Statistics and Applications {ISSN 2454-7395 (online)}
Volume 22, No. 2, 2024 (New Series), pp 231–258
http://www.ssca.org.in/journal

Bayesian Analysis of Exponentiated Exponential Power Distribution under Hamiltonian Monte Carlo Method

Laxmi Prasad Sapkota and Vijay Kumar
Department of Mathematics and Statistics
Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur (UP)-273009, India

Received: 10 November 2022; Revised: 01 November 2023; Accepted: 09 January 2024

Abstract

In this article, we present a new univariate probability distribution containing three parameters named exponentiated exponential power distribution. The density function and failure rate function of this new distribution accommodate broad varieties of shapes. Some mathematical and statistical properties of the proposed model are provided. Also, we have performed a full Bayesian analysis of the proposed model. Using Stan software whose Markov chain Monte Carlo (MCMC) techniques are based on a No-U-Turn sampler (NUTS) which is an adaptive variant of Hamiltonian Monte Carlo (HMC); a more robust and efficient sampler. We have presented the numerical as well as graphical analysis of the EEP model and found that all chains are well mixed and conversed. Further, we have estimated the parameters of the model and performed posterior predictive checks, and found that the underlying model can be used to generate reliable samples. The developed techniques are applied to a real data set, thus we can apply for full Bayesian analysis for the proposed model using these Bayesian techniques. Hence it is expected that the EEP model will be a choice in the fields of the theory of probability, applied statistics, bayesian inferences, and survival analysis.

Key words: Exponential power distribution; Posterior distribution; Bayesian analysis; HMC.

AMS Subject Classifications: 62F10, 62F15, 62E10

1. Introduction

Lifetime distributions are typically adapted to study the length of the lifetime of parts of a system, or a device, and usually, we conduct the survival and reliability analysis. Generally, lifetime models are extensively employed in fields like bioscience, medicine, demography, engineering, biology, insurance, etc. Several continuous probability distributions like exponential, Weibull, Cauchy, gamma, etc. are generally found in the literature of probability and applied statistics to study real-life data. Since the last decade, most scientists have been paying attention to the family of exponential models for their capability to model real-life data, and it has been observed that this model has performed well in several applications because of the bearing of closed-type solutions to several survival analyses. It will simply be

Corresponding Author: Laxmi Prasad Sapkota
Email: laxmisapkota75@gmail.com

even underneath the assumption of a constant failure rate however in practice, the failure rates don't seem to be continuously constant. Hence, the chaotic use of the exponential model appears to be insufficient and inappropriate. In this paper, we have presented a new model by extending the exponential power (EP) distribution defined by (Smith and Bain, 1975). The shape of the hazard function of this distribution depends on the value of the shape parameter α. For $\alpha \geq 1$, the hazard function is increasing and for $\alpha < 1$, it has a U-shape and exponentially increasing (towards the right) hazard function (Chen, 1999; Barriga *et al.*, 2011). The distribution and density function of EP distribution having parameters α and λ are as follows

$$F_{EP}(x) = 1 - \exp\left\{1 - e^{(\lambda\, x)^{\alpha}}\right\};\ (\alpha,\ \lambda) > 0, \quad x \geq 0. \tag{1}$$

$$f_{EP}(x) = \alpha\, \lambda^{\alpha}\, x^{\alpha-1}\, e^{(\lambda\, x)^{\alpha}} \exp\left\{1 - e^{(\lambda\, x)^{\alpha}}\right\} \quad ;\ (\alpha,\ \lambda) > 0, \quad x \geq 0. \tag{2}$$

Using the EP distribution (Barriga *et al.*, 2011) has defined a flexible lifetime model named complementary exponential power (CEP) distribution. The CDF of CEP is

$$F\left(t; \alpha, \beta, \theta\right) = \left[1 - \exp\left(1 - \exp\left\{\left(\frac{t}{\alpha}\right)^{\beta}\right\}\right)\right]^{\theta};\ t > 0.$$

To define the proposed new lifetime distribution we have used the technique presented by (Gupta and Kundu, 1999). They defined the generalized exponential (GE) distribution by inserting a shape parameter to the exponential distribution and it is superior to an exponential distribution, having decreasing and increasing failure rate hazard function. The cumulative density function (CDF) and its probability density function (PDF) of GE distribution are

$$F_{GE}\left(x;\ \alpha,\ \lambda\right) = \left\{1 - e^{-\lambda\, x}\right\}^{\alpha}; (\alpha,\ \lambda) > 0,\ x > 0.$$

$$f_{GE}\left(x;\ \alpha,\ \lambda\right) = \alpha\ \lambda\ e^{-\lambda\, x}\left\{1 - e^{-\lambda\, x}\right\}^{\alpha-1} (\alpha,\ \lambda) > 0,\ x > 0.$$

Using the same technique (Mudholkar and Srivastava, 1993) developed a three-parameter exponentiated Weibull (EW) distribution by inserting one additional shape parameter to the Weibull distribution. The CDF of EW is

$$F\left(x\right) = \left\{1 - \exp\left(-\alpha x^{\beta}\right)\right\}^{\lambda}; x > 0.$$

Another extension of exponential distribution has been developed by (Nadarajah and Haghighi, 2011) which can be taken alternative to exponentiated exponential, Weibull, and gamma distributions. Similarly exponentiated Chen (EC) distribution was defined by (Chaubey and Zhang, 2015) with either unimodel or decreasing density shape and decreasing or bathtub hazard shape. Dey *et al.* (2017) have redefined the exponentiated Chen distribution and extensively investigated the properties, and estimated the parameters using different methods. The CDF of EC is

$$F_{EC}\left(x; \alpha, \delta, \lambda\right) = \left\{1 - \exp\left[\lambda\left(1 - e^{x^{\delta}}\right)\right]\right\}^{\alpha}; (\alpha, \delta, \lambda) > 0,\ x > 0.$$

The exponentiated exponential Poisson (EEP) was introduced by (Ristić and Nadarajah, 2014) with a flexible hazard function. Using the same technique (Ashour and Eltehiwy, 2015) has defined the exponentiated power Lindley having CDF as

$$F\left(t; \alpha, \beta, \theta\right) = \left[1 - \left(1 + \frac{\theta t^{\beta}}{\theta + 1}\right) e^{-\theta t^{\beta}}\right]^{\alpha};\ t > 0.$$

Another extension of exponential distribution has been defined by (Almarashi *et al.*, 2019) whose hazard function can have a variety of shapes. Similarly, EP distribution is also used by (Joshi *et al.*, 2020) and generated a flexible model named logistic-exponential power distribution that can have decreasing or increasing or bathtub-shaped hazard function. Sapkota (2020) has defined exponentiated exponential logistic distribution and introduced a flexible hazard function. Hence we are motivated to generalize the EP distribution to get a versatile model by inserting only one shape parameter.

Further in this study, we have analyzed the suggested new model under the Bayesian approach. It is a fundamental framework for reasoning about uncertainty in statistical modeling and decision-making. It is a flexible and coherent approach that can handle various statistical problems, ranging from simple parameter estimation to complex hierarchical modeling and machine learning tasks. It provides a principled way to incorporate prior knowledge, update beliefs based on data, and quantify uncertainty in the results (Lambert, 2018; McElreath, 2020). Under this approach, we have used the HMC algorithm which is a powerful MCMC algorithm used to sample from complex probability distributions, especially in Bayesian statistics and machine learning. Unlike traditional MCMC methods, which often suffer from slow exploration of high-dimensional spaces, HMC leverages the concept of Hamiltonian dynamics from physics to efficiently explore the target distribution. HMC treats the probability distribution as a potential energy surface, and the Markov chain as a particle moving through this surface (Neal, 2011; Sapkota, 2022). HMC is more efficient than traditional MCMC methods because it generates less correlated samples and requires fewer evaluations of the target distribution's gradient (Carpenter *et al.*, 2017).

The remaining sections of this article are structured as follows: In the second section, we introduce the new distribution and examine its statistical characteristics. Moving on to the third section, we present some statistical properties of the EEP model. Section 4 is dedicated to discussing the application of the suggested model under the classical approach. Under the Bayesian approach, we formulate the proposed model, and its posterior analysis is presented in sections 5, 6, and 7, respectively. In section 8, we showcase the compatibility of the model, while section 9 delves into concluding remarks.

2. Exponentiated exponential power (EEP) distribution

Let $X \sim EEP\,(\alpha, \lambda, \theta)$ then the CDF of EEP distribution can be obtained by using Equation (1) and written as

$$F(x) = \left[1 - \exp\left\{1 - \exp\left(\lambda x^{\alpha}\right)\right\}\right]^{\theta}; x > 0, (\alpha, \lambda, \theta) > 0. \tag{3}$$

The PDF of EEP is obtained using Equation (2) as

$$f(x) = \alpha\lambda\theta x^{\alpha-1} \exp\left\{1 + \lambda x^{\alpha} - e^{\lambda x^{\alpha}}\right\} \left[1 - \exp\left(1 - e^{\lambda x^{\alpha}}\right)\right]^{\theta-1}; x > 0. \tag{4}$$

Different shapes of PDF curves of $EEP\,(\alpha, \lambda, \theta)$ distribution are presented in Figure 1.

2.1. Some special cases

- When $\theta = 1$, obviously EEP distribution reduces to EP distribution (Smith and Bain, 1975).

- When $\lambda = 1$, the EEP distribution reduces to EC distribution (Chaubey and Zhang, 2015).
- When $\lambda = 1$ and $\theta = 1$, the EEP distribution reduces to Chen distribution (Chen, 2000).
- If $\lambda = \frac{1}{\alpha^\beta}$, then the EEP distribution reduces to CEP distribution (Barriga *et al.*, 2011).

2.2. Survival function of EEP distribution

The survival function for the time t is

$$S(t) = \overline{F}(t) = 1 - [1 - \exp\{1 - \exp(\lambda t^\alpha)\}]^\theta ; t > 0. \tag{5}$$

2.3. The hazard function of EEP distribution

Suppose t be the time of an item or component or an event that will survive and we would like to compute the probability of failing at time $t + \Delta t$ then the hazard function can be defined as

$$h(t) = \alpha\lambda\theta t^{\alpha-1} \exp\left\{1 + \lambda t^\alpha - e^{\lambda t^\alpha}\right\} \frac{\left[1 - \exp\left(1 - e^{\lambda t^\alpha}\right)\right]^{\theta-1}}{1 - [1 - \exp\{1 - \exp(\lambda t^\alpha)\}]^\theta} ; t > 0. \tag{6}$$

2.4. Reverse hazard function of EEP distribution

The reverse hazard function of EEP distribution is

$$\begin{aligned} P_{rev}(x) &= \frac{f(x; \alpha, \lambda, \theta)}{F(x; \alpha, \lambda, \theta)} \\ &= \alpha\lambda\theta x^{\alpha-1} \exp\left\{1 + \lambda x^\alpha - e^{\lambda x^\alpha}\right\} \left[1 - \exp\left(1 - e^{\lambda x^\alpha}\right)\right]^{\theta-1}, x > 0. \end{aligned}$$

2.5. Quantile function

Suppose X be a non-negative continuous random variable with a CDF $F_X(x)$ and $U \in (0, 1)$, then the u^{th} quantile of X is,

$$Q(u) = \left[\frac{1}{\lambda} \ln\left\{1 - \ln\left(1 - u^{1/\theta}\right)\right\}\right]^{1/\alpha} ; 0 < u < 1. \tag{7}$$

We can also calculate the median through Equation (7) as

$$Median = \left[\frac{1}{\lambda} \ln\left\{1 - \ln\left(1 - 2^{-1/\theta}\right)\right\}\right]^{1/\alpha}$$

To generate the random numbers for EEP distribution we can use

$$x = \left[\frac{1}{\lambda} \ln\left\{1 - \ln\left(1 - v^{1/\theta}\right)\right\}\right]^{1/\alpha} ; 0 < v < 1. \tag{8}$$

2.6. Skewness of EEP distribution

Using quartiles, Bowley's coefficient of skewness can be computed as,

$$S_k\left(B\right)=\frac{Q\left(1/4\right)-2Q\left(1/2\right)+Q\left(3/4\right)}{Q\left(0.75\right)-Q\left(0.25\right)}.$$

2.7. Kurtosis of EEP distribution

The coefficient of kurtosis using octiles (Moors, 1988) is

$$K_u\left(M\right)=\frac{Q\left(0.875\right)+Q\left(0.375\right)-Q\left(0.625\right)-Q\left(0.125\right)}{Q\left(0.75\right)-Q\left(0.25\right)}.$$

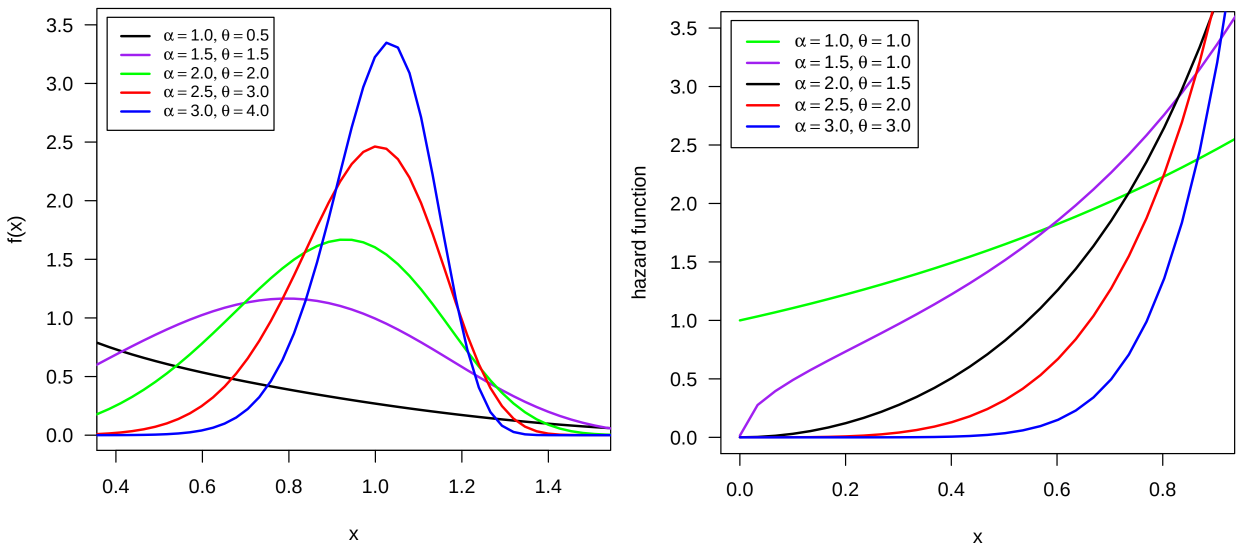

Figure 1: Graphs of PDF (left panel) and HRF (right panel) for different values of α and θ and fixed $\lambda = 1$

3. Some statistical properties of EEP distribution

3.1. Moments

The K^{th} moment about origin using the quantile function see for details (Balakrishnan and Cohen, 2014) and (Dey *et al.*, 2017) can be computed as

$$\mu_k^{raw}=E(X^r)=\int_0^\infty x^k f\left(x\right)dx=\int_0^1\left[Q\left(v\right)\right]^k dv$$

$$=\int_0^1\lambda^{-k/\alpha}\left[\log\left\{1-\log\left(1-v^{1/\theta}\right)\right\}\right]^{k/\alpha}dv \tag{9}$$

$$=\lambda^{-k/\alpha}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}W_p(k)W_q\left(\frac{k}{\alpha}+p\right)(-1)^{\frac{2k}{\alpha}+p}\int_0^1 v^{\frac{1}{\theta}\left(\frac{k}{\alpha}+p\right)+q}dv.$$

$$\therefore \quad \mu_k^{raw} = \lambda^{-k/\alpha} \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} W_p(k) W_q\left(\frac{k}{\alpha}+p\right)(-1)^{\frac{2k}{\alpha}+p}\left(C_{pq}\right)^{-1}. \tag{10}$$

here $W_p(k)$ is the coefficient of $\log\left(1-v^{1/\theta}\right)$ after the expansion of $\left[\log\left\{1-\log\left(1-v^{1/\theta}\right)\right\}\right]^{k/\alpha}$, $W_q\left(\frac{k}{\alpha}+p\right)$ is the coefficient of v^q after the expansion of $\left[\log\left(1-v^{1/\theta}\right)\right]^{k/\alpha}$ and $C_{pq} = \frac{1}{\theta}\left(\frac{k}{\alpha}+p\right)+q+1$.

Remark: $Mean(\mu_1)$ and $Variance(\mu_2)$ of EEP distribution can be computed as

$$\mu_1 = \lambda^{-1/\alpha} \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} W_p(1) W_q\left(\frac{1}{\alpha}+p\right)(-1)^{\frac{2}{\alpha}+p}\left(\frac{1}{\theta}\left(\frac{1}{\alpha}+p\right)+q+1\right)^{-1}.$$

$$\mu_2 = \lambda^{-2/\alpha} \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} W_p(2) W_q\left(\frac{2}{\alpha}+p\right)(-1)^{\frac{4}{\alpha}+p}\left(\frac{1}{\theta}\left(\frac{2}{\alpha}+p\right)+q+1\right)^{-1} - \left(\mu_1\right)^2.$$

3.2. Moment generating function (MGF)

The expression for MGF of EEP distribution can be obtained by using the Equation (10) as

$$M_X(t) = \sum_{s=0}^{\infty} \frac{t^s}{s!}\mu_k^{raw} = \lambda^{-k/\alpha} \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \sum_{s=0}^{\infty} W_p(k) W_q\left(\frac{k}{\alpha}+p\right)(-1)^{\frac{2k}{\alpha}+p}\frac{t^s}{s!}\left(C_{pq}\right)^{-1}. \tag{11}$$

3.3. Conditional moments (CM)

Let Y be a random variable from the EEP distribution, and then the CM for Y can be expressed as

$$\begin{aligned} E\left(Y^k/Y>y\right) &= \frac{1}{S(y)}\int_x^{\infty} y^k f(y)\,dy \\ &= \frac{1}{S(y)}\int_v^{1} Q^k(v)dv \\ &= \frac{1}{S(y)}\lambda^{-k/\alpha} \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \sum_{s=0}^{\infty} W_p(k) W_q\left(\frac{k}{\alpha}+p\right)(-1)^{\frac{2k}{\alpha}+p}\frac{1-F(y)^{C_{pq}}}{C_{pq}}, \end{aligned} \tag{12}$$

here $S(y)$ and $F(y)$ are survival functions and CDF of EEC distribution.

3.4. Average residual life (ARL) function

ARL function of a component using quantile function can be calculated by

$$\mu_{ARL}(x) = \frac{1}{S(x)}\lambda^{-1/\alpha} \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} W_p(1) W_q\left(\frac{1}{\alpha}+p\right)(-1)^{\frac{2}{\alpha}+p}\frac{1-F^{\left(\frac{1}{\theta}\left(\frac{1}{\alpha}+p\right)+q+1\right)}(x)}{\left(\frac{1}{\theta}\left(\frac{1}{\alpha}+p\right)+q+1\right)} - x.$$

3.5. Mean deviation (MD)

Let μ and $F(.)$ denote the mean and CDF of EEP distribution then MD can be expressed as

$$\mu_{MD\_Mean}\left(x\right) = 2\mu F\left(\mu\right) - 2\mu + 2\int\limits_{\mu}^{1} Q^{k}\left(v\right)dv. \tag{13}$$

$$= 2\mu F\left(\mu\right) - 2\mu + 2\lambda^{-1/\alpha}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty} W_p(1) W_q\left(\frac{1}{\alpha}+p\right)$$

$$\times \left(-1\right)^{\frac{2}{\alpha}+p}\frac{1 - F^{\left(\frac{1}{\theta}\left(\frac{1}{\alpha}+p\right)+q+1\right)}\left(\mu\right)}{\left(\frac{1}{\theta}\left(\frac{1}{\alpha}+p\right)+q+1\right)}.$$

4. Classical analysis of the proposed model

4.1. Parameter estimation

In this subsection, we have used the maximum likelihood estimation (MLE) method which is the most frequently used method for the point and interval estimation of the parameters of the model. Let $\underline{x} = (x_1, \ldots, x_n)$ be a non-negative observed sample of size 'n' following the $EEP\left(\alpha, \lambda, \theta\right)$ then we can define the likelihood function for the parameter vector $\chi = \left(\alpha, \lambda, \theta\ \right)^{T}$ as

$$L(\chi) = \left(\alpha\lambda\theta\right)^{n}\prod_{i=1}^{n} x_i^{\alpha-1}\exp\left\{1 + \lambda x_i^{\alpha} - e^{\lambda x_i^{\alpha}}\right\}\left[1 - \exp\left(1 - \mathrm{e}^{\lambda x_i^{\alpha}}\right)\right]^{\theta-1}. \tag{14}$$

Taking the logarithm to (14) we get the log-likelihood function as

$$\ell\left(\chi\right) = n\ln\left(\alpha\lambda\theta\right) + \left(\alpha-1\right)\sum_{i=1}^{n}\ln x_i + n + \lambda\sum_{i=1}^{n} x_i^{\alpha} - \sum_{i=1}^{n} e^{\lambda x_i^{\alpha}} + \left(\theta-1\right)\sum_{i=1}^{n}\ln\left[1 - K\left(x_i\right)\right]. \tag{15}$$

Differentiating (15) with respect to parameters $\alpha,\ \lambda\ and\ \theta$, we get

$$\frac{\partial\ell}{\partial\alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n}\ln x_i + \lambda\sum_{i=1}^{n}\left[x_i^{\alpha}\ln x_i\left\{1 - \ln x_i\left\{1 - (\theta-1)K\left(x_i\right)\left\{1 - K\left(x_i\right)\right\}^{-1}\right\}\right\}\right],$$

$$\frac{\partial\ell}{\partial\lambda} = \frac{n}{\lambda} + \sum_{i=1}^{n} x_i^{\alpha} - \sum_{i=1}^{n} x_i^{\alpha} e^{\lambda x_i^{\alpha}}\left\{1 - \left(\theta-1\right)\left\{1 - K\left(x_i\right)\right\}^{-1} K\left(x_i\right)\right\},$$

$$\frac{\partial\ell}{\partial\theta} = \frac{n}{\theta} + \sum_{i=1}^{n}\ln\left[1 - K\left(x_i\right)\right],$$

where $K\left(x_i\right) = \exp\left(1 - e^{\lambda x_i^{\alpha}}\right)$. Manually it is quite difficult to solve these equations for the parameters $\alpha,\ \lambda\ and\ \theta$. Using the appropriate software like R, Python, Matlab, etc. we can solve them manually. Let $\chi = \left(\alpha, \lambda, \theta\ \right)^{T}$ be the parameter vector and MLEs of χ is $\hat{\chi} = (\hat{\alpha}, \hat{\lambda}, \hat{\theta})$, then $(\hat{\chi} - \chi) \rightarrow N_3\left[0, \left(M\left(\chi\right)\right)^{-1}\right]$ distributed as the normal distribution, here

$M(\chi)$ is known as Fisher's information matrix computed as,

$$M = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix},$$

where,

$$M_{11} = \frac{\partial^2 l}{\partial \alpha^2},\ \mathrm{M}_{12} = \frac{\partial^2 l}{\partial \alpha \partial \lambda},\ \mathrm{M}_{13} = \frac{\partial^2 l}{\partial \alpha \partial \theta},$$
$$M_{21} = \frac{\partial^2 l}{\partial \lambda \partial \alpha},\ \mathrm{M}_{22} = \frac{\partial^2 l}{\partial \lambda^2},\ \mathrm{M}_{23} = \frac{\partial^2 l}{\partial \theta \partial \lambda},$$
$$M_{31} = \frac{\partial^2 l}{\partial \lambda \partial \alpha},\ \mathrm{M}_{32} = \frac{\partial^2 l}{\partial \theta \partial \lambda},\ \mathrm{M}_{33} = \frac{\partial^2 l}{\partial \theta^2},$$

which can be calculated as,

$$\frac{\partial^2 \ell}{\partial \alpha^2} = -\frac{n}{\alpha^2} + \lambda \sum_{i=1}^{n} \left\{ x_i^\alpha (\ln x_i)^2 \right\} - (\theta - 1) \sum_{i=1}^{n} [1 + \lambda \ln x_i + \alpha x_i] \left(\lambda x_i^\alpha e^{\lambda x_i^\alpha} \ln x_i \right)^2$$
$$K(x_i) \{1 - K(x_i)\}^{-1} \left[1 + K(x_i) \{1 - K(x_i)\}^{-1} \right].$$

$$\frac{\partial^2 \ell}{\partial \lambda^2} = -\frac{n}{\lambda^2} - \sum_{i=1}^{n} x_i^{2\alpha} e^{\lambda x_i^\alpha} (\theta - 1) K(x_i) \left\{ \{1 - K(x_i)\}^{-1} K(x_i) - \{1 - K(x_i)\}^{-2} K(x_i) e^{\lambda x_i^\alpha} \right\}.$$

$$\frac{\partial^2 \ell}{\partial \theta^2} = -\frac{n}{\theta^2}.$$

$$\frac{\partial^2 \ell}{\partial \theta \partial \alpha} = \lambda \sum_{i=1}^{n} \left[x_i^\alpha e^{\lambda x_i^\alpha} \ln(x_i) K(x_i) \{1 - K(x_i)\}^{-1} \right].$$

$$\frac{\partial^2 \ell}{\partial \theta \partial \lambda} = \sum_{i=1}^{n} \left[x_i^\alpha e^{\lambda x_i^\alpha} \{1 - K(x_i)\}^{-1} K(x_i) \right].$$

$$\frac{\partial^2 \ell}{\partial \alpha \partial \lambda} = -\sum_{i=1}^{n} x_i^\alpha \ln x_i e^{\lambda x_i^\alpha} \left\{ (1 + \lambda x_i^\alpha) \left\{ 1 - (\theta - 1) \{1 - K(x_i)\}^{-1} K(x_i) \right\} + Z_i \right\},$$

where

$$Z_i = (\theta - 1) \lambda x_i^\alpha e^{\lambda x_i^\alpha} K(x_i) \{1 - K(x_i)\}^{-1} \{1 + K(x_i)\}.$$

Now the observed information matrix can be calculated through algorithms like Newton-Raphson and can be computed as,

$$[M(\chi)]^{-1} = \begin{pmatrix} V(\hat{\alpha}) & \mathrm{cov}(\hat{\alpha}, \hat{\lambda}) & \mathrm{cov}(\hat{\alpha}, \hat{\theta}) \\ \mathrm{cov}(\hat{\lambda}, \hat{\alpha}) & V(\hat{\lambda}) & \mathrm{cov}(\hat{\lambda}, \hat{\theta}) \\ \mathrm{cov}(\hat{\theta}, \hat{\alpha}) & \mathrm{cov}(\hat{\theta}, \hat{\lambda}) & V(\hat{\theta}) \end{pmatrix}.$$

Hence, estimated $100(1-\delta)\%$ CI for α, λ *and* θ can be created as, $\hat{\alpha} \pm z_{\delta/2} SE(\hat{\alpha})$, $\hat{\lambda} \pm z_{\delta/2} SE(\hat{\lambda})$, and $\hat{\theta} \pm z_{\delta/2} SE(\hat{\theta})$.

4.2. Illustration with real dataset

In this section, we examine a real dataset previously utilized by various researchers to showcase the capabilities and applicability of the EEP distribution. Additionally, we present the EEP distribution alongside several competing distributions, listed below, to offer a comprehensive comparison.

- Exponential power (EP) distribution by (Smith and Bain, 1975).
- Power Lindley distribution (PL) by (Ghitany *et al.*, 2015).
- Generalized Rayleigh (GR) distribution by (Kundu and Raqab, 2005).
- Marshall-Olkin Extended Exponential (MOEE) distribution by (Marshall and Olkin, 1997).

To compare the proposed distribution with the distributions as mentioned above we have computed the Bayesian information criterion (BIC), Akaike information criterion (AIC), negative log-likelihood (-LL), Hannan-Quinn information criterion (HQIC), and Corrected Akaike Information criterion (CAIC) statistic. These statistics are obtained by using the following expressions

$$AIC = -2\ell(\hat{\chi}) + 2d.$$
$$BIC = -2\ell(\hat{\chi}) + d\log(n).$$
$$CAIC = \frac{2d(d+1)}{n-d-1} + AIC.$$
$$HQIC = -2\ell(\hat{\chi}) + 2d\log[\log(n)].$$

Here $\hat{\chi}$ denotes estimated parameter space, n is the size of the sample and d is the number of parameters of the model under study. In addition, to judge the goodness-of-fit of EEP distribution Cramer-Von Mises (A^2), Kolmogorov-Smirnov (KS), and Anderson-Darling (W) statistics are presented and calculated as

$$KS = \max_{1\le j\le n}\left(d_j - \frac{j-1}{n}, \frac{j}{n} - d_j\right).$$
$$W = -n - \frac{1}{n}\sum_{j=1}^{n}(2j-1)\left[\ln d_j + \ln\left(1 - d_{n+1-j}\right)\right].$$
$$A^2 = \frac{1}{12n} + \sum_{j=1}^{n}\left[\frac{2j-1}{2n} - d_j\right]^2.$$

where the d_j's are the ordered observations, and $d_j = CDF\left(x_j\right)$.

4.2.1. Dataset

The dataset represents the waiting time (in minutes) of 100 clients (Ghitany *et al.*, 2008) before the client received service in a bank. The data set is,
"0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9, 3.1, 3.2, 3.3, 3.5, 3.6, 4.0, 4.1, 4.2, 4.2, 4.3,

4.3, 4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9, 5.0, 5.3, 5.5, 5.7, 5.7, 6.1, 6.2, 6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1, 7.1, 7.1, 7.4, 7.6, 7.7, 8.0, 8.2, 8.6, 8.6, 8.6, 8.8, 8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11.0, 11.0, 11.1, 11.2, 11.2, 11.5, 11.9, 12.4, 12.5, 12.9, 13.0, 13.1, 13.3, 13.6, 13.7, 13.9, 14.1, 15.4, 15.4, 17.3, 17.3, 18.1, 18.2, 18.4, 18.9, 19.0, 19.9, 20.6, 21.3, 21.4, 21.9, 23.0, 27.0, 31.6, 33.1, 38.5”

4.2.2. Exploratory study of the dataset

The main aim of the exploratory data analysis is to explore more information about the data. The latest statistical tools for data analysis incorporate exploratory data analysis. The descriptive statistics of the dataset are presented in Table 1. The basic exploratory

Table 1: Summary statistics of the dataset

Minimum	Q1	Median	Mean	Q3	Maximum	Skewness	Kurtosis
0.800	4.675	8.100	9.877	13.025	38.500	1.451	2.430

data analysis technique is applied to study the data and results are displayed respectively in Table 1. Efficient modeling requires an excellent understanding of the properties of different types of models. The parameters of the proposed model are estimated using the maximum likelihood (ML) estimation method. To evaluate the validity of the model, we calculate the Kolmogorov-Smirnov (KS) distance between the fitted distribution function and empirical distribution function where the parameters are estimated by the ML estimation method. The probability–probability (PP) plot and quantile-quantile (QQ) plot are used to check the suitability of the proposed model.

4.2.3. Computation of MLE

The MLEs of EEP distribution are calculated with the help of R programming software using `maxLik()` an R package developed by (Henningsen and Toomet, 2011) and they are uniquely determined (see Figure 2). In Table 2, the MLEs with 95% confidence interval (CI) and standard errors (SE) are presented.

Table 2: MLE and SE for α, λ, and θ of EEP distribution

Parameter	MLE	SE	95% CI
α	0.3407	0.0590	(0.2252, 0.4562)
λ	0.6068	0.1259	(0.3600, 0.8535)
θ	7.6150	2.6998	(2.3234, 12.9065)

4.2.4. Model validation

To check the validity of the proposed model we performed the Kolmogorov-Smirnov (KS) test and we found that $KS = 0.0358$ and p-value $= 0.9995$ which indicates that the proposed model can fit the data well. Further, we have presented the K-S plot (right panel) and quantile-quantile (Q-Q) plot (left panel) to evaluate the validity of the model in Figure 3 and it also verifies the validity of the model.

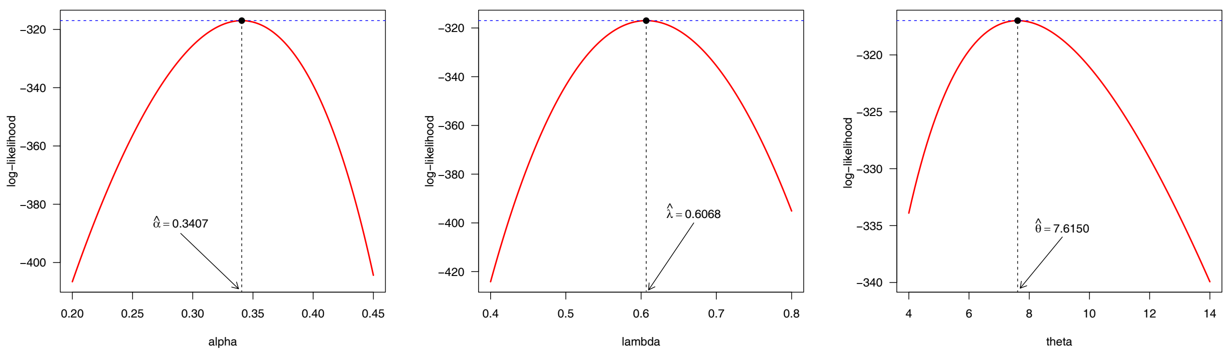

Figure 2: The graph of profile log-likelihood for α, λ, and θ

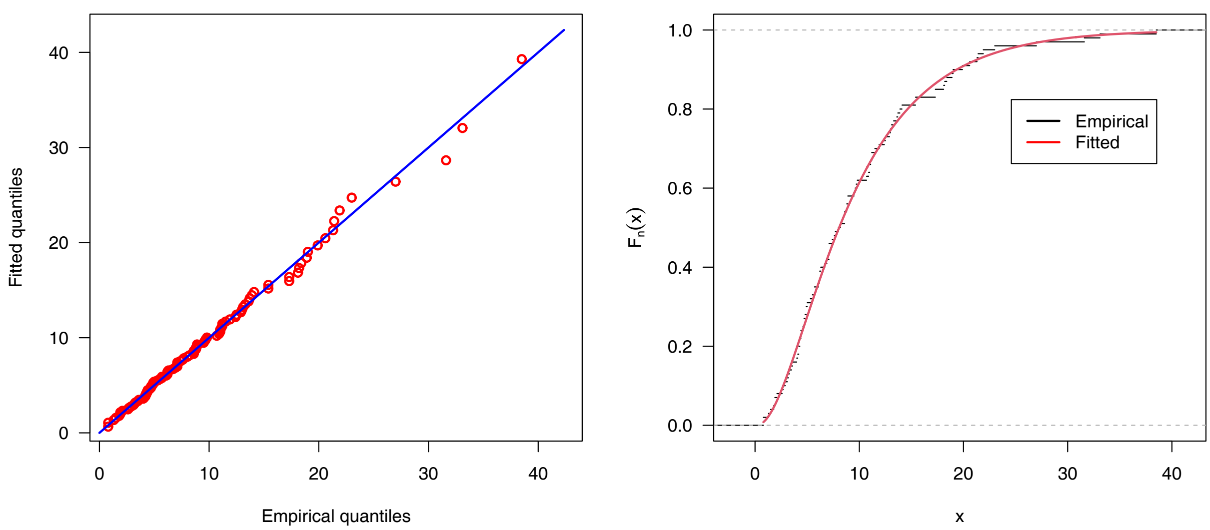

Figure 3: The graph of Q-Q (left side) and K-S (right side) of the EEP distribution

4.2.5. Model selection

To select a good model we have computed the AIC, CAIC, BIC, and HQIC for the proposed model as well as the other four models taken for comparison, and it is observed that the EEP distribution has the highest value of LL and lowest values of AIC, CAIC, BIC, and HQIC, hence we confirmed that the proposed distribution is better than the competing distribution (Table 3) for more detail see (Lambert, 2018). To evaluate the fit attained by EEP distribution among challenging distributions, the Anderson-Darling (W), Kolmogorov-Smirnov (KS), and the Cramer-Von Mises (A^2) tests are conducted and results are reported in Table 4. We have seen that the EEP distribution gets the smallest test statistic with a higher p-value which indicates the EEP distribution gets consistently better fit than those models taken under consideration. We have also presented the graphs to evaluate the goodness-of-fit of EEP distribution with distributions that are taken for comparison in Figure 4 (left panel) and the Kaplan-Meier (KM) estimate which is used to estimate the reliability function of EEP distribution in Figure 4 (right panel) and exhibits good fit.

Table 3: AIC, CAIC, BIC, and HQIC, Log-likelihood (LL) statistics

Model	AIC	CAIC	BIC	HQIC	LL
EEP	639.9793	640.2293	647.7948	643.1420	-316.9897
PL	640.6372	640.7609	645.8475	642.7460	-318.3186
MOEE	645.4241	645.5453	650.6344	647.5330	-320.7120
GR	647.0364	647.1601	652.2467	649.1450	-321.5182
EP	654.0395	654.1607	659.2499	656.1480	-325.0198

Table 4: Value of W, KS and A^2 statistics with p-value

Model	W(*p-value*)	KS(*p-value*)	A^2(*p-value*)
EEP	0.0173(0.9990)	0.0358(0.9995)	0.1274(0.9997)
PL	0.0458(0.9025)	0.0520(0.9498)	0.3028(0.9359)
MOEE	0.0760(0.7164)	0.0596(0.8690)	0.6351(0.6150)
GR	0.2043(0.2595)	0.0945(0.3337)	1.0911(0.3126)
EP	0.2549(0.1822)	0.0930(0.3532)	1.6490(0.1447)

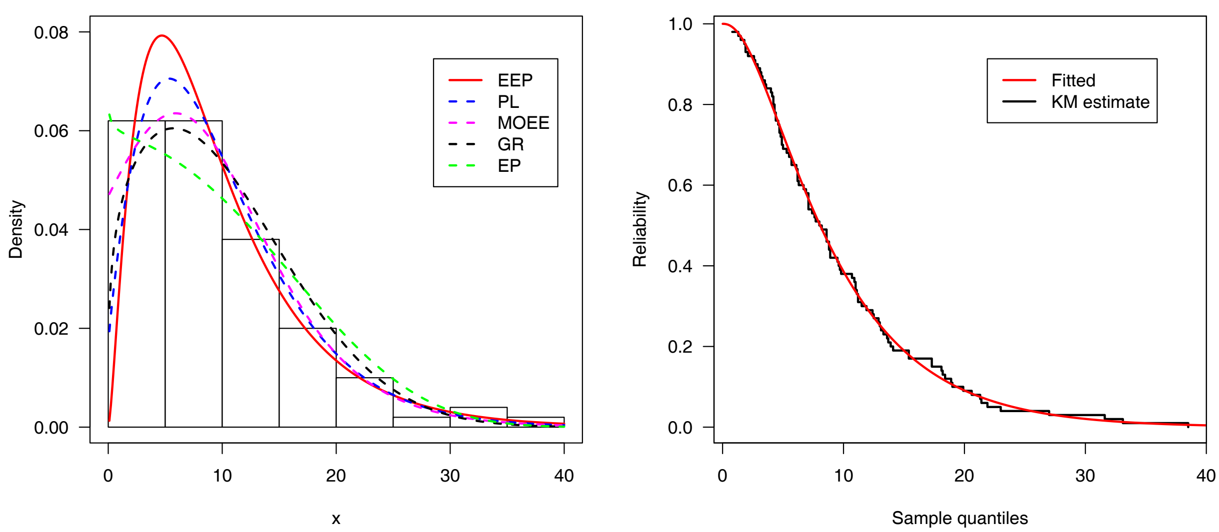

Figure 4: PDF plot with a histogram of fitted distributions (left side) and KM estimate with fitted quantiles (right side) of EEP distribution

5. Model formulation under the Bayesian approach

We usually assume the parameters $\Theta = (\alpha, \lambda, \theta)$ (for our study) as a constant in the classical approach and the goal is to investigate the distribution of the observed data set given Θ using the likelihood of the data sample. But the parameter Θ is considered as a random variable whereas the observed data set is taken as constant in the Bayesian approach (Lambert, 2018). In this type of modeling, prior information is used to support our assumption about the parameters of the distribution (Gelman *et al.*, 2013). In Bayesian modeling, the posterior distribution function is obtained by multiplying the prior distribution function and the likelihood function of the model under consideration for more detail see (McElreath, 2020). For Bayesian inference, we need the following elements

- The probability distribution function: $f(x/\Theta)$
- Prior distribution: $p(\Theta)$
- Likelihood: $p(Data/\Theta)$
- Data: $(x_1, ..., x_n)$

5.1. Prior distribution $p(\Theta)$

In Bayesian inference, a prior distribution (simply called prior) is the unconditional probability distribution that is used to express our beliefs about the true value of the parameters before the data is taken into account. The term $p(\Theta)$ denotes the probability distribution which represents our pre-data beliefs depending upon the different values of the parameters $\Theta = (\alpha, \lambda, \theta)$ of our model. In this study, we have taken the weakly informative Gamma prior for the parameters $\Theta = (\alpha, \lambda, \theta)$ as $\alpha \sim G(a_1, b_1)$, $\lambda \sim G(a_2, b_2)$ and $\theta \sim G(a_3, b_3)$. Particularly we have chosen $(a_1 = 0.001, b_1 = 0.001)$, $(a_2 = 0.001, b_2 = 0.001)$, and $(a_3 = 0, b_3 = 0.001)$ respectively for Gamma prior and most commonly used as weak prior on variance which is nearly flat as in Figure (5). The prior distributions can be written as

$$p(\alpha) = \frac{b_1^{a_1}}{\Gamma(a_1)} \alpha^{a_1 - 1} \exp(-b_1 \alpha); \ \alpha > 0, \ (a_1, b_1) > 0.$$

$$p(\lambda) = \frac{b_2^{a_2}}{\Gamma(a_2)} \lambda^{a_2 - 1} \exp(-b_2 \lambda); \ \lambda > 0, \ (a_2, b_2) > 0.$$

$$p(\theta) = \frac{b_3^{a_3}}{\Gamma(a_3)} \theta^{a_3 - 1} \exp(-b_3 \theta); \ \theta > 0, \ (a_3, b_3) > 0.$$

5.2. Likelihood $p(Data/\Theta)$

Given a set of data $(x_1, ..., x_n)$, the likelihood function of EEP distribution can be computed as

$$L(x) = (\alpha\lambda\theta)^n \prod_{i=1}^{n} x_i^{\alpha-1} \exp\left(1 + \lambda x_i^{\alpha} - e^{\lambda x_i^{\alpha}}\right) \left[1 - \exp\left(1 - e^{\lambda x_i^{\alpha}}\right)\right]^{\theta-1}. \tag{16}$$

5.3. Posterior distribution $p(\Theta/Data)$

Let $p(\alpha, \lambda, \theta/\underline{x})$ denote the posterior distribution and it can be obtained by using Bayes' rule as

$$p(\alpha, \lambda, \theta/\underline{x}) \propto L(\alpha, \lambda, \theta/\underline{x}) \times p(\alpha, \lambda, \theta).$$

In the Bayesian inference technique, we use Bayes' rule to estimate probability distribution called posterior distribution which can be obtained as

$$p(\Theta/data) \propto p(data/\Theta)\, p(\Theta).$$

$$p\left(\alpha,\lambda,\theta/\underline{x}\right)\propto\alpha^{n+a_1-1}\theta^{n+a_3-1}\lambda^{n+a_2-1}\prod_{i=1}^{n}e^{-b_1\alpha-b_2\lambda-b_3\theta}x_i^{\alpha-1}\times \exp\left(1+\lambda x_i^{\alpha}-e^{\lambda x_i^{\alpha}}\right)\left[1-\exp\left(1-e^{\lambda x_i^{\alpha}}\right)\right]^{\theta-1}. \tag{17}$$

All the information needed for Bayesian analyses is contained in the posterior distribution

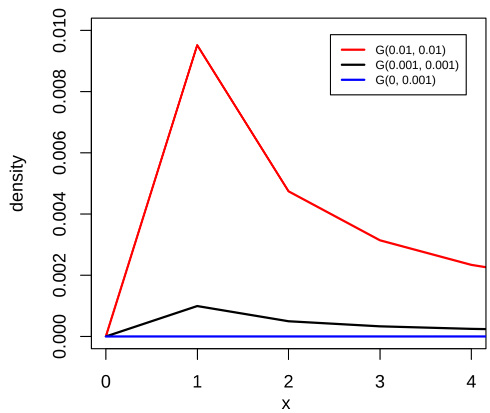

Figure 5: Graph of Gamma prior for various values of the parameters

and the aim is to compute the numeric as well as graphic summaries of it through integration. But the posterior distribution is quite complicated and could not draw any inferences. Hence we propose an alternative technique known as the simulation technique. This technique is based on the Markov Chain Monte Carlo (MCMC) method. MCMC draws samples by running a cleverly constructed Markov Chain that eventually converges to the target distribution i.e. posterior distribution $p\left(\alpha,\lambda,\theta/\underline{x}\right)$ (Brooks, 1998).

There are many different techniques to construct such chains some of them are, Gibbs sampler (Geman and Geman, 1984; Gelfand and Smith, 1990) are special cases of the general framework of (Metropolis *et al.*, 1953) and (Hastings, 1970). In this article, we implement MCMC algorithms through Stan (a probabilistic programming language) (Stan Development Team, 2022), the HMC algorithm, and its adaptive variant the NUTS for more detail see (Hoffman *et al.*, 2014; Carpenter *et al.*, 2017). Also Chaudhary and Kumar (2020) presented the Bayesian estimate of Gompertz extension distribution having three parameters. Also, Alizadeh *et al.* (2020) discussed the technique for estimating the model parameters of the odd log-logistic Lindley-G family of distribution.

6. MCMC method

6.1. HMC method

HMC is computationally a bit costly as compared to Metropolis and Gibbs sampling but its proposals are much more efficient (Gelman *et al.*, 2015). As a result, HMC doesn't require as many samples to explore the posterior distribution. For more detail about the HMC algorithm see (Beskos *et al.*, 2013).

6.1.1. No-U-Turn Sampler (NUTS)

NUTS engine routinely selects a suitable value for leapfrog step L in every iteration to maximize the distance at every L and control the random walk behavior. Let ω_1 and ω_0 be the current position and initial position of a particle and D be half of the distance between the positions ω_1 and ω_0 at each leapfrog step. The aim is to run leapfrog steps until ω_1 starts to move backward towards ω_0, which is achieved using the following algorithm, where leapfrog steps are run until the derivative of D with respect to time becomes less than 0.

$$\frac{\partial D}{\partial t} = \frac{\partial}{\partial t}\left[\frac{1}{2}\left(\omega_1 - \omega_0\right)^T\left(\omega_1 - \omega_0\right)\right] = \left(\omega_1 - \omega_0\right)^T p < 0.$$

However, this algorithm doesn't assure convergence or reversibility to the target distribution. The NUTS solves this type of problem by performing a doubling method for slice sampling (Neal, 2003). To generate the samples using NUTS, see (Hoffman *et al.*, 2014). For more details about NUTS, readers can go through (Nishio and Arakawa, 2019) and Devlin *et al.* (2021).

6.1.2. Defining the model in STAN

For the Bayesian analysis of the EEP model, we have used the latest Bayesian analysis software called Stan a high-level programming language that uses NUTS which is a variant of HMC simulation (Hoffman *et al.*, 2014). We have used the Rstan package (Stan Development Team, 2020) to run STAN in R software (R Core Team, 2022). The Stan scripts in R for the EEP model for the Bayesian analysis are presented in the appendix. We run the Stan using the algorithm HMC and engine NUTS having 4 chains for 2000 iterations. By default, Stan generates 1000 warm-up samples and 1000 real samples for a chain which are used for inferences.

6.2. Convergence and efficiency diagnostics for NUTS/ HMC and Markov chains

In the convergence diagnostic, we monitor the performance of NUTS/ HMC and MCMC sampling as

NUTS/ HMC: Here we study the information about divergence, energy, `tree-depth`, `step-size`, and `acceptance statistic`. Figure 6 (left panel) is the plot of the overlaid histograms of the marginal energy distribution π_E and the energy transition distribution $\pi_{\Delta E}$ for all 4 chains. The plot shows the histograms that look well-matched and indicate that the Hamiltonian Monte Carlo has performed robustly and Figure 6 (right panel) indicates that there are no divergent transitions. In Figure 7 we have displayed the performance of the NUTS sampling algorithm and Figure 8 are plots of the histogram of Rhat statistic, the ratio of effective sample size and sample size, and the ratio of Monte Carlo Standard Error (MCSE) and posterior SD. These plots show the good efficiency of the sampling algorithm NUTS for detail see Betancourt (2017).

MCMC: The MCMC draws can be monitored by plotting the following graphs autocorrelation plots, rank plots, trace plots, ergodic mean plots, and pairs plots.

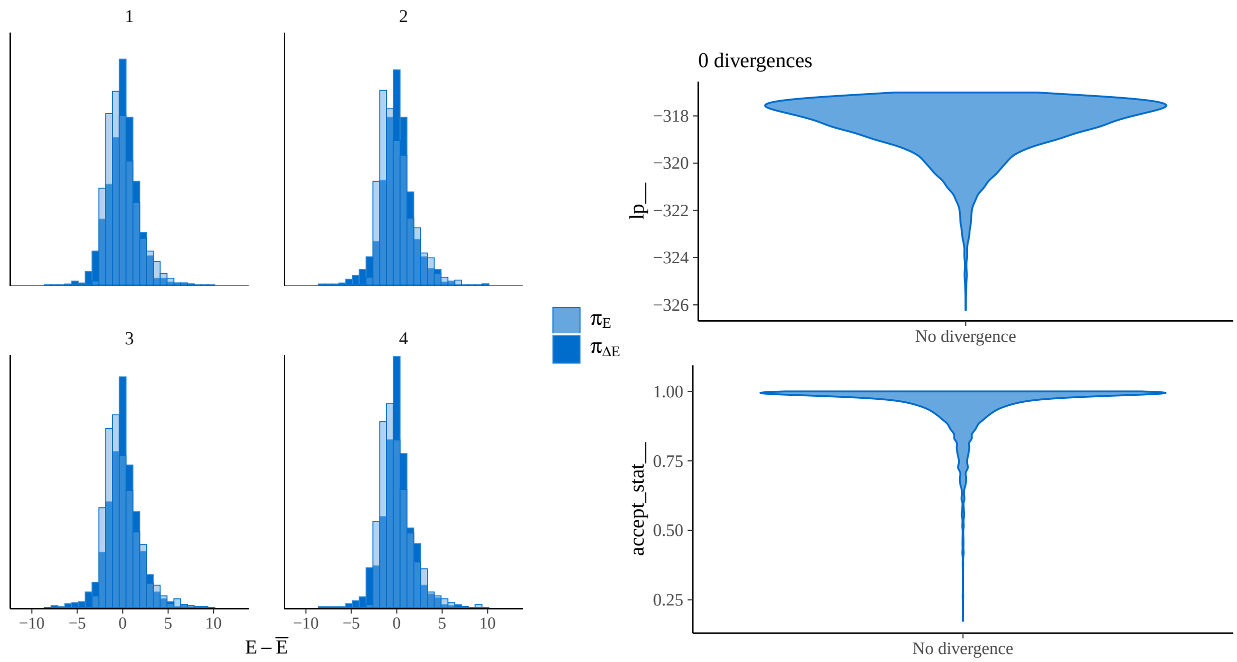

Figure 6: Histograms of π_E and $\pi_{\Delta E}$ for all 4 chains (left panel) and the divergent transition status (x-axis) against the log-posterior and the acceptance statistic (right panel) of the sampling algorithm for all chains

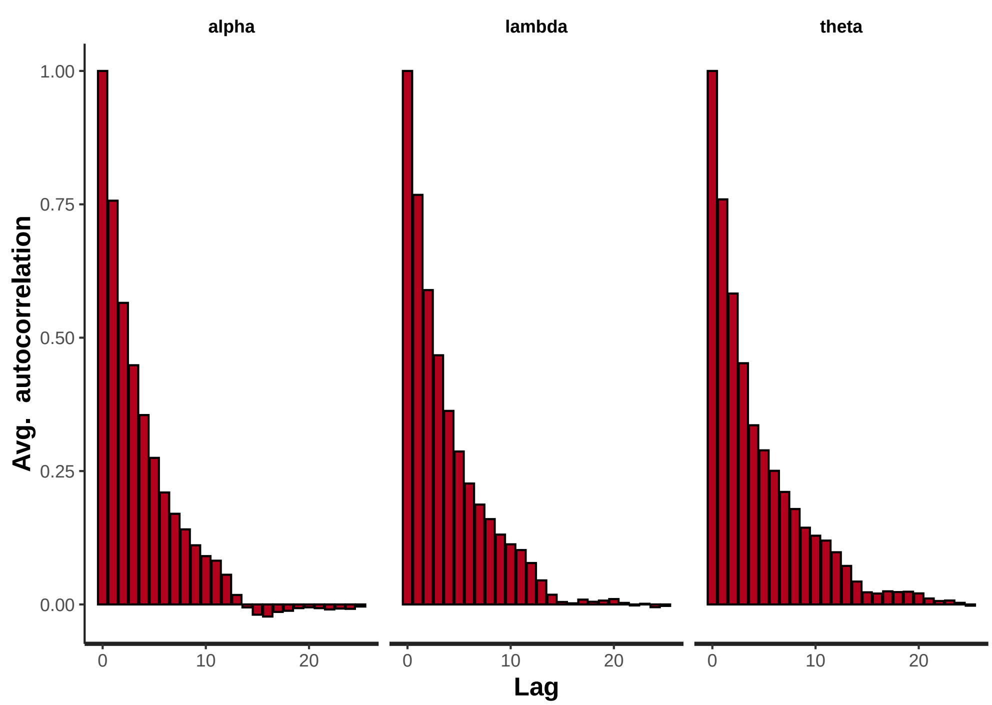

Figure 9: Autocorrelation plots of the parameters α, λ, and θ for all chains

These are autocorrelation plots for all chains and indicate that the samples of a Monte Carlo simulation are independent (Figure, 9). In Figure 10 we have displayed the histogram

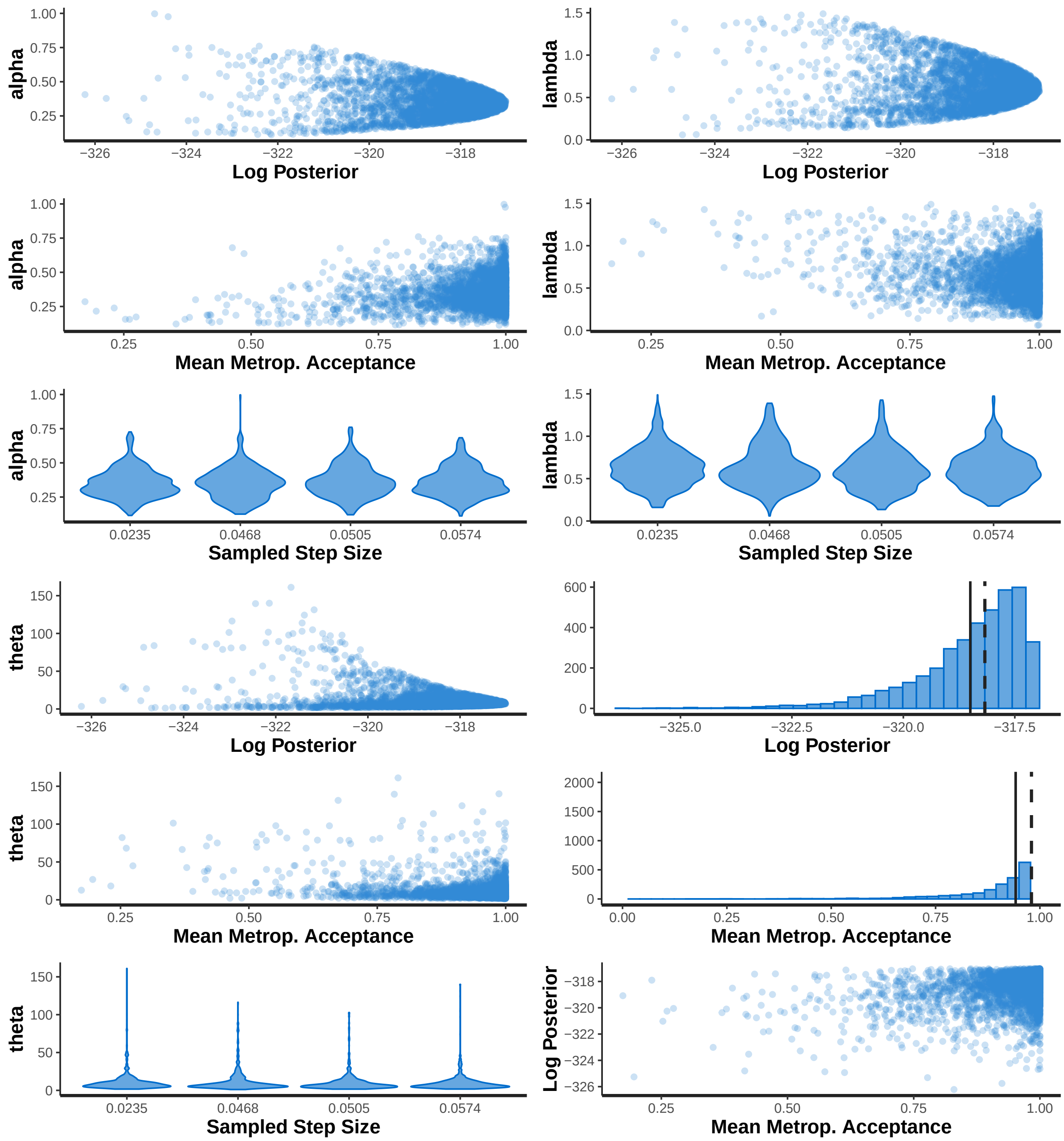

Figure 7: Average metropolis acceptance and step size for the parameters α, λ, θ and log posterior

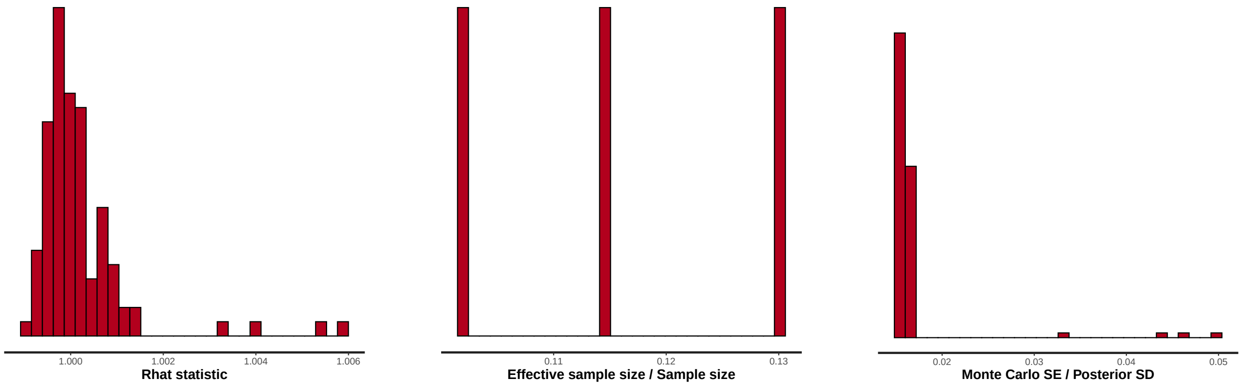

Figure 8: Histogram of Rhat statistic, the ratio of effective sample size and sample size, and the ratio of MCSE and posterior SD

of rank plots of α, λ, and θ for all four chains. Rank histograms visualize how the values from the chains mix in terms of ranking. An ideal plot would show the rankings mixing or overlapping in a uniform distribution. See (Vehtari *et al.*, 2021) for details. In general, we look for three possessions in the trace plots good mixing, stationarity, and convergence.

Good mixing implies that the chain quickly explores the full posterior region. It doesn't slowly wander, but rather rapidly zig-zags around, as a good Hamiltonian chain should. Stationarity indicates the path of each chain staying within the same high probability portion of the posterior distribution. Another way to imagine this is that the average value of the chain is relatively stable from start to end. Convergence represents that independent, multiple chains attach around the same area of high probability. Figure 11 shows the trace plots for alpha, theta, and lambda are well mixing and convergent for all chains.

The Ergodic mean is computed as the average of all values of the samples for all chains corresponding iterations. Figure 12 indicates that all chains converged smoothly around the mean value. Figure 13 is a pairs plot of MCMC draws of α, λ, and θ. Univariate marginal posteriors are shown along the diagonal as histograms. Bivariate plots are displayed above and below the diagonal as scatter plots. The red colored draws represent, if present, the divergent transitions. Divergent transitions can indicate problems with the validity of the results. A good plot would show no divergent transitions. A bad plot would show divergent transitions in a systematic pattern. We have also presented a detailed numerical summary of the HMC and NUTS algorithm in Table 5 and statistics related to the posterior summary are presented in Table 6.

Table 5: Informational statistic of NUTS/HMC for convergence of chains

	accept_stat	stepsize	treedepth	n_leapfrog	divergent	energy
All chains	0.9419	0.0446	3.9813	36.2755	0	319.991
chain1	0.9443	0.0574	3.8890	33.0920	0	319.913
chain2	0.9294	0.0505	3.9140	32.6420	0	320.035
chain3	0.9536	0.0468	4.0110	36.7580	0	320.111
chain4	0.9403	0.0235	4.1110	42.6100	0	319.907

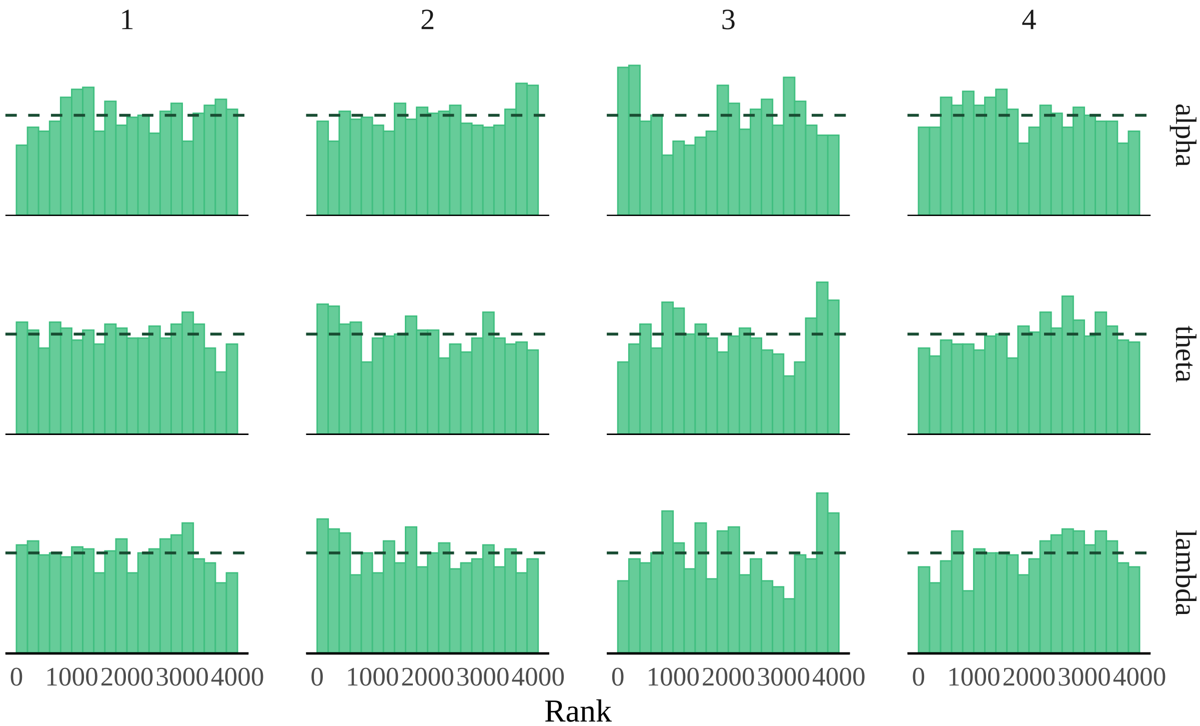

Figure 10: Rank histograms of α, λ, and θ

7. Posterior analysis

7.1. Numerical summary

Using the `stan()` function in R-Software we have estimated the posterior density of the fitted EEP model. The numerical summaries of the posterior distribution after fitting the EEP model for the data taken under study for all merged chains are reported in Table 6. The MCMC estimate for α is 0.36 ± 0.109 which is statistically significant. Similarly, the estimate for λ is 0.62 ± 0.233, which is statistically significant. The estimate for θ is 10.926 ± 12.646 which is also statistically significant. No parameters have an effective sample size (`n_eff`) for estimating the posterior mean less than 10 % of the total sample size indicating that the samples are efficient and `Rhat`($\hat{R}$) (estimated potential scale reduction statistic) provides the analysis of sampling and its efficiency. Here `Rhat` is less than 1.01 indicating convergence of all chains. Also, we have depicted the highest posterior density (HPD) credible interval and credible interval in Table 7.

Table 6: Output summary of posterior samples for the EEP model

Parameters	mean	se_mean	sd	2.50%	50%	97.50%	n_eff	Rhat
alpha	0.3553	0.0048	0.1094	0.1716	0.3454	0.6085	520	1.0039
lambda	0.6212	0.0109	0.2334	0.2339	0.5966	1.1632	458	1.0055
theta	10.9263	0.6267	12.6463	2.4253	7.2978	45.8048	407	1.0059
Log-posterior	-318.4981	0.0410	1.2386	-321.6890	-318.1700	-317.1100	911	1.0033

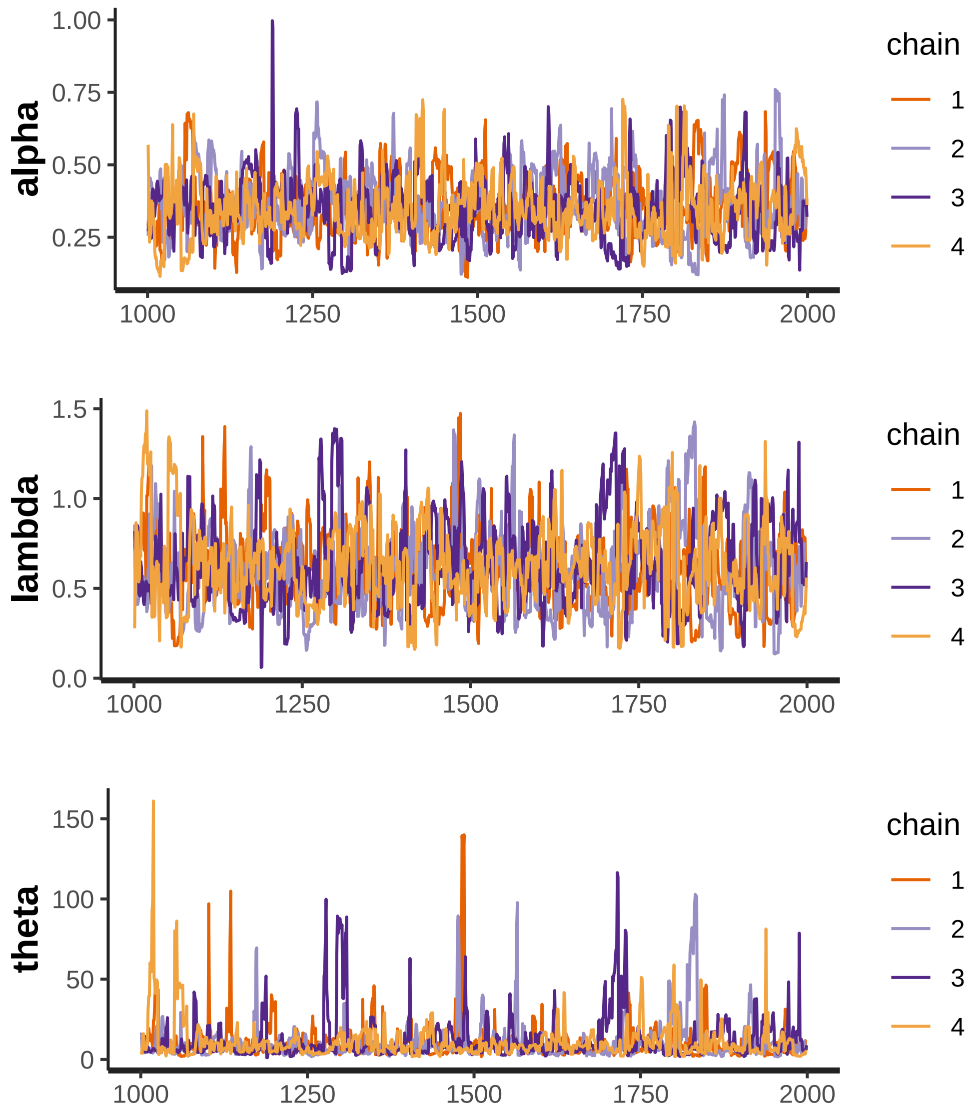

Figure 11: Trace plot of the parameters α, λ, and θ for all chains

7.2. Visual summary

Various graphical representations can be employed to visually summarize the posterior distribution, such as histograms, boxplots, caterpillar plots, and density plots. In this study, we utilized Gamma priors to plot histograms and kernel density estimates for α, λ, and θ (Figure, 14), based on a total of 4000 posterior samples. These graphical presentations offer comprehensive insights into the parameters' posterior distribution. Histograms are particularly useful for understanding the distribution's tail behavior, skewness, kurtosis, the presence of outliers, and whether multi-modal behavior exists. Our analysis reveals that α and λ exhibit almost symmetrical distributions, whereas θ demonstrates positive skewness under Gamma prior. Furthermore, in Figure (15), we present histograms of posterior parameters using a Uniform prior. It is evident that the choice of prior significantly impacts the resulting posterior distribution.

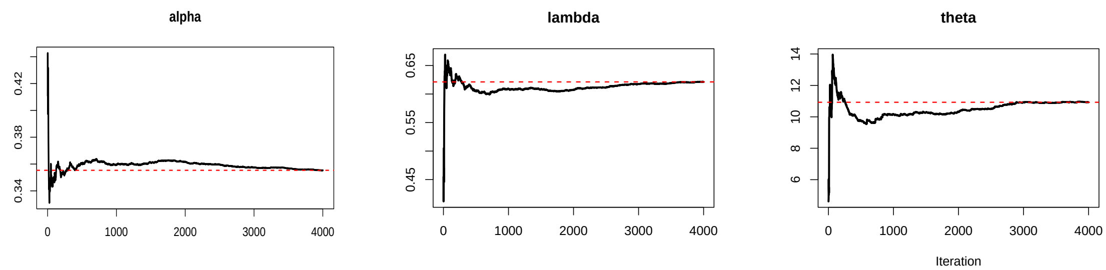

Figure 12: The Ergodic mean plots for α, λ, and θ

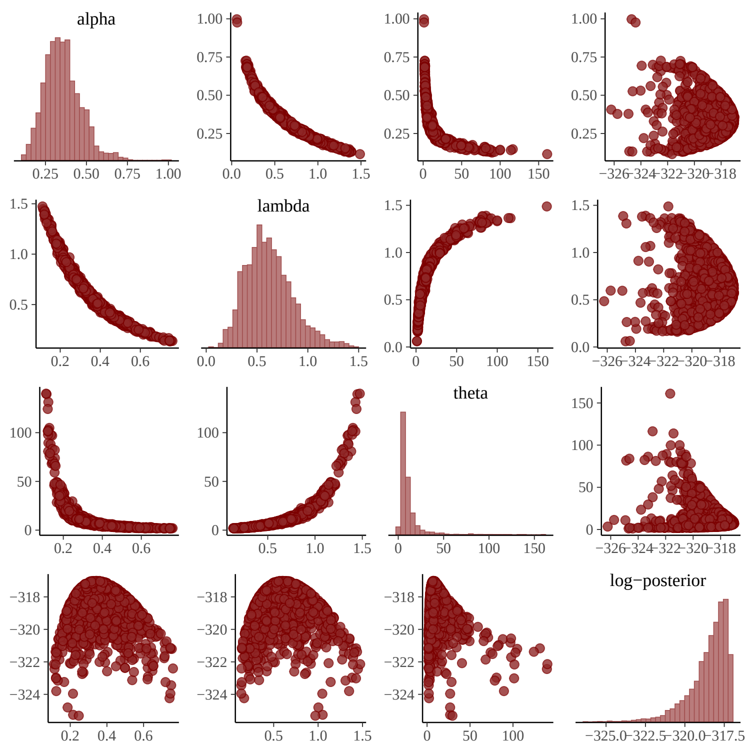

Figure 13: Pairs plot of α, λ, θ and log-posterior

8. Model compatibility

8.1. Posterior predictive checks (PPCs)

A usual way to access the fit of a Bayesian model is to observe how well the predictions can be made from the model that agrees with the observed data (Gelman, 2003; Gelman *et al.*, 2004). If our model is capable of fitting the data then it should generate data that are quite similar to the observed data. The data that are used for posterior predictive checks (PPCs) we can generate them by simulating the posterior predictive distribution. The R package bayesplot presents different plotting functions for visual posterior predictive checking; using observed data and simulated data from the posterior predictive distribution, we can generate these graphical displays (Gabry *et al.*, 2017).

Table 7: HPD interval and credible interval for model parameters α, λ, and θ

Parameters	HPD interval	Credible Interval
alpha	(0.141, 0.554)	(0.1716, 0.6085)
lambda	(0.168, 1.070)	(0.2339, 1.1632)
theta	(1.650, 31.20)	(2.4253, 45.8048)

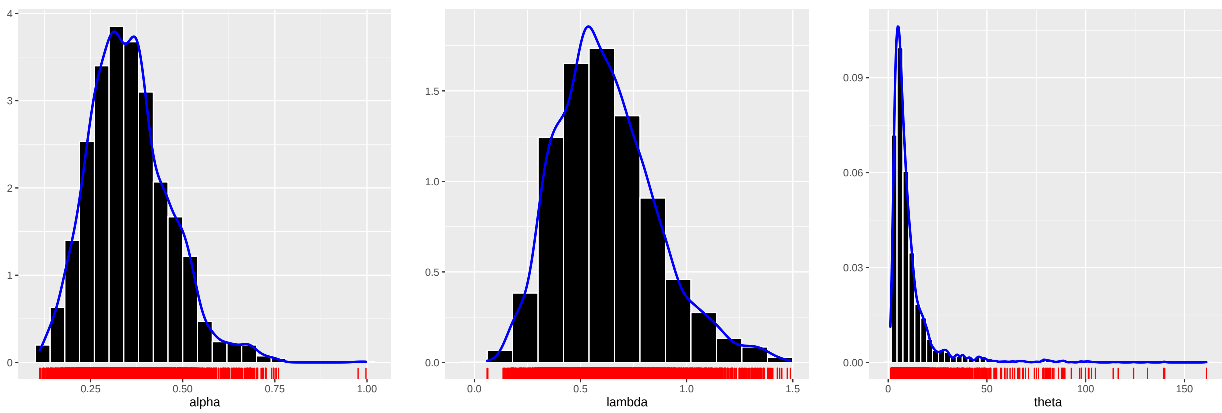

Figure 14: Histogram with kernel density estimates of posterior samples for the parameters α, λ, and θ respectively under a gamma prior

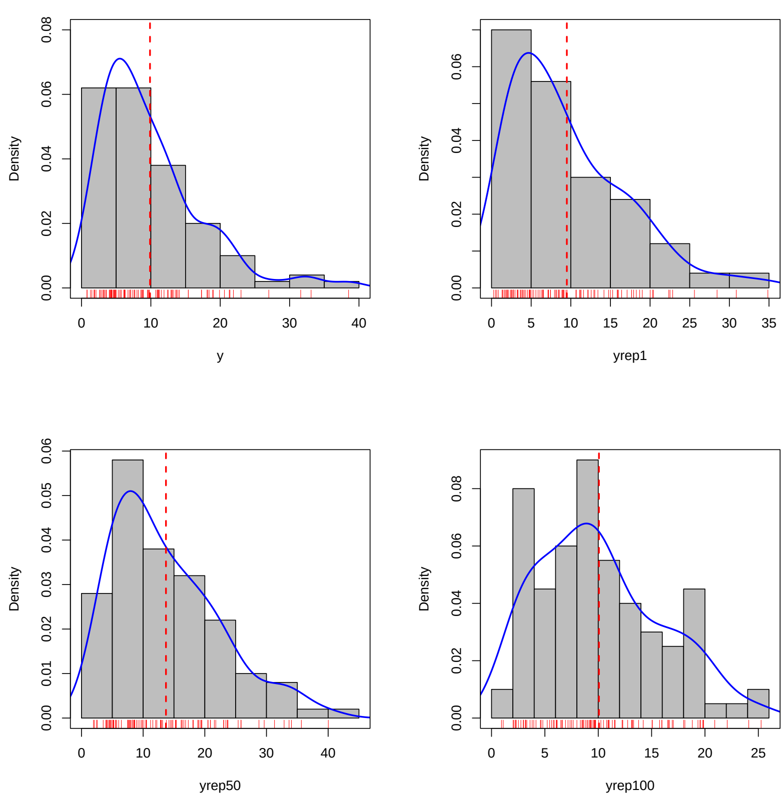

Figure 17: Histogram of observed data (y) and replicated data $y_{rep}[1]$, $y_{rep}[50]$ *and* $y_{rep}[100]$ with point estimate mean (red vertical line)

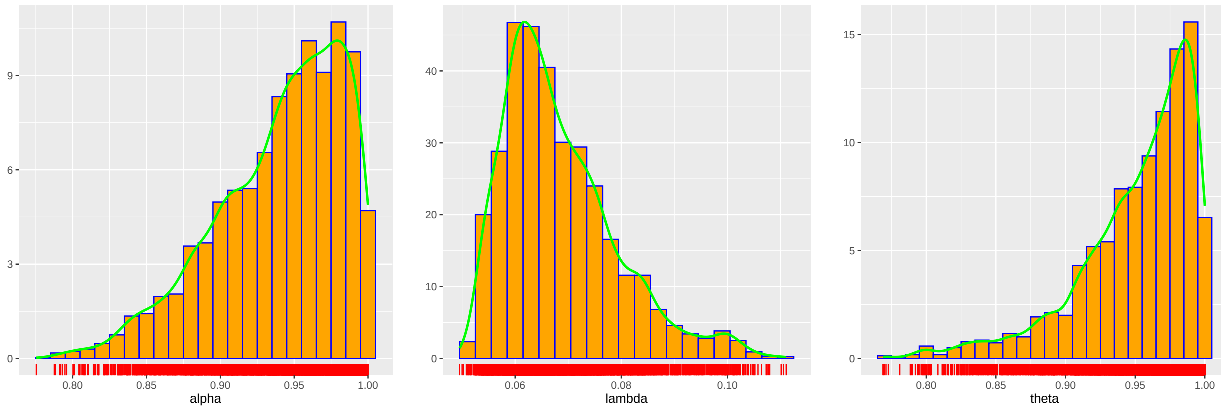

Figure 15: Histogram with kernel density estimates of posterior samples for the parameters α, λ, and θ respectively under a uniform prior

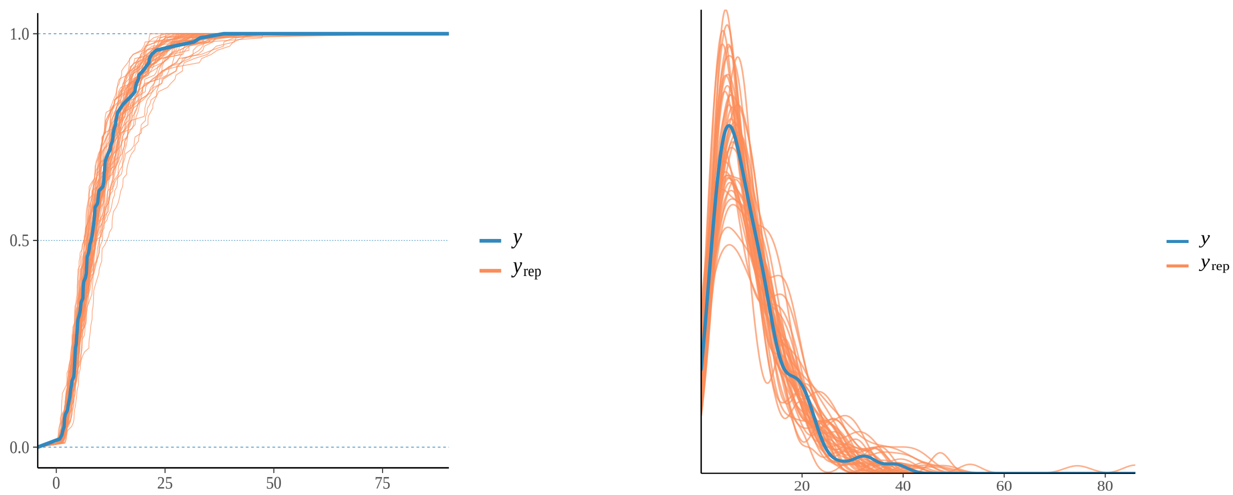

Figure 16: CDF plot of the observed dataset y (blue), with 40 simulated datasets y_{rep} (left panel) and kernel density estimate of the observed dataset y (blue), with density estimates for 40 simulated datasets y_{rep} drawn from the posterior predictive distribution (left panel)

The posterior predictive distribution is the distribution of the outcome variable implied by a model after using the observed data y (a vector of length $N = 100$) to update our beliefs about unknown parameters $\Theta = (\alpha, \lambda, \theta)$ of the model. The posterior predictive distribution for observation y_{rep} can be written as,

$$p(\tilde{y}/y) = \int p(\tilde{y}/\Theta)\, p(\Theta/y) d\Theta.$$

For every simulation (draw) $s = 1, ..., S$ of the parameters from the posterior distribution $\Theta(s) \sim p(\Theta|y)$, we generate a vector of N outcomes $\tilde{y}^{(\mathrm{s})}$ using the posterior predictive distribution by simulating from the data model conditional on parameters $\Theta_{(s)}$.

The result is an $S \times N (4000 \times 100)$ matrix of draws $\tilde{y}$. We have denoted the resulting simulation matrix by y_{rep}, this matrix is the replication of the observed data y rather than predictions for future observations. To attain further clarity on our decision for the study of the posterior predictive checks we have taken the smallest, middle, and largest, i.e. ($y_{rep}[1],\ y_{rep}[50]\ and\ y_{rep}[100]$) replicated observations. We have presented a wide variety of graphical model checks based on comparing observed data to draws from the posterior (or prior) predictive distribution. To Compare the empirical distribution of the data y to the distributions of simulated/replicated data y_{rep} from the posterior predictive distribution an empirical CDF estimate of each dataset (row) in y_{rep} are overlaid with the distribution of y (blue curve) is displayed in (Figure, 16, left panel) and kernel density estimate of the observed dataset y (blue), with density estimates for 40 simulated datasets y_{rep} drawn from the posterior predictive distribution (Figure, 16, right panel). To analyze the predicting capacity of posterior samples we have presented the visual summaries such as a histogram with kernel density plot for observed data y and simulated data $y_{rep}[1],\ y_{rep}[50]\ and\ y_{rep}[100]$ (Figure 17).

8.2. Model selection

The WAIC (Widely Applicable Information Criterion) is used to compare different statistical models based on their out-of-sample predictive accuracy. The lower the WAIC value, the better the model's predictive performance. Hence EEP model is better than the EP model see (Table, 8). Where `elpd_waic` is the estimated log pointwise predictive density using the WAIC. It represents the model's fit to the data and is measured in terms of log-likelihood. `p_waic` is the effective number of parameters computed from the WAIC. It takes into account both the actual number of parameters in the model and the model's complexity and `waic` is the value of the WAIC itself, which is a combination of the model fit (`elpd_waic`) and the effective number of parameters `p_waic`. A lower WAIC indicates better predictive performance.

Table 8: Model selection statistics

Estimate	EEP distribution	EP distribution
`elpd_waic`	-319.6	-326.9
`p_waic`	1.5	1.2
`waic`	639.1	653.9

9. Conclusion

In this research work, we put forward a new distribution using the exponential power model as a baseline distribution and named it exponentiated exponential power (EEP) distribution. We have explored some properties including the hazard rate function, cumulative distribution function, survival function, probability density function, cumulative hazard function, order statistics, quantiles, the measures of skewness based on quartiles, and median, and kurtosis based on octiles.

Also we have performed a full Bayesian analysis for the proposed model. Using Stan software whose MCMC techniques are based on the NUTS which is an adaptive variant of HMC; a more robust and efficient sampler. We have presented the numerical as well as graphical analysis of the EEP model and found that all chains are well mixed and conversed. Further, we have estimated the parameters of the model and performed posterior predictive checks, and found that the underlying model can be used to generate reliable samples. The developed techniques are applied to a real data set, thus we can apply for full Bayesian analysis for the proposed model using these Bayesian techniques. Hence it is expected that the EEP model will be a choice in the fields of the theory of probability, applied statistics, bayesian inferences, and survival analysis.

Acknowledgements

We are very grateful to the Chair Editor and the reviewer for valuable comments and suggestions which have improved considerably the first version of the manuscript.

References

Alizadeh, M., Afify, A. Z., Eliwa, M. S., and Ali, S. (2020). The odd log-logistic Lindley-G family of distributions: properties, Bayesian and non-Bayesian estimation with applications. *Computational Statistics*, **35**, 281–308.

Almarashi, A. M., Elgarhy, M., Elsehetry, M. M., Kibria, B., and Algarni, A. (2019). A new extension of exponential distribution with statistical properties and applications. *Journal of Nonlinear Sciences and Applications (JNSA)*, **12**, 135–145.

Ashour, S. K. and Eltehiwy, M. A. (2015). Exponentiated power Lindley distribution. *Journal of Advanced Research*, **6**, 895–905.

Balakrishnan, N. and Cohen, A. C. (2014). *Order Statistics and Inference: Estimation Methods*. Elsevier.

Barriga, G. D., Louzada-Neto, F., and Cancho, V. G. (2011). The complementary exponential power lifetime model. *Computational Statistics & Data Analysis*, **55**, 1250–1259.

Beskos, A., Pillai, N., Roberts, G., Sanz-Serna, J.-M., and Stuart, A. (2013). Optimal tuning of the hybrid Monte Carlo algorithm. *Bernoulli*, **19**, 1501–1534.

Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. *arXiv preprint arXiv:1701.02434*, **22**.

Brooks, S. (1998). Markov chain Monte Carlo method and its application. *Journal of the Royal Statistical Society: series D (the Statistician)*, **47**, 69–100.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, **76**.

Chaubey, Y. P. and Zhang, R. (2015). An extension of chen's family of survival distributions with bathtub shape or increasing hazard rate function. *Communications in Statistics-Theory and Methods*, **44**, 4049–4064.

Chaudhary, A. K. and Kumar, V. (2020). A Bayesian estimation and prediction of Gompertz extension distribution using the MCMC method. *Nepal Journal of Science and Technology*, **19**, 142–160.

Chen, Z. (1999). Statistical inference about the shape parameter of the exponential power distribution. *Statistical Papers*, **40**, 459–468.

Chen, Z. (2000). A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. *Statistics and Probability Letters*, **49**, 155–161.

Devlin, L., Horridge, P., Green, P. L., and Maskell, S. (2021). The No-U-Turn sampler as a proposal distribution in a sequential Monte Carlo sampler with a near-optimal L-kernel. *arXiv preprint arXiv:2108.02498*, **Preprint**, 1–5.

Dey, S., Kumar, D., Ramos, P. L., and Louzada, F. (2017). Exponentiated Chen distribution: Properties and estimation. *Communications in Statistics-Simulation and Computation*, **46**, 8118–8139.

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., and Gelman, A. (2017). Visualization in Bayesian workflow. *arXiv preprint arXiv:1709.01449*, **182**.

Gelfand, A. E. and Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, **85**, 398–409.

Gelman, A. (2003). A Bayesian formulation of exploratory data analysis and goodness-of-fit testing. *International Statistical Review*, **71**, 369–382.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). Bayesian data analysis chapman & hall. *CRC Texts in Statistical Science*, **136**.

Gelman, A., J. B., C., Hal S., S., David B., D., Vehtari, A., and Donald B., R. (2013). *Bayesian data analysis*. Chapman and Hall/CRC.

Gelman, A., Lee, D., and Guo, J. (2015). Stan: A probabilistic programming language for bayesian inference and optimization. *Journal of Educational and Behavioral Statistics*, **40**, 530–543.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721–741.

Ghitany, M., Al-Mutairi, D. K., and Aboukhamseen, S. (2015). Estimation of the reliability of a stress-strength system from power Lindley distributions. *Communications in Statistics-Simulation and Computation*, **44**, 118–136.

Ghitany, M. E., Atieh, B., and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, **78**, 493–506.

Gupta, R. D. and Kundu, D. (1999). Theory and methods: Generalized exponential distributions. *Australian and New Zealand Journal of Statistics*, **41**, 173–188.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrica*, **57**, 97–109.

Henningsen, A. and Toomet, O. (2011). maxlik: A package for maximum likelihood estimation in R. *Computational Statistics*, **26**, 443–458.

Hoffman, M. D., Gelman, A., et al. (2014). The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, **15**, 1593–1623.

Joshi, R. K., Sapkota, L. P., and Kumar, V. (2020). The logistic-exponential power distribution with statistical properties and applications. *International Journal of Emerging Technologies and Innovative Research*, **7**, 629–641.

Kundu, D. and Raqab, M. Z. (2005). Generalized Rayleigh distribution: different methods of estimations. *Computational Statistics and Data Analysis*, **49**, 187–200.

Lambert, B. (2018). *A Student's Guide to Bayesian Statistics*. SAGE Publications Ltd.

Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application. *Biometrika*, **84**, 641–652.

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Chapman and Hall/CRC.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of Chemical Physics*, **21**, 1087–1092.

Moors, J. (1988). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, **37**, 25–32.

Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, **42**, 299–302.

Nadarajah, S. and Haghighi, F. (2011). An extension of the exponential distribution. *Statistics*, **45**, 543–558.

Neal, R. (2011). *MCMC Using Hamiltonian Dynamics*. Chapman & Hall/CRC.

Nishio, M. and Arakawa, A. (2019). Performance of hamiltonian monte carlo and no-u-turn sampler for estimating genetic parameters and breeding values. *Genetics Selection Evolution*, **51**, 1–12.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ristić, M. M. and Nadarajah, S. (2014). A new lifetime distribution. *Journal of Statistical Computation and Simulation*, **84**, 135–150.

Sapkota, L. P. (2020). Exponentiated–exponential logistic distribution: Some properties and application. *Janapriya Journal of Interdisciplinary Studies*, **9**, 100–108.

Sapkota, L. P. (2022). A Bayesian analysis and estimation of weibull inverse rayleigh distribution using hmc method. *Nepal Journal of Mathematical Sciences*, **3**, 39–58.

Smith, R. M. and Bain, L. J. (1975). An exponential power life-testing distribution. *Communications in Statistics-Theory and Methods*, **4**, 469–481.

Stan Development Team (2020). RStan: the R interface to Stan. R package version 2.21.1.

Stan Development Team (2022). The Stan Core Library. Version 2.31.0.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., and Bürkner, P.-C. (2021). Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with discussion). *Bayesian Analysis*, **16**, 667–718.

ANNEXURE

```
functions{real expexp_power_lpdf(real y, real alpha,
        real lambda, real theta){
        return log(alpha*lambda*theta) + (alpha-1)*
        log(y)+1+lambda*y^alpha - exp(lambda*y^alpha)+
                (theta-1)*log(1-exp(1-exp(lambda*y^alpha)));
        }
        real expexp_power(real alpha, real lambda, real theta){
        return ((1/lambda)*log(1-log(1-
        (uniform_rng(0,1))^(1/theta))))^(1/alpha);
        }
}
```

```
    data{
        int N;
        real y[N];
    }

    parameters{
        real <lower=0> alpha;
        real <lower=0> lambda;
        real <lower=0> theta;
    }

    model{
        for(i in 1 : N) {
            y[i] ~ expexp_power(alpha, lambda, theta);
        }
        alpha ~ gamma(0.001, 0.001);
        lambda ~ gamma(0.001, 0.001);
        theta ~ gamma(0, 0.001);
    }

    generated quantities{
    vector [N] yrep;
    for(i in 1 : N)
    {
        yrep[i]= expexp_power_rng(alpha, lambda, theta);
    }
}
```

Data Creation in R software

```
y = c(0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9,
3.1, 3.2, 3.3, 3.5, 3.6, 4.0, 4.1, 4.2, 4.2, 4.3, 4.3,
4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9, 5.0, 5.3, 5.5,
5.7, 5.7, 6.1, 6.2, 6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1,
7.1, 7.1, 7.4, 7.6, 7.7, 8.0, 8.2, 8.6, 8.6, 8.6, 8.8,
8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11.0,
11.0, 11.1, 11.2, 11.2, 11.5, 11.9, 12.4, 12.5, 12.9,
13.0, 13.1, 13.3, 13.6, 13.7, 13.9, 14.1, 15.4, 15.4,
17.3, 17.3, 18.1, 18.2, 18.4, 18.9, 19.0, 19.9, 20.6,
21.3, 21.4, 21.9, 23.0, 27.0, 31.6, 33.1, 38.5)
N <- length(y)
Data = list(y=y, N=N).
```