ISSN: 0017-2294

DOI Link: https://doi.org/10.61586/5h16b
Vol.65, Issue.5, Part.1, May 2025, PP. 17-36

Inference drawn on Logistic Distribution in the Presence of Fuzzy Data: Classical and Bayesian Approach

Lingutla Vijay Kumar, Nadiminti Nagamani*

Department of Mathematics, School of Advanced Sciences, VIT-AP University, Inavolu, Beside AP Secretariat, Amaravati AP, India.

Abstract

This paper investigates parameter estimation for the logistic distribution in the presence of fuzzy data, which introduces uncertainty and imprecision into the analysis. To handle this, we employ both classical and Bayesian frameworks. As closed-form solutions for Maximum Likelihood Estimators (MLEs) are not available under fuzzy conditions, the Expectation-Maximization (EM) algorithm is utilized to obtain numerical estimates. Asymptotic confidence intervals are then constructed based on these estimates. For Bayesian inference, we consider a conjugate prior and apply Lindley's approximation and Approximate Bayesian Computation (ABC) to derive Bayes estimators, addressing the analytical intractability of the posterior distributions. Furthermore, advanced Markov Chain Monte Carlo (MCMC) methods namely Hamiltonian Monte Carlo (HMC) and the Metropolis-Hastings (MH) algorithm are employed to sample from posterior distributions and to construct Highest Posterior Density (HPD) intervals. The performance of all proposed estimators is assessed through extensive simulation studies. Finally, the practical utility of the methodologies is illustrated using a real-world dateset.

Keywords: Logistic Distribution, Fuzzy data, Bayesian estimation, Lindley approximation, Approximate Bayesian Computation, Metropolis-Hastings, Hamiltonian Monte Carlo.

1 Introduction

In recent years, analyzing data that is uncertain or imprecise has become increasingly important across diverse domains such as engineering, economics, and reliability analysis. In many real-world scenarios, limitations in measurement instruments, subjective assessments, or incomplete observations lead to data that cannot be captured precisely. To effectively address this challenge, fuzzy data—grounded in fuzzy set theory—offers a robust mathematical framework for modeling such imprecision.

The logistic distribution has been widely studied in the context of precise data, due to its versatility and applications in fields such as growth modeling, survival

*Corresponding author E-mail address:maniinadiminti999@gmail.com

analysis, and logistic regression. Seminal works like those of [1] and [2] have thoroughly detailed its statistical properties. Classical estimation techniques, including the method of moments, maximum likelihood estimation (MLE), and Bayesian inference, have been extensively applied to this distribution [3, 4].

As the relevance of imprecise and vague data becomes more pronounced, fuzzy data modeling has gained momentum. Introduced by Zadeh [5], fuzzy sets provide a formal way to represent ambiguity and vagueness in data. Numerous studies have since extended traditional estimation methods to fuzzy settings, particularly for distributions like the normal, exponential, gamma, and Weibull. For example, [6] analyzed the Weibull distribution using MLE and Bayesian approaches under fuzzy observations, while [7] applied Bayesian estimation to the weighted exponential distribution under different loss functions using Lindley's approximation. Other notable works include [8] on the Lindley distribution under fuzzy data and [9] who incorporated Monte Carlo simulations for Bayesian inference. More recently, [10] and [11] investigated the generalized exponential and Weibull models using Bayesian methods for fuzzy datasets. [12] addressed parameter estimation in the gamma distribution within fuzzy frameworks. These studies highlight the growing interest in fuzzy data analysis, particularly for single-population models.

However, the logistic distribution has received comparatively less attention in the context of fuzzy data. Existing works rarely combine both MLE and Bayesian techniques under fuzzy uncertainty. Furthermore, there is a lack of comprehensive studies validating these estimation procedures through simulation and real-world applications. To bridge this gap, the present study proposes robust parametric estimation methods for the logistic distribution under triangular fuzzy data representation.

While MLE remains a popular and asymptotically efficient method for parameter estimation in classical statistics, it is sensitive to deviations from model assumptions and may perform poorly when data are uncertain or contain outliers [13]. This motivates the need to explore Bayesian computational techniques, which offer greater flexibility under uncertainty.

This study investigates three Bayesian estimation approaches—Approximate Bayesian Computation (ABC), Hamiltonian Monte Carlo (HMC), and the Metropolis-Hastings (MH) algorithm. ABC is particularly advantageous when the likelihood function is difficult or impossible to compute, relying instead on summary statistics [14]. HMC utilizes gradient information to improve sampling efficiency in high-dimensional spaces, reducing autocorrelation and improving convergence speed [15, 16]. Betancourt [17] further elaborated on the conceptual strengths of HMC in handling complex posterior landscapes. Meanwhile, the MH algorithm, despite being older, remains foundational in Bayesian inference due to its simplicity and general-purpose adaptability [18, 19].

This paper aims to estimate the location and scale parameters of the logistic distribution when data are represented as triangular fuzzy numbers. Both classical (MLE) and Bayesian techniques are utilized. For the Bayesian framework, prior information is incorporated, and Bayes estimators are obtained using Lindley's approximation, ABC, and MCMC techniques such as HMC and MH.

To evaluate the effectiveness of these estimators, an extensive Monte Carlo simulation study is conducted under varying levels of fuzziness and sample sizes. Addi-

 ISSN: 0017-2294

tionally, a real-life dataset is used to demonstrate practical applicability by converting crisp data into triangular fuzzy values. The results indicate that the proposed methods perform well in estimating logistic parameters even in the presence of uncertainty.

1.1 Key Contributions

The major contributions of this work are summarized below:

- A novel estimation framework for incorporating fuzzy data into logistic distribution analysis.
- A comparative study of classical MLE and advanced Bayesian techniques, including Lindley's approximation, ABC, HMC, and MH, under fuzzy data scenarios.
- Empirical validation through extensive simulation studies and real-world data application.
- Advancement of statistical methodologies to enhance decision-making under uncertainty.

By integrating fuzzy data models with modern Bayesian computational techniques, this research contributes to more robust statistical inference, particularly in applications where data ambiguity is common, such as biomedical sciences, reliability engineering, and the social sciences.

The remainder of the paper is structured as follows: Section 2 introduces foundational concepts related to fuzzy data and triangular fuzzy numbers. Section 3 presents the MLE approach for logistic parameters using the Expectation-Maximization (EM) algorithm. Section 4 discusses Bayesian estimation techniques including Lindley's approximation, ABC, and MCMC methods (HMC and MH). Section 5 contains simulation studies evaluating estimator performance. Section 6 applies the methods to real-world data. Finally, Section 7 concludes with a summary of findings and suggestions for future research.

2 Theoretical Background

Fuzzy set theory, first introduced by [5], offers a mathematical framework designed to address uncertainty and imprecision by allowing elements to have partial membership in sets, unlike classical set theory which uses a binary membership approach. In this framework, each element x of a fuzzy set is assigned a degree of membership through a membership function $\mu_A(x)$, which maps x to a value within the interval $[0, 1]$.

For any fuzzy set A, the membership function $\mu_A(x)$ satisfies the condition:

$$\mu_A(x) \in [0,1], \quad \text{for all } x.$$

To handle the imprecision inherent in fuzzy data, [20] extended the concept to fuzzy probability. In this context, the probability of a fuzzy event $\tilde{A}$ is expressed as:

$$P(\tilde{A}) = \int \mu_{\tilde{A}}(x)\, dP(x), \tag{2.1}$$

ISSN: 0017-2294

where $\mu_{\tilde{A}}(x)$ represents the membership function associated with the fuzzy event $\tilde{A}$.

This membership function quantifies the degree of truth that an element belongs to the set, offering a way to handle situations where exact values are unavailable. For example, in medical studies, it can be used to represent the level of certainty about a patient's survival time or diagnosis when dealing with uncertain or incomplete data.

In survival analysis, fuzzy methods are particularly valuable in situations where data is incomplete or uncertain—such as imprecise survival times, often resulting from missing records, recall bias, or measurement errors. These techniques enable researchers to model and analyze data that is inherently imprecise, a common occurrence in fields like medicine and clinical research.

Triangular Membership Function

Triangular membership functions are commonly employed to represent fuzzy numbers, particularly when data uncertainty is derived from expert judgments or subjective assessments. A triangular fuzzy number is defined by three parameters: the lower bound a, the most likely or peak value b, and the upper bound c. These parameters represent the minimum, the most probable, and the maximum values of the fuzzy quantity, respectively.

The membership function $\mu_A(x)$ for a triangular fuzzy number is given by:

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a} & \text{for } a \le x \le b, \\ \frac{c-x}{c-b} & \text{for } b < x \le c, \\ 0 & \text{otherwise.} \end{cases}$$

This piecewise linear function illustrates the degree of membership of an element x in the fuzzy set A. Triangular fuzzy numbers are particularly advantageous in cases where precise measurements are unavailable, but approximate values based on expert opinion or judgment are accessible.

For instance, if a clinician estimates that a patient's survival time is most likely 12 months but could range from 10 to 15 months, a triangular fuzzy number $(10, 12, 15)$ can be used to capture this uncertainty.

In the following sections, we will explore the estimation of the parameters of the logistic distribution under fuzzy data using both Maximum Likelihood and Bayesian methods. Specifically, we apply the Expectation-Maximization (EM) algorithm for MLE, while for Bayesian inference, we use Lindley's approximation, Approximate Bayesian Computation (ABC), and Markov Chain Monte Carlo (MCMC) techniques, including Metropolis-Hastings (MH) and Hamiltonian Monte Carlo (HMC).

3 Maximum Likelihood Estimation

Consider a random sample $X = (x_1, x_2, \ldots, x_m)$ drawn from a logistic distribution with location parameter μ and scale parameter s. The probability density function (PDF) of the logistic distribution is given by:

$$f_X(x; \mu, s) = \frac{e^{-(x-\mu)/s}}{s\left(1 + e^{-(x-\mu)/s}\right)^2}, \quad -\infty < x < \infty.$$

In the presence of fuzzy data, each observation x_i is represented by a fuzzy number $\tilde{X}_i$ with an associated membership function $\mu_{\tilde{X}_i}(x)$. The fuzzy likelihood function is then defined as:

$$L(\mu, s \mid \tilde{X}) = \prod_{i=1}^{m} P(\tilde{X}_i), \tag{3.2}$$

where $P(\tilde{X}_i)$ denotes the fuzzy probability of the i-th observation, evaluated as:

$$P(\tilde{X}_i) = \int_{a_i}^{c_i} \mu_{\tilde{X}_i}(x) f_X(x; \mu, s)\, dx.$$

Substituting the logistic PDF into the above equation, the likelihood function becomes:

$$L(\mu, s \mid \tilde{X}) = \frac{1}{s^m} \prod_{i=1}^{m} \int_{a_i}^{c_i} \frac{e^{-(x-\mu)/s}}{\left(1 + e^{-(x-\mu)/s}\right)^2} \mu_{\tilde{X}_i}(x)\, dx.$$

The log-likelihood function is obtained by taking the natural logarithm of the likelihood function:

$$\ell(\mu, s) = -m \log s + \sum_{i=1}^{m} \log \int_{a_i}^{c_i} \frac{e^{-(x-\mu)/s}}{\left(1 + e^{-(x-\mu)/s}\right)^2} \mu_{\tilde{X}_i}(x)\, dx. \tag{3.3}$$

To obtain the maximum likelihood estimates (MLEs) of μ and s, we differentiate the log-likelihood function and solve the following score equations:

$$\sum_{i=1}^{m} \frac{\int_{a_i}^{c_i} \left(\frac{x-\mu}{s^2}\right) \frac{e^{-(x-\mu)/s}}{\left(1+e^{-(x-\mu)/s}\right)^2} \mu_{\tilde{X}_i}(x)\, dx}{\int_{a_i}^{c_i} \frac{e^{-(x-\mu)/s}}{\left(1+e^{-(x-\mu)/s}\right)^2} \mu_{\tilde{X}_i}(x)\, dx} = 0,$$

$$-\frac{m}{s} + \sum_{i=1}^{m} \frac{\int_{a_i}^{c_i} \left(\frac{(x-\mu)^2}{s^3}\right) \frac{e^{-(x-\mu)/s}}{\left(1+e^{-(x-\mu)/s}\right)^2} \mu_{\tilde{X}_i}(x)\, dx}{\int_{a_i}^{c_i} \frac{e^{-(x-\mu)/s}}{\left(1+e^{-(x-\mu)/s}\right)^2} \mu_{\tilde{X}_i}(x)\, dx} = 0.$$

These equations are nonlinear, and they can be solved using numerical optimization techniques. In this study, the Expectation-Maximization (EM) algorithm is used to estimate the parameters μ and s under fuzzy observations.

3.1 Expectation-Maximization Algorithm for Logistic Distribution

The Expectation-Maximization (EM) algorithm [21] is an iterative method for estimating parameters in the presence of incomplete or imprecise data. For a single logistic population, the complete-data log-likelihood is given by:

$$\ell = -m \log s + \sum_{i=1}^{m} \log \int_{a_i}^{c_i} \frac{\exp\left(-\frac{x-\mu}{s}\right)}{\left(1 + \exp\left(-\frac{x-\mu}{s}\right)\right)^2} \mu_{\tilde{X}_i}(x)\, dx. \tag{3.4}$$

The steps of the EM algorithm are as follows:

- **Initialization:** Select initial values $\mu^{(0)}$, $s^{(0)}$, and set the iteration index $h = 0$.

- **E-step:** Compute the expected values using the following formula:

$$E(x_i \mid \tilde{X}_i) = \frac{\int_{a_i}^{c_i} x\, f_X(x; \mu^{(h)}, s^{(h)}) \mu_{\tilde{X}_i}(x)\, dx}{\int_{a_i}^{c_i} f_X(x; \mu^{(h)}, s^{(h)}) \mu_{\tilde{X}_i}(x)\, dx}.$$

- **M-step:** Update the parameters as follows:

$$\mu^{(h+1)} = \frac{1}{m} \sum_{i=1}^{m} E(x_i \mid \tilde{X}_i), \tag{3.5}$$

$$s^{(h+1)} = \frac{1}{m} \sum_{i=1}^{m} \left| E(x_i \mid \tilde{X}_i) - \mu^{(h+1)} \right|. \tag{3.6}$$

- **Convergence Check:** If $|\mu^{(h+1)} - \mu^{(h)}| < \epsilon$ and $|s^{(h+1)} - s^{(h)}| < \epsilon$, terminate the algorithm. Otherwise, increment h and repeat the steps.

After obtaining the MLEs, the asymptotic 95% confidence intervals for μ and s are computed using the Fisher information matrix as follows:

$$\hat{\mu} \pm 1.96\sqrt{I_{11}^{-1}}, \quad \hat{s} \pm 1.96\sqrt{I_{22}^{-1}},$$

where I^{-1} is the inverse of the Fisher information matrix. The Fisher information for the logistic distribution can be derived analytically and used to compute the asymptotic confidence intervals.

4 Bayesian Estimation

Bayesian estimation offers a robust framework for parameter inference by combining prior knowledge with observed data. Consider a logistic distribution with fuzzy data $\tilde{X}$ and parameter vector $\theta = (\mu, s)$, where μ is the location parameter and s is the scale parameter. The posterior distribution of θ given the fuzzy observations is expressed as:

$$p(\theta \mid \tilde{X}) = \frac{p(\tilde{X} \mid \theta) p(\theta)}{\int p(\tilde{X} \mid \theta) p(\theta)\, d\theta}, \tag{4.7}$$

where $p(\tilde{X} \mid \theta)$ denotes the likelihood of the fuzzy data, and $p(\theta)$ represents the prior distribution for the parameters.

The Bayesian estimator of a parameter function $\pi(\theta)$, such as a parameter or a function of the parameters, is the posterior expectation under the squared error loss function (SELF). This is given by:

$$\hat{\pi}_{\text{Bayes}} = E[\pi(\theta) \mid \tilde{X}] = \frac{\int \pi(\theta) p(\tilde{X} \mid \theta) p(\theta)\, d\theta}{\int p(\tilde{X} \mid \theta) p(\theta)\, d\theta}. \tag{4.8}$$

Due to the complexity of the integrals involving fuzzy data, exact evaluation of the posterior distribution is often impractical. Thus, several numerical methods are applied to approximate the Bayesian estimators:

- **Lindley's Approximation:** This method provides an approximation of the posterior distribution by employing a second-order Taylor series expansion around the mode, which simplifies the integral evaluation.
- **Approximate Bayesian Computation (ABC):** A simulation-based technique that approximates the posterior distribution by comparing summary statistics from simulated and observed data, which is particularly useful when the likelihood function is not readily available.
- **Hamiltonian Monte Carlo (HMC):** A sophisticated Markov Chain Monte Carlo (MCMC) method that utilizes Hamiltonian dynamics to efficiently explore complex posterior distributions, especially in high-dimensional parameter spaces.
- **Metropolis-Hastings Algorithm:** A popular MCMC method for generating posterior samples based on an acceptance-rejection mechanism, which helps in approximating the posterior distribution when direct computation is infeasible.

For the sake of analytical simplicity, a conjugate prior distribution is assumed in all methods. The estimations are performed under the squared error loss function, which yields the posterior mean as the Bayes estimate of the parameters.

4.1 Lindley's Approximation

Lindley's approximation [23] is a numerical method used to approximate the Bayes estimator when the posterior distribution cannot be expressed in a closed form. For the parameter vector $\theta = (\mu, s)$, the Bayes estimator under the squared error loss function is the posterior mean:

$$\hat{\theta}_{\text{Bayes}} = \mathbb{E}[\theta \mid \tilde{X}] = \int \theta \, p(\theta \mid \tilde{X}) \, d\theta. \tag{4.9}$$

The procedure to compute Lindley's approximation is as follows:

1. Assume independent priors for the parameters. Let the location parameter μ follow a normal distribution $\mu \sim \mathcal{N}(\alpha, \tau^2)$ and the scale parameter s follow a Gamma distribution $s \sim \text{Gamma}(b, c)$, where α and τ^2 represent the mean and variance of the normal prior, and b, c are the shape and rate parameters of the gamma prior.
2. Write the posterior distribution in terms of the likelihood and prior as:
$$p(\theta \mid \tilde{X}) \propto p(\tilde{X} \mid \theta) p(\theta),$$
where $p(\tilde{X} \mid \theta)$ is the fuzzy likelihood and $p(\theta)$ is the prior distribution for the parameters.
3. Define the log-posterior function as $\ell(\theta) = \log p(\theta \mid \tilde{X})$.
4. Find the mode $\hat{\theta}$ of the posterior distribution by maximizing the log-posterior $\ell(\theta)$.

5. Compute the Hessian matrix H, which is the matrix of second derivatives of the negative log-posterior $-\ell(\theta)$, evaluated at the mode $\hat{\theta}$.

6. Approximate the posterior distribution by performing a second-order Taylor expansion of the log-posterior function around the mode $\hat{\theta}$. This approximation leads to a normal distribution:

$$p(\theta \mid \tilde{X}) \approx \mathcal{N}(\hat{\theta}, -H^{-1}).$$

7. Use the approximate posterior distribution to compute the Bayes estimator, which is the mode of the posterior:

$$\hat{\theta}_{\text{Bayes}} \approx \hat{\theta}.$$

8. If the posterior distribution is highly non-linear, refine the mode $\hat{\theta}$ by using gradient-based optimization techniques to obtain a more accurate estimate.

4.2 Approximate Bayesian Computation (ABC)

Approximate Bayesian Computation (ABC) is a powerful technique particularly useful when the exact likelihood function is either difficult or infeasible to compute, which is often the case with fuzzy or uncertain data, such as those modeled by the Logistic distribution. ABC circumvents the need for explicit likelihood calculations by leveraging simulations alongside the observed data to approximate the posterior distribution. This technique allows prior beliefs about the parameters to be incorporated and uncertainty in the posterior to be quantified. ABC is particularly advantageous in fields like reliability and survival analysis, where decision-making occurs under uncertainty. By using summary statistics and a suitable distance metric, ABC can handle fuzzy data and irregular data structures with ease.

The steps involved in applying the ABC method for a single logistic population are outlined below:

1. Model the data with a logistic distribution:

$$X \sim \text{Logistic}(\mu, s),$$

where μ and s represent the location and scale parameters, respectively.

2. Define prior distributions for the parameters:

$$\mu \sim \mathcal{N}(\alpha, \tau^2), \quad s \sim \text{Gamma}(b, c),$$

where α and τ^2 are the mean and variance of the normal prior for μ, and b, c are the shape and rate parameters for the gamma prior on s.

3. Draw samples for μ and s from their respective prior distributions, and use these to simulate synthetic fuzzy data $\tilde{X}^*$ based on the logistic model.

4. Calculate summary statistics (such as the mean and variance) for both the observed fuzzy data $\tilde{X}$ and the simulated data $\tilde{X}^*$, denoted as S_{obs} and S_{sim}, respectively.

ISSN: 0017-2294

5. Compute the distance between the observed and simulated summary statistics using an appropriate distance metric, such as the Euclidean distance:
$$d = \|S_{\text{obs}} - S_{\text{sim}}\|.$$
6. Accept the sampled parameters if the distance d is less than or equal to a pre-determined tolerance level ϵ, i.e., $d \leq \epsilon$.
7. Repeat this process many times to gather a set of accepted parameter samples, which will approximate the posterior distribution.
8. Estimate the parameters of interest by computing summary statistics (e.g., the posterior mean or median) from the accepted samples. Credible intervals can also be derived from the empirical distribution of these samples.

In summary, ABC provides a versatile and effective framework for Bayesian inference in complex data scenarios, such as when working with fuzzy data. By approximating the posterior distribution without needing the explicit likelihood function, ABC successfully manages parameter uncertainty and aids in reliability analysis within logistic models.

4.3 Hamiltonian Monte Carlo (HMC)

Hamiltonian Monte Carlo (HMC) is an advanced Markov Chain Monte Carlo (MCMC) technique designed to efficiently sample from intricate posterior distributions in Bayesian inference. Unlike traditional MCMC methods, which propose new samples based solely on previous ones, HMC leverages Hamiltonian dynamics to guide the exploration of parameter space. This approach improves the efficiency of sampling by reducing the correlation between successive samples, making it particularly beneficial when working with complex, high-dimensional data such as fuzzy data modeled by the Logistic distribution.

The procedure for implementing HMC with a fuzzy logistic population is outlined in the following steps:

1. Model the fuzzy data $\tilde{X}$ using a logistic distribution:
$$X \sim \text{Logistic}(\mu, s),$$
where μ represents the location parameter and s represents the scale parameter. The data points are represented as triangular fuzzy numbers with lower bounds L_i, peak values P_i, and upper bounds U_i.
2. Specify prior distributions for the parameters:
$$\mu \sim \mathcal{N}(\alpha, \tau^2), \quad s \sim \text{Gamma}(b, c),$$
where α and τ^2 are the mean and variance of the normal prior for μ, and b, c are the shape and rate parameters of the gamma prior for s.
3. Combine the fuzzy likelihood and prior distributions to construct the posterior distribution:
$$\text{Posterior}(\mu, s \mid X) \propto L(\mu, s \mid X) \cdot \pi(\mu) \cdot \pi(s).$$

4. Compute the gradient of the log-posterior with respect to the parameters:

$$\frac{\partial}{\partial\theta}\log\text{Posterior}(\mu, s), \quad \text{where } \theta = (\mu, s).$$

5. Define the Hamiltonian function as:

$$H(\theta, p) = -\log\text{Posterior}(\theta) + \frac{1}{2}p^\top p,$$

where p is the momentum vector, sampled from a standard normal distribution, and has the same dimension as θ.

6. Initialize the parameter vector $\theta_0 = (\mu_0, s_0)$, and sample the momentum $p \sim \mathcal{N}(0, I)$.

7. Use the leapfrog integration method to update θ and p iteratively:

$$p^{(t+\frac{1}{2})} = p^{(t)} - \frac{\epsilon}{2}\nabla_\theta \log\text{Posterior}(\theta^{(t)}),$$

$$\theta^{(t+1)} = \theta^{(t)} + \epsilon \cdot p^{(t+\frac{1}{2})},$$

$$p^{(t+1)} = p^{(t+\frac{1}{2})} - \frac{\epsilon}{2}\nabla_\theta \log\text{Posterior}(\theta^{(t+1)}),$$

where ϵ is the step size, and t is the current iteration number.

8. Apply the Metropolis acceptance criterion to decide whether to accept or reject the proposed sample, based on the Hamiltonian value.

9. Repeat the process over many iterations to generate a set of samples from the posterior distribution.

10. Estimate the parameters by calculating the posterior mean or median from the accepted samples. Credible intervals, such as the Highest Posterior Density (HPD) intervals, can also be derived from the posterior distribution.

HMC offers an efficient and robust framework for sampling from posterior distributions in Bayesian analysis, especially when dealing with complex models, such as those involving fuzzy data. Its utilization of gradient information accelerates convergence, making it highly effective for inference in high-dimensional, uncertain models.

4.4 Metropolis-Hastings (MH) Algorithm

The Metropolis-Hastings (MH) algorithm is a Markov Chain Monte Carlo (MCMC) method used to generate samples from a posterior distribution when direct sampling is not feasible. This method is particularly valuable for complex models, such as the Logistic distribution with fuzzy data. The algorithm constructs a Markov chain whose stationary distribution corresponds to the desired posterior, utilizing a proposal mechanism and an acceptance rule.

The steps to apply the MH algorithm to a single fuzzy logistic population are as follows:

ISSN: 0017-2294

1. Assume the observed fuzzy data $\tilde{X}$ follows a Logistic distribution:

$$X \sim \text{Logistic}(\mu, s),$$

where μ is the location parameter and s is the scale parameter. The data points are modeled as triangular fuzzy numbers, with lower L_i, peak P_i, and upper U_i values.

2. Specify the prior distributions for the parameters:

$$\mu \sim \mathcal{N}(\alpha, \tau^2), \quad s \sim \text{Gamma}(b, c),$$

where α, τ^2, b, and c are the hyperparameters of the prior distributions.

3. The posterior distribution is proportional to the product of the fuzzy likelihood and the prior:

$$p(\mu, s \mid \tilde{X}) \propto L(\tilde{X} \mid \mu, s) \cdot p(\mu) \cdot p(s).$$

4. Use a symmetric proposal distribution, such as a bivariate normal distribution, to propose new parameter values:

$$q((\mu^*, s^*) \mid (\mu, s)) = \mathcal{N}((\mu, s), \Sigma),$$

where Σ is the covariance matrix that controls the step size for the proposal.

5. Generate a proposed sample (μ^*, s^*) from the proposal distribution.

6. Compute the acceptance probability:

$$\alpha = \min\left(1, \frac{L(\tilde{X} \mid \mu^*, s^*) \cdot p(\mu^*) \cdot p(s^*)}{L(\tilde{X} \mid \mu, s) \cdot p(\mu) \cdot p(s)}\right).$$

7. Sample a uniform random number $u \sim U(0,1)$. If $u \leq \alpha$, accept the proposed sample:

$$(\mu, s) \leftarrow (\mu^*, s^*).$$

Otherwise, retain the current values.

8. Repeat the sampling process for N iterations to generate a sequence of samples from the posterior distribution.

9. Use the collected samples to estimate posterior quantities such as:
   - Posterior means or medians of μ and s,
   - Credible intervals (e.g., 95% HPD or quantile-based).

The MH algorithm provides a flexible and practical approach for Bayesian inference in logistic models involving fuzzy data. Its simplicity and versatility make it a valuable tool for parameter estimation in fields such as reliability and survival analysis.

ISSN: 0017-2294

5 Simulation Study

In this section, we present the results of a simulation study designed to evaluate the performance of various estimation methods for a single logistic population under different sample sizes. We estimate the location parameter μ and the scale parameter s using Maximum Likelihood (ML) estimates with the Expectation-Maximization (EM) algorithm, as well as Bayesian estimates obtained via Lindley's approximation, Approximate Bayesian Computation (ABC), Hamiltonian Monte Carlo (HMC), and the Metropolis-Hastings (MH) algorithm.

Two sets of parameter values are considered in the simulations:

- Set 1: $(\mu, s) = (1, 2)$ (see Table 1)
- Set 2: $(\mu, s) = (1.5, 2.5)$ (see Table 2)

Fuzzy numbers for the observed data are generated by sampling from a logistic distribution, where each fuzzy number is represented by its lower, peak, and upper values:

- The peak values are drawn from the logistic distribution.
- The lower and upper bounds are generated by subtracting and adding small uniform random values to the peak values, respectively.

The membership values for these fuzzy numbers are defined using a triangular membership function, allowing the representation of the degree of membership within each fuzzy interval.

To assess the performance of the estimators, we use the following evaluation metrics:

- Average Value (Avg)
- Mean Squared Error (MSE)

For comparison, 10,000 random samples are generated from the logistic population under varying sample sizes and parameter configurations. Specific hyperparameters ($b_1 = b_2 = 1$, $c_1 = c_2 = 2$) are used to compute bias and MSE for all estimators. The results are summarized in Tables 1 and 2.

Table 1 and Table 2 present the average estimates and corresponding MSEs for each method under squared error loss. We observe that the EM and HMC estimators generally produce more accurate estimates with lower MSEs, especially as the sample size increases. The MH estimates perform slightly worse for smaller sample sizes but improve as the sample size grows.

Additionally, to quantify the uncertainty associated with the estimators, we compute the 95% confidence intervals using asymptotic theory (for ML estimates) and the 95% Highest Posterior Density (HPD) intervals using HMC and MH estimates. These intervals are reported in Tables 3 and 4. As expected, the interval widths decrease with larger sample sizes, reflecting improved precision of the estimates.

Among the Bayesian methods, HMC consistently produces narrower HPD intervals than MH, indicating better convergence and efficiency. The asymptotic intervals based on ML estimates tend to be wider for smaller sample sizes but become comparable with the Bayesian intervals as the sample size increases.

The simulation results lead to the following key observations:

- As the sample size increases, the MSE of all methods decreases, indicating that estimation improves with more data.
- The EM algorithm produces accurate and stable estimates across all sample sizes, with relatively low computational cost.
- Lindley's approximation provides reasonable estimates but is less accurate for small sample sizes due to linearization error.
- Approximate Bayesian Computation (ABC) performs well for moderate sample sizes, though its accuracy is influenced by the choice of summary statistics and tolerance levels.
- Hamiltonian Monte Carlo (HMC) offers the most accurate estimates, with the lowest MSE and narrowest HPD intervals, albeit at a higher computational cost.
- Metropolis-Hastings (MH) sampling yields competitive estimates but may require careful tuning to ensure convergence, particularly for smaller sample sizes.
- HPD intervals obtained from HMC and MH methods are generally narrower and more informative than the asymptotic confidence intervals from the EM method, reflecting better uncertainty quantification.
- Bayesian methods, particularly HMC and MH, are more robust to smaller sample sizes compared to the EM and Lindley methods.
- While the EM algorithm is efficient and scalable, Bayesian methods offer richer inference through posterior distributions and are better suited for incorporating prior information.

6 Real Data Analysis

To assess the performance of our model, we utilized data from an engineering dataset provided by [6], which includes a case study on the light-emitting diodes (LED) manufacturing process, specifically focusing on the luminous intensities of LED sources. The data is shown in Table 5.

In our approach, each triplet is modeled by triangular fuzzy numbers $\tilde{x}_i$ and $\tilde{y}_j$, representing possibility distributions corresponding to unknown values x_i, which are realizations of the random variable X_i. To estimate the parameters (μ, s) of the Logistic distribution, we employed both MLE and Bayesian estimation techniques, considering the fuzzy nature of the dataset. Specifically, Bayesian estimation was carried out using Lindley's approximation, ABC, and MCMC methods, including HMC and the MH algorithm.

The results, as presented in Table 6, provide valuable insights into the performance of each estimation method under conditions of uncertainty and imprecision. Furthermore, Table 7 reports the asymptotic confidence intervals obtained via MLE and the 95% HPD intervals computed using the HMC and MH algorithms, offering a comparative assessment of the precision and reliability of the parameter estimates.

Table 1: For various sample sizes, we compare Estimated value and mean square errors of multiple estimators under squared error loss for $\theta = (\mu, s)$=(1, 2).

	EM		Lindley		ABC		HMC		MH	
Size	Avg	Error	Avg	Error	Avg	Error	Avg	Error	Avg	Error
20	1.0207	0.6196	1.2954	0.0917	1.0858	0.0474	0.9555	0.4699	0.9487	0.5983
	1.9320	0.1400	1.7305	0.2232	1.6989	0.1956	2.2705	0.2566	2.3048	0.4029
50	1.0046	0.2189	1.2406	0.0616	1.0563	0.0427	1.005	0.2125	0.9969	0.2102
	1.9808	0.0612	1.8192	0.1163	1.7436	0.1397	2.052	0.0618	2.0573	0.0716
70	0.9969	0.1739	1.2106	0.0477	1.0028	0.0387	0.9911	0.1491	0.9939	0.1497
	1.9901	0.0403	1.8311	0.0911	1.6266	0.1214	1.959	0.0394	1.9565	0.0394
80	0.9750	0.1571	1.1510	0.0254	1.0100	0.0366	0.8932	0.1446	0.9945	0.1252
	1.9909	0.0357	1.8269	0.0855	2.2728	0.0936	1.986	0.0347	1.9899	0.0369
100	0.9916	0.1228	1.0553	0.0043	0.9748	0.0331	0.8865	0.1185	0.9965	0.1217
	1.9920	0.0266	1.8174	0.0752	2.0099	0.0811	1.9534	0.0283	1.9562	0.0282

Table 2: For various sample sizes, we compare Estimated value and mean square errors of multiple estimators under squared error loss for $\theta = (\mu, s)$=(1.5, 2.5).

	EM		Lindley		ABC		HMC		MH	
Size	Avg	Error	Avg	Error	Avg	Error	Avg	Error	Avg	Error
20	1.4928	0.9601	1.0491	0.2041	1.5509	0.0600	1.5541	0.8268	1.5535	0.8016
	2.4289	0.2216	2.2265	0.2879	2.5377	0.2774	2.7791	0.3434	2.8218	0.5152
50	1.5020	0.3740	1.1355	0.1343	1.5483	0.0573	1.5549	0.3530	1.5154	0.3818
	2.4712	0.0865	2.4074	0.1414	2.2659	0.2276	2.5467	0.0939	2.5585	0.1128
70	1.4970	0.2691	1.1913	0.0972	1.6193	0.0425	1.5226	0.2444	1.5	0.2809
	2.4801	0.0621	2.4363	0.1057	2.8744	0.1278	2.4362	0.0626	2.4441	0.0683
80	1.5037	0.2350	1.2192	0.0811	1.5134	0.0401	1.3916	0.2294	1.3643	0.2547
	2.4788	0.0553	2.4418	0.0951	2.3865	0.0602	2.4710	0.0544	2.4825	0.0605
100	1.4981	0.1849	1.2737	0.0538	1.4938	0.0301	1.3780	0.1871	1.3525	0.2205
	2.4848	0.0437	2.4487	0.0796	2.6498	0.0586	2.4328	0.0457	2.4381	0.0478

Table 3: 95% Asymptotic and HPD intervals using ML estimates, HMC and MH estimates for the parameters (μ, s)=(1, 2) at various sample sizes

(m,n)	Observed	Asymptotic	HMC	MH
20	μ=1	[-0.5423, 2.5307]	[-0.2879, 2.7172]	[-0.7749, 2.5659]
	s=2	[1.2135, 2.6727]	[1.5171, 3.1040]	[1.4911, 3.3946]
50	μ=1	[0.0427, 1.9605]	[0.1923, 2.1031]	[0.0008, 2.0835]
	s=2	[1.5179, 2.4361]	[1.5961, 2.5334]	[1.6004, 2.5956]
70	μ=1	[0.1842, 1.8110]	[0.2913, 1.8567]	[0.2045, 1.8402]
	s=2	[1.5945, 2.3737]	[1.5987, 2.3562]	[1.6113, 2.3688]
80	μ=1	[0.2427, 1.7631]	[0.2312, 1.6943]	[0.1718, 1.6580]
	s=2	[1.6159, 2.3501]	[1.6380, 2.3592]	[1.6356, 2.3562]
100	μ=1	[0.3242, 1.6728]	[0.2995, 1.5998]	[0.2312, 1.5617]
	s=2	[1.6609, 2.3147]	[1.6586, 2.2815]	[1.6544, 2.2890]

Table 4: 95% Asymptotic and HPD intervals using ML estimates, HMC and MH estimates for the parameters (μ, s)=(1.5, 2.5) at various sample sizes

(m,n)	Observed	Asymptotic	HMC	MH
20	μ=1.5	[-0.4278, 3.4134]	[-0.2378, 3.2617]	[-0.4279, 3.2812]
	s=2.5	[1.5169, 3.3409]	[1.8853, 3.8125]	[1.7914, 4.2509]
50	μ=1.5	[0.3033, 2.7007]	[0.3347, 2.6289]	[0.2400, 2.6841]
	s=2.5	[1.8973, 3.0451]	[1.9602, 3.1283]	[1.9335, 3.2118]
70	μ=1.5	[0.4803, 2.5137]	[0.5813, 2.5043]	[0.4940, 2.5616]
	s=2.5	[1.9932, 2.9670]	[1.9848, 2.9216]	[1.9646, 2.9502]
80	μ=1.5	[0.5534, 2.4540]	[0.4827, 2.2894]	[0.3568, 2.2658]
	s=2.5	[2.0199, 2.9377]	[2.0197 2.9185]	[2.0467, 2.9841]
100	μ=1.5	[0.6553, 2.3409]	[0.5487, 2.1649]	[0.4645, 2.2023]
	s=2.5	[2.0762, 2.8934]	[2.0460, 2.8373]	[2.0603, 2.8608]

Table 5: The breakdown lifetimes at 32 KV and 34 KV of an insulating fluid.

Data Set
(2.163, 2.738, 3.068), (5.972, 6.353, 8.150), (1.032, 1.971, 2.642), (1.336, 2.750, 3.284), (0.628, 0.964, 1.735), (2.995, 3.442, 5.066), (3.766, 5.814, 6.212), (5.434, 7.093, 7.655), (0.974, 1.839, 2.045), (4.352, 5.206, 5.988), (3.920, 4.762, 6.121), (4.065, 5.312, 7.480), (1.375, 2.195, 3.086), (0.618, 0.839, 2.217), (4.575, 6.050, 6.734), (2.821, 3.409, 5.272), (2.247, 2.990, 4.128), (7.125, 8.470, 9.044), (5.443, 6.231, 7.395), (1.766, 2.190, 2.638), (7.261, 8.325, 8.871), (7.155, 8.013, 8.352), (0.830, 1.288, 2.541), (3.590, 4.169, 4.899), (4.634, 5.780, 7.058), (5.965, 7.344, 8.019), (3.177, 3.600, 4.213), (1.027, 1.218, 3.116), (6.032, 7.746, 8.529), (6.279, 8.156, 9.435).

Table 6: MLE and Bayesian estimators of the combined model

Method	μ	s
EM	4.5224	1.4388
Lindley	4.5296	1.4260
ABC	4.5105	1.414
HMC	4.4485	1.5673
MH	4.3711	1.5857

Table 6 presents the estimates of the location parameter μ and the scale parameter s of the Logistic distribution obtained through various estimation techniques. The EM algorithm yields $\mu = 4.5224$ and $s = 1.4388$, serving as the MLE-based benchmark. Among the Bayesian methods, Lindley's approximation provides estimates close to the EM results, indicating its effectiveness under fuzzy data conditions. The ABC method also yields similar values but slightly underestimates μ and s compared to Lindley. The HMC and MH algorithms, which are more robust to complex posterior distributions, produce slightly lower estimates of μ and higher estimates of s, suggesting a broader distribution under the Bayesian framework.

In Table 7, the 95% asymptotic confidence intervals obtained via the EM algorithm and the HPD intervals from the HMC and MH algorithms are reported. The EM-based intervals for μ and s are relatively narrower, reflecting the asymptotic nature of MLE. In contrast, the HPD intervals obtained from HMC and MH are wider, indicating the Bayesian methods' ability to better capture the uncertainty

Table 7: 95% Asymptotic and HPD intervals using EM, HMC, and MH Algorithms

Method	μ	s
EM	[3.5959, 5.4490]	[1.0268, 1.8508]
HMC	[3.3749, 5.4019]	[1.1402, 2.0458]
MH	[3.3033, 5.5320]	[1.1376, 2.1509]

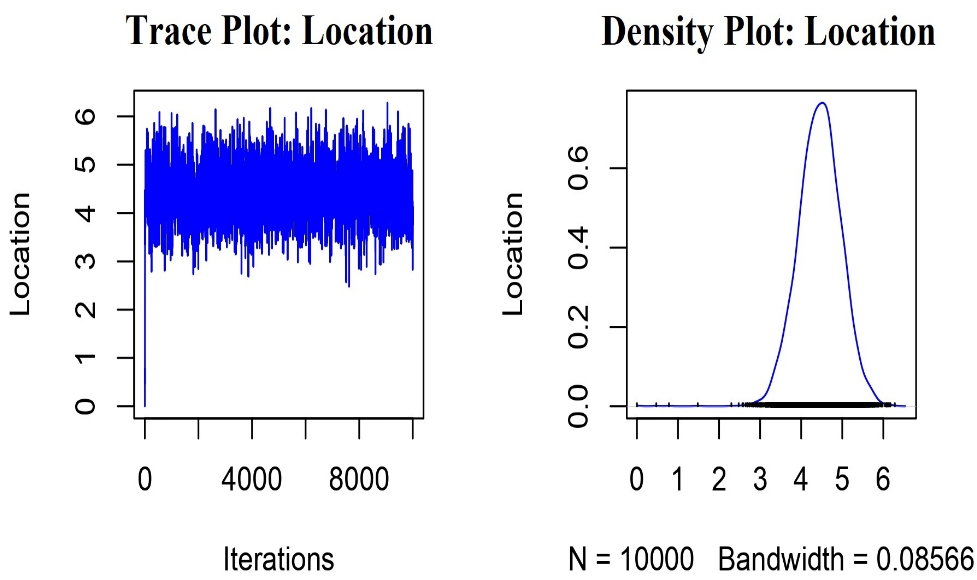

Figure 1: Illustration of traceplots of parameters.

in parameter estimates under imprecise data conditions. The MH method, in particular, produces the widest HPD intervals for both parameters, implying a more conservative estimate of uncertainty. Overall, the results highlight that while all methods provide comparable point estimates, the Bayesian approaches—especially those using MCMC—offer richer information about the uncertainty surrounding the parameters.

To further validate the Bayesian estimation results, diagnostic plots including trace plots, density plots, and autocorrelation plots for the parameters (μ, s) are shown in Figures 1, 2, and 3. The trace plots for both the location and scale parameters exhibit good mixing and stationarity, with no apparent trends, suggesting that the Markov chains have converged to the target posterior distribution. The corresponding density plots are smooth and uni-modal, indicating that the posterior distributions are well-behaved and concentrated around specific parameter values. Additionally, the autocorrelation plots for both parameters display a rapid decline in autocorrelation as the lag increases, with values dropping close to zero after a few lags. This demonstrates low serial dependence among successive samples and confirms the efficiency of the sampling process. Collectively, these diagnostic plots provide strong evidence that the HMC algorithm has performed effectively, yielding reliable posterior estimates for the location and scale parameters.

7 Conclusion

In this study, we considered parameter estimation for the logistic distribution using both classical and Bayesian approaches. Maximum Likelihood Estimation was performed via the Expectation-Maximization (EM) algorithm, while Bayesian estimation techniques included Lindley's approximation, Approximate Bayesian Computation (ABC), Hamiltonian Monte Carlo (HMC), and Metropolis-Hastings (MH).

ISSN: 0017-2294

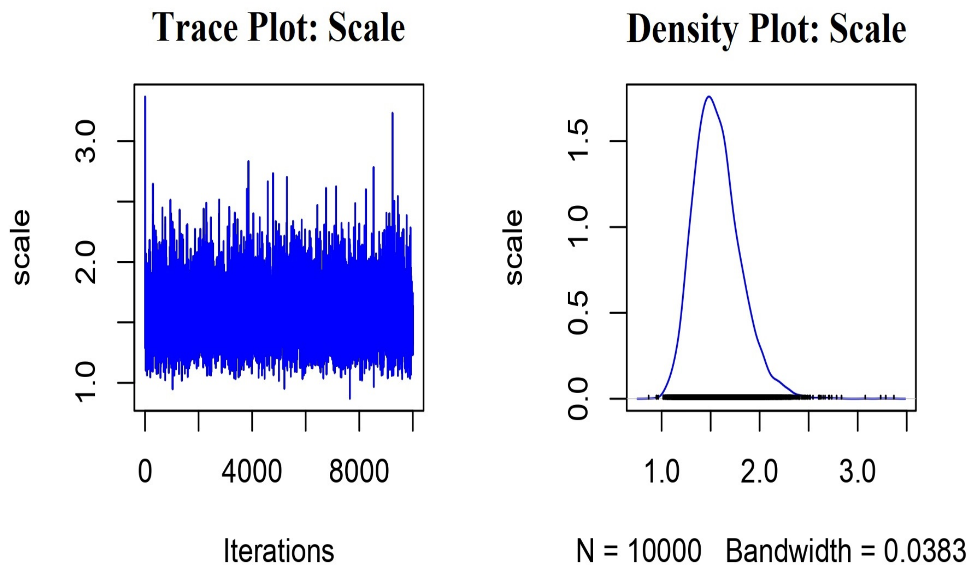

Figure 2: Illustration of Density plots of parameters.

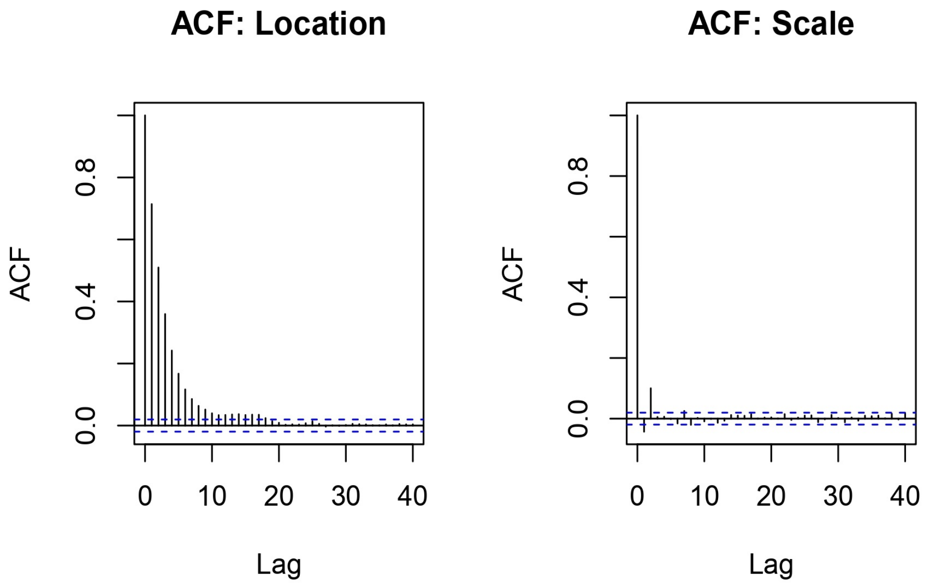

Figure 3: Illustration of Auto Correlation plots of parameters.

The Bayesian approaches yielded consistent and robust estimates, with the diagnostic results confirming convergence and accuracy across methods. The comparative analysis demonstrated that the Bayesian methods, particularly HMC and MH, are effective in capturing posterior uncertainty and offer enhanced performance, especially in complex parameter spaces.

For future work, the current methodology can be extended to models involving two or more populations, allowing for the comparison of group-specific parameters and more complex survival structures. Additionally, different types of fuzzy sets—such as trapezoidal, Gaussian, and interval-valued fuzzy sets—can be incorporated to better represent uncertainty in real-world data. Such extensions will enhance the flexibility and applicability of the proposed methods, especially in handling imprecise or ambiguous information often encountered in survival analysis and other applied statistical domains.

References

[1] Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). Continuous Univariate Distributions, Volume 2 (2nd ed.). Wiley.

[2] Kotz, S., Balakrishnan, N., & Johnson, N. L. (2004). Continuous Multivariate Distributions, Volume 1: Models and Applications. John Wiley & Sons.

[3] Balakrishnan, N. (2013). Handbook of the Logistic Distribution. CRC Press.

[4] Rashad, A., Mahmoud, M. and Yusuf, M., 2016. Bayes estimation of the logistic distribution parameters based on progressive sampling. Appl Math Inf Sci, 10(6), pp.2293-2301.

[5] Zadeh LA. Fuzzy sets. *Inf Control*. 1965;8(3):338-353.

[6] Pak A, Parham G A and Saraj M 2013 Inference for the Weibull distribution based on fuzzy data, *Rev. Colomb. Estad.*, **36**(2), 337–356. https://doaj.org/article/0e408ddcb692431ea0c60785820e877c

[7] Khoolenjani N B and Shahsanaie F 2016 Estimating the parameter of exponential distribution under Type-II censoring from fuzzy data, *J. Stat. Theory Appl.*, **15**(2), 181. https://doi.org/10.2991/jsta.2016.15.2.8

[8] Nadia H A-N and Subhi R S 2018 Two parameter Lindley distribution: Estimating the reliability function with fuzzy data, *ARPN J. Eng. Appl. Sci.*, **13**(18), 7670–7676. http://dx.doi.org/10.3923/jeasci.2018.7670.7676

[9] Al-Noor N H and Hussein L K 2018 Weighted exponential distribution: Approximate Bayes estimations with fuzzy data, *J. Al-Nahrain Univ. Sci.*, **00**(1), 174–185. https://doi.org/10.22401/anjs.00.1.24

[10] Pak A, Khoolenjani N B, Alamatsaz M H and Mahmoudi M R 2020 Bayesian method for the generalized exponential model using fuzzy data, *Int. J. Fuzzy Syst.*, **22**(4), 1243–1260. https://doi.org/10.1007/s40815-020-00843-8

[11] Roohanizadeh, Z., Baloui Jamkhaneh, E. and Deiri, E., 2022. Parameters and reliability estimation for the weibull distribution based on intuitionistic fuzzy lifetime data. *Complex & Intelligent Systems*, **8**(6), pp.4881-4896.

[12] Ashok, A. and Nagamani, N., 2025. Adaptive estimation: Fuzzy data-driven gamma distribution via Bayesian and maximum likelihood approaches. *AIMS Mathematics*, **10**(1), pp.438-459.

[13] Huber PJ. Robust Statistics. Hoboken: Wiley; 2009.

[14] Beaumont M, Zhang W, Balding DJ. Approximate Bayesian computation in population genetics. *Genetics*. 2002;162(4):2025–2035.

[15] Neal RM. MCMC using Hamiltonian dynamics. In: Brooks S, Gelman A, Jones GL, Meng XL, editors. Handbook of Markov Chain Monte Carlo. Boca Raton: CRC Press; 2011. p. 113–162.

[16] Robert CP, Casella G, Chopin N. Monte Carlo Statistical Methods. 3rd ed. New York: Springer; 2013.

[17] Betancourt M. A conceptual introduction to Hamiltonian Monte Carlo. *arXiv Preprint*. 2017. arXiv:1701.02434. https://doi.org/10.48550/arXiv.1701.02434

[18] Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*. 1970;57(1):97–109. https://doi.org/10.1093/biomet/57.1.97

[19] Chib S. Marginal likelihood from the Metropolis-Hastings output. *J Am Stat Assoc*. 1995;90(432):1313–1321. https://doi.org/10.1080/01621459.1995.10476635

[20] Zadeh L. Probability measures of fuzzy events. *Math Anal Appl*. 1968;23(2):421-427. https://doi.org/10.1016/0022-247x(68)90078-4

[21] Dempster A P, Laird N M and Rubin D B 1977 Maximum likelihood from incomplete data via the EM algorithm, *J. R. Stat. Soc. Ser. B*, **39**(1), 1–22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

[22] Nadarajah S (2004) Information matrix for logistic distributions. Math Comput Model 40(9–10):953–958

[23] Lindley DV. Approximate bayes method. *Trab Estud Estad*. 1980;31:223-237. https://doi.org/10.1007/BF02888353