	0.0836	2.6196	0.1732	0.7331	0.0014
40	2.7108	0.0836	1.9768	0.0009	0.7526	0.0014
50	2.7108	0.0159	1.9768	0.0005	0.7526	0.0004
70	2.8737	0.0114	2.0306	0.0005	0.7693	0.0002
100	2.8930	0.0001	2.0027	0.0001	0.7750	0.0001
n			*priorIII*			
10	2.3191	0.4635	1.8712	0.0270	0.7080	0.0067
20	2.6611	0.2570	1.8932	0.0169	0.7526	0.0057
30	2.5557	0.1973	2.1302	0.0165	0.7140	0.0034
40	2.6613	0.1147	2.1151	0.0132	0.7310	0.0026
50	2.4930	0.1146	1.8355	0.0113	0.7389	0.0014
70	2.6819	0.1011	1.9200	0.0063	0.7557	0.0011
100	2.7654	0.0550	1.9227	0.0059	0.7669	0.0005

7. Conclusion

In this paper, we have examined the classical and Bayesian inference procedures for the BT XII distribution parameters, as well as the corresponding reliability parameter when the available data are described regarding fuzzy numbers. In this context, we considered three priors as **noninformative prior**, i.e. $a_1 = b_1 = a_2 = b_2 = 0$, **less informative prior**, i.e. $a_1 = b_1 = a_2 = b_2 = 0.01$, and **informative prior**, i.e. $a_1 = b_1 = a_2 = b_2 = 4$. The general results can be made from Tables 1-3 as follows. Considering the criterion MSE for all methods, with increasing n, the estimates are improved. The performance of the Bayes estimates with assumptions of noninformative prior and less informative prior regarding

AVs and MSEs, are almost identical. So, we prefer the prior XII since it will make the priors proper. The simulation study for all methods shows that the estimate of R is satisfactory, even for samples with sizes small and moderate. Using the NR or EM algorithms for the computation of MLEs gives similar estimation results. Because these two procedures have different features in the complexity of the iterative numerical search, we let users choose which to be used based on their preferences. The Bayes estimates obtained by Tierney and Kadane's approximation and the MCMC method behave in a very similar manner. However, from the computational point of view, Tierney and Kadane's procedure is easier to obtain. Note that these estimation results cannot be attributed to the assumed fuzzy numbers in Fig. 2. We have implemented the estimation procedures for different fuzzy numbers (not reported here) and found that the rationale for such fuzzy numbers, which are characterized by the membership functions $\mu_{\tilde{x}}(.)$ will not influence the estimate results.

Table 2: The AV and MSE of the estimate of parameters $\theta = (c,k) = (2,3)$, and $R = 0.5120$.

n			*priorI*			
	AV(c)	MSE(c)	AV(k)	MSE(k)	AV(R)	MSE(R)
10	2.0763	0.0333	6.2226	1.3855	0.3228	0.0357
20	2.0096	0.0058	3.9389	1.1535	0.4333	0.0126
30	1.9341	0.0043	4.0740	0.8816	0.3995	0.0067
40	2.0035	0.0029	3.8693	0.7557	0.4298	0.0061
50	1.8174	0.0002	3.0318	0.0300	0.4725	0.0015
70	1.9459	0.0002	3.1732	0.0010	0.4832	0.0008
100	1.9836	0.0001	3.0189	0.0003	0.5076	0.0001
n			*priorII*			
10	2.0566	0.0309	6.0537	9.3255	0.3306	0.0328
20	2.0021	0.0049	3.9133	1.1654	0.4336	0.0129
30	1.9297	0.0032	4.0795	0.8342	0.3982	0.0068
40	1.9996	0.0031	3.8646	0.7476	0.4292	0.0061
50	1.8241	0.0001	3.0440	0.0260	0.4727	0.0015
70	1.9438	0.0002	3.1613	0.0019	0.4841	0.0007
100	1.9867	0.0001	3.0146	0.0002	0.5088	0.0001
n			*priorIII*			
10	1.4479	0.3048	2.1876	0.6598	0.5174	0.0017
20	1.6671	0.1170	2.4551	0.2968	0.5173	0.0010
30	1.6578	0.1108	2.7872	0.1753	0.4702	0.0001
40	1.7714	0.0880	2.8966	0.0610	0.4794	0.0001
50	1.7032	0.0522	2.5812	0.0487	0.5034	0.0001
70	1.8432	0.0245	2.7791	0.0452	0.5069	0.0001
100	1.9136	0.0074	2.7528	0.0106	0.5242	0.0001

Table 3: The AV and MSE of the estimate of parameters $\theta = (c,k) = (1,1)$, and $R = 0.6667$.

n			*priorI*			
	AV(c)	MSE(c)	AV(k)	MSE(k)	AV(R)	MSE(R)
10	1.4393	0.1930	2.0352	1.0717	0.5414	0.0156
20	1.2573	0.1049	1.3105	0.0964	0.6350	0.0053
30	1.1349	0.0873	1.1559	0.0243	0.6497	0.0036
40	1.2375	0.0662	1.1183	0.0140	0.6740	0.0011
50	1.1726	0.0564	0.9733	0.0110	0.6999	0.0009
70	1.2954	0.0298	0.9305	0.0048	0.7274	0.0002
100	1.3239	0.0182	0.8951	0.0007	0.7399	0.0001
n			*priorII*			
10	1.4468	0.1996	2.0446	1.0913	0.5417	0.0156
20	1.2510	0.1058	1.3056	0.0934	0.6348	0.0052
30	1.1340	0.0891	1.1553	0.0241	0.6497	0.0036
40	1.2430	0.0630	1.1158	0.0134	0.6754	0.0012
50	1.1782	0.0590	0.9689	0.0103	0.7019	0.0010
70	1.2986	0.0317	0.9340	0.0043	0.7269	0.0002
100	1.3254	0.0179	0.8980	0.0009	0.7393	0.0001
n			*priorIII*			
10	1.2976	0.0903	1.5289	0.2797	0.5985	0.0046
20	1.2102	0.0886	1.2267	0.0514	0.6454	0.0045
30	1.1188	0.0732	1.1325	0.0175	0.6527	0.0029
40	1.2164	0.0468	1.1010	0.0102	0.6750	0.0009
50	1.1589	0.0441	0.9768	0.0086	0.6969	0.0004
70	1.2705	0.0252	0.9427	0.0032	0.7210	0.0002
100	1.3005	0.0141	0.9071	0.0005	0.7339	0.0001

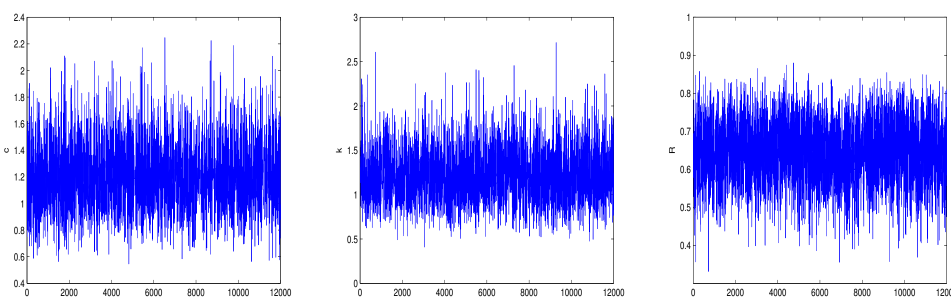

Figure 3: Plots of generated c versus iteration of MCMC (Gibss algorithm).

Acknowledgements

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

References

Akbari, M. G., Rezaei, A., (2007). A uniformly minimum variance unbiased point estimator using fuzzy observations. *Austrian Journal of Statistics*, 36 (4), pp. 307–317.

Ali Mousa, M. A. M., Jaheen, Z. F., (2002). Statistical inference for the Burr model based on progressively censored data functions. *Comput. Math. Applic.*, 43, pp. 1441–1449.

Burr, I. W., (1942). Cumulative frequency functions. *Annals of Mathematical Statistics*, 13 , pp. 215–222.

Burr, I. W., (1973). Parameters for a general system of distributions to match a grid of $\alpha_3$and α_4. *Communications in Statistics-Theory and Methods*, 2 , pp. 1–21.

Coppi, R., Gilb, M. A., Kiersc, H. A. L., (2006). The fuzzy approach to statistical analysis. *Computational Statistics and Data Analysis*, 51 (1), pp. 1–14.

Denoeux, T., (2011). Maximum likelihood estimation from fuzzy data using the EM algorithm, *Fuzzy Sets and Systems*, doi:10.1016/j.fss.2011.05.022.

Gertner, G. Z., Zhu, H., (1996). Bayesian estimation in forest surveys when samples or prior information are fuzzy. *Fuzzy Sets and Systems.*, 77(3), pp. 277–290.

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

Hanagl, D. D., Ahmadi, K. A., (2009). Bayesian estimation of the parameters of bivariate exponential distribution, *Communication in Statistics-Simulation and Computation*, 38, pp. 1391–1413.

Hate, M. A., (1949). A certain cumulative probability function. *Ann. Math. Statist.*, 20, pp. 461–63.

Huang, H., Zuo, M., Sun, Z., (2006). Bayesian reliability analysis for fuzzy lifetime data. *Fuzzy Sets and Systems*, 157(3), pp. 1674–1686.

Moore, D., Papadopoulos, A. S., (2000). The Burr type XII distribution as a failure model under various loss functions. *Microelectron. Reliab.*, 40, pp. 2117–2122, doi: 10.1016/S0026-2714(00)00031-7.

Pak, A., Parham, G. H., Saraj, M., (2013). Inference for the Weibull Distribution Based on Fuzzy Data. *Revista Colombiana de Estadistica*, 36(2), pp. 339–358.

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

Pak, A., (2016). Statistical Inference for the Parameter of Lindley Distribution Based on Fuzzy data. To appear in *Brazilian Journal of Probability and Statistics*.

Rodriguez, R. N., (1977). A guide to the Burr type XII Distributions. *Biometrika*, 64(1) , 129-134, doi:10.1093/biomet/64.1.129.

Rubinstein, R. Y., Kroese, D. P., (2006). *Simulation and the Monte Carlo method*. 2nd edition, John Wiley and Sons, Inc., Hoboken, New Jersey.

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

Shafiq, M., Atif, M., (2015). On the survival models for step-stress experiments based on fuzzy life time data. *Qual Quant*, doi: 10.1007/s11135-015-0295-9.

Singh S., Singh U., Kumar M., (2016). Bayesian estimation for Poission-exponential model under Progressive type-II censoring data with Binomial removal and its application to ovarian cancer data, *Communications in Statistics-Simulation and Computation*, 45, pp. 3457–3475.

Singpurwalla, N. D., Booker, J. M., (2004). Membership functions and probability measures of fuzzy sets, *Journal of the American Statistical Association*, 99(467), pp. 867–877.

Soliman, A. A., (2005). Estimation of parameters of life from progressively censored data using Burr-XII model. *IEEE Trans. Reliab.*, 54, 34–42.

Tadikamalla, P. R., (1980). A Look at the Burr and Related Distributions. *International Statistical Review*, 48(3) , pp. 337–344.

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

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

Viertl, R., (2011). *Statistical Methods for Fuzzy Data*, Wiley, Chichester.

Viertl, R., (2006). Univariate statistical analysis with fuzzy data. *Computational Statistics & Data Analysis*, 55(1), pp. 133–147.

Wingo, D. R., (1983). Maximum likelihood methods for fitting the Burr Type XII distribution to life test data. *Biom J.*, 25, pp. 77–84.

Wingo, D. R., (1993). Maximum likelihood methods for fitting the Burr type XII distribution to multiply(progressively) censored life test data. *Metrika*, 40, pp. 203–210.

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

Wu, J. W., Yu, H. Y., (2005). Statistical inference about the shape parameter of the Burr type XII distribution under the failure-censored sampling plan. *Applied Math. Computat.*, 163, pp. 443–482.

Xiuchun, L., S. Yimin, W. Jieqiong, C., Jian, (2007). Empirical Bayes estimators of reliability performances using LINEX loss under progressively Type-II censored samples. *Math. Comput. Simulat.*, 73, pp. 320–326.

Zadeh, L., (1965). Fuzzy sets *Information and Control*, 8(3), pp. 338–353.

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

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