



# Bayesian Estimation and Prediction of Three-Parameter Complementary Exponential Power Distribution using MCMC Technique

Arun Kumar Chaudhary, Vijay Kumar

**Abstract:** The Markov chain Monte Carlo (MCMC) technique is applied for estimating the Complementary Exponential Power (CEP) distribution's parameters through the analysis of complete sample in this article. With the help of the Bayesian and the Maximum Likelihood techniques, the unknown parameters of the model are estimated. To find Complementary Exponential Power distribution's parameters' Bayesian estimates, a new methodology is developed, via simulation method of MCMC through the application of OpenBUGS platform. To demonstrate under the gamma and uniform sets of priors, a real data set is taken. The generations of posterior MCMC samples is conducted with OpenBUGS software. For analyzing the output of so generated MCMC samples, and studying the statistical properties, distribution's comparison tools and model validation the functions of R have been used. The credible interval and predicted of the reliability, hazard and modal parameters' values are also estimated. We have shown that Bayesian estimators are more efficient than classical estimators for any real data set.

**Keywords:** Complementary exponential power model, Markov chain Monte Carlo, Bayesian estimation, Maximum likelihood estimation, Gamma Prior.

## I. INTRODUCTION

Smith & Bain (1975) first developed lifetime model of the distribution of exponential power (EP). Several writers (Leemis, 1986; Rajarshi & Rajarshi, 1988; Chen, 1999) have examined EP distribution. Novel models were developed in past years centered on changes of the Exponential Power model. Including two shape parameters, Chen (2000) presented a novel model. Though the number of shape parameters was only two, the new model has some nice features. First its capacity to handle two hazard functions (bathtub-shaped and increasing) is attractive in the model, and second, two shape parameters' joint confidence region and confidence interval have a closed form. Yet for the evaluation of a number of datasets, it requires a scale parameter for more flexibility. A new model known as the model of Weibull extension, introduced by (Xie et al., 2002), which has an

extra scale parameter to the model proposed by Chen, to address such a drawback. This led to the model becoming more flexible and the practitioners found it to be more persuasive. But the Xie model version, like its predecessors, just lodges two hazard functions (bathtub-shaped and increasing) Though the distribution of EP and its variants are often applied to evaluate survival time data, this distribution fails to give a suitable parametric fit in real world situations where hazard functions could be unimodal or decreasing shaped. The CEP distribution, that is primarily associated with (Smith & Bain, 1975)'s distribution that has an additional shape parameter, was presented not long ago by (Barriga et al., 2011). There are several flexible distributions in the survival literature that can handle decreasing, unimodal, bathtub-shaped and increasing failure rate functions (Mudholkar & Srivastava, 1993; Mudholkar et al., 1996; Pham & Lai, 2007; Carrasco et al., 2008). Although certain distributions include four parameters, a distribution of three parameters is the Complementary Exponential Power (CEP). This is practically a benefit, because it would be critical to analyze close models in as few parameters as feasible. Especially with moderate and small sample sizes, which is a typical survival analysis condition, the parameters is not calculated precisely (Xie et al., 2002).

The development of exponentiated distribution is generally performed as per the description of (Marshall & Olkin, 2007). Based on the observation, a novel CDF  $F(x) = (F_{\text{baseline}}(x))^{\theta}$ ,  $\theta > 0$  can be obtained by taking up an arbitrary power  $\theta > 0$  to any baseline CDF  $F_{\text{baseline}}(x)$ , however  $F(x)$  can be termed resilience parameter family as this extra parameter  $\theta$  is termed resilience parameter.

Even though this may not be our situation, considering  $\theta$  as an integer, word resilience readily emerges. We can  $F(x)$  see as cumulative distribution function of a concurrent model with independent components that is less prone to failure with increasing no. of components, contributing in this case to a resilient design. Many writers following this idea have described extensions of survival distributions. (Mudholkar et al., 1995) first presented the exponentiated Weibull distribution which generalizes the Weibull distribution, three-parameter generalized exponentiated exponential distribution which is the standard exponential distribution generalization, presented by (Gupta & Kundu, 1999), and four additional exponentiated distributions which generalized the standard Weibull, Gumbel, Fréchet and the gamma distributions,

Revised Manuscript Received on December 25, 2020.

\* Correspondence Author

**Arun Kumar Chaudhary\***, is an Associate Professor of Department of Management Science(Statistics), Nepal Commerce Campus, Tribhuvan University, Nepal, PH- +9779851084055, Email: akchaudhary1@yahoo.com, chaudharyak111@gmail.com

**Vijay Kumar**, Department of Mathematics and Statistics, DDU Gorakhpur University, Gorakhpur, India. Email: vijay.mathstat@ddugu.ac.in, vkgkp@rediffmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

# Bayesian Estimation and Prediction of Three-Parameter Complementary Exponential Power Distribution using MCMC Technique

presented by (Nadarajah & Kotz,2006), are some examples. Three-parameter Complementary Exponential Power Model is presented in this paper along with analysis of maximum likelihood estimate (MLE). Bayesian model for proposed distribution is formulated. For Bayesian analysis, a real set of data is illustrated.

Maximum Likelihood Estimation, corresponding model validation and exploratory data analysis have been conducted. The independent gamma and uniform set of priors of the data set are used for the complete Bayesian analysis with the help of MCMC technique using the OpenBugs platform. The parameters' Bayes estimates, reliability as well as hazard functions have been acquired and their credible interval are described. For the set of data, we have also addressed the question of model compatibility

## II. MODEL ANALYSIS

### A. Cumulative distribution function (cdf):

(Smith & Bain, 1975) introduced the CDF of the EP distribution which is presented in the expression below,  

$$F_{EP}(x, \lambda, \alpha) = 1 - e^{(1-\exp(x\lambda)^\alpha)} \quad (1)$$

Where,  $x > 0, \lambda > 0; \alpha > 0$ .

The Cumulative distribution function of the Complementary Exponential Power distribution is obtained by taking up  $F_{EP}(x)$  to the power of  $\theta$  namely  $F(x) = (F_{EP}(x))^\theta$ . The distribution function of CEP with three parameters is given by, (Barriga *et al.*, 2011)  

$$F(\theta, \lambda, \alpha; x) = (1 - e^{(1-\exp(x\lambda)^\alpha)})^\theta \quad (2)$$
 Here, shape parameters are  $\alpha > 0, \theta > 0$  and scale parameter is  $\lambda > 0$ . If we take  $\theta = 1$ , the function (2) diminishes to EP distribution.  $CEP(\alpha, \lambda, \theta)$  represents CEP distribution.

### B. Probability density function (PDF):

The PDF can be expressed as,  

$$f(\theta, \lambda, \alpha; x) = P_1 \exp[1 - e^{(x\lambda)^\alpha} + (x\lambda)^\alpha] x^{\alpha-1} \lambda^\alpha \theta \alpha \quad (3)$$
 Where,  $P_1 = [1 - e^{(1-\exp(x\lambda)^\alpha)}]^\theta - 1$

When  $\alpha = 1, \lambda = 1$ , for different values of  $\theta$ , some typical PDF of CEP are portrayed in Figure 1. In Figure 1, we can observe that the Complementary Exponential Power distribution's PDF can exhibit various geometry.

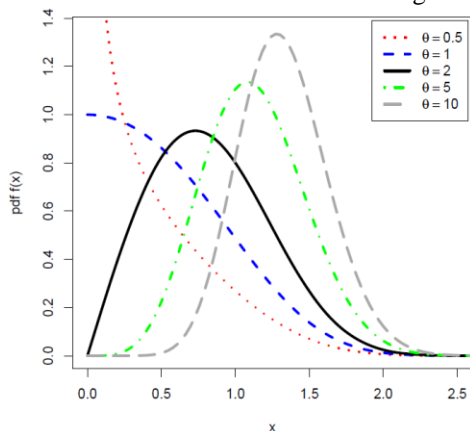


Fig 1. The PDF of CEP distribution for various values of  $\theta, \alpha = 1, \lambda = 1$ .

### C. Reliability/Survival function:

Reliability function is given by the following expression  

$$R(\theta, \lambda, \alpha; x) = (1 - e^{(1-\exp(x\lambda)^\alpha)})^\theta; x \geq 0 \quad (4)$$

### D. The Hazard-Rate Function (HRF):

The expression for HRF is provided below.

$$h(\theta, \lambda, \alpha; x) = \frac{x^{\alpha-1} \lambda^\alpha \theta \alpha * \exp[(x\lambda)^\alpha + 1 - e^{(x\lambda)^\alpha}] * P_2}{1 - [1 - e^{(1-\exp(x\lambda)^\alpha)}]^\theta} \quad (5)$$

Where,  $P_2 = [1 - e^{(1-\exp(x\lambda)^\alpha)}]^\theta - 1$

Based on parameter values lying on the shape parameters ( $\alpha > 0$  and  $\theta > 0$ )' domain area, which is bounded by  $\theta = 1, \alpha = 1$  and the curve  $\alpha\theta = 1$ , unimodal, bathtub, and various monotone hazard functions including the bathtub and unimodal functions are accommodated by the CEP distribution. Some of the characteristics of the CEP distribution hazard function are as follows:

- (i)  $h(0) = 0$ , when  $h(0) = \lambda$  &  $\alpha > 1$ , when  $\theta, \alpha = 1$ .
- (ii) When  $\alpha \leq 1; \theta > 1$ , hazard function shows decreasing curve.
- (iii) When  $\alpha > 1$  and  $\theta \leq 1$ , hazard function shows increasing curve.
- (iv) When  $\alpha < 1$  and  $\theta \leq 1$  (or  $\alpha > 1$  and  $\alpha\theta < 1$ ), then  $h(x)$  shows bathtub shaped curve.
- (v) When  $\alpha < 1$  and  $\alpha\theta > 1$ , then  $h(x)$  shows upside-down bathtub curve.

Since the function of the hazard rate in (5) is complicated, some characteristics of (ii-v) were determined mathematically. For several parameter varieties of the CEP distribution, the different kinds of failure rate function shapes are displayed in Figure 2.

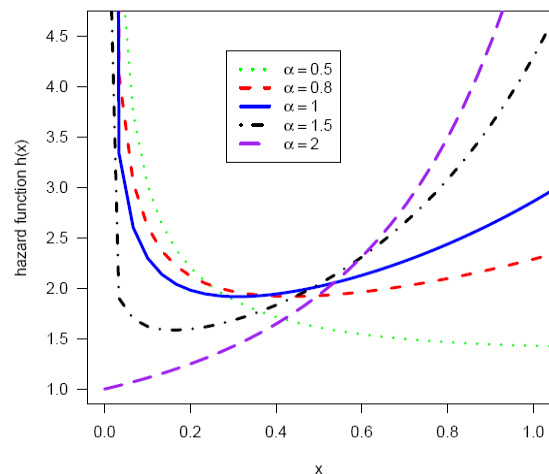


Figure 2. The  $h(x)$  of CP distribution for  $\theta = 1, \lambda = 1$  and varying values of  $\alpha$

### E. The Quantile Function:

The quantile function is given by,

$$x_p = \frac{1}{\lambda} \left\{ \ln \left[ 1 - \ln \left( 1 - p^{1/\theta} \right) \right] \right\}^{1/\alpha}; 0 < p < 1 \quad (6)$$

### F. The Random Deviate Generation

Consider  $U$  denote uniform random variable  $(0, 1)$ . Also, consider  $F(\cdot)$  as Cumulative Distribution Function whose inverse is  $F^{-1}(\cdot)$ . Next,  $F^{-1}(u)$  can be drawn using the distribution  $F(\cdot)$ . Hence, from  $CEP(\alpha, \lambda, \theta)$ , the random deviate can be generated as



$$x = \left\{ \log_e \left[ 1 - \log_e \left( 1 - u^{\frac{1}{\theta}} \right) \right] \right\}^{\frac{1}{\alpha}} \frac{1}{\lambda}; 1 > u > 0 \quad (7)$$

Here, u belongs to the U(0, 1) distribution

### III. MLE AND ASYMTOTIC INTERVALS

Here,  $CEP(\alpha, \lambda, \theta)$  distribution's MLE is reviewed. Then corresponding asymptotic characteristics to achieve estimated confidence intervals centered on maximum likelihood estimators is examined.

Consider a random sample  $\underline{x}=(x_1, \dots, x_n)$  from  $CEP(\alpha, \lambda, \theta)$ , having size 'n'; then the log-likelihood function  $\ell(\alpha, \lambda, \theta | \underline{x})$  is given by the following expression

$$\ell(\alpha, \lambda, \theta) = n \log_e \alpha + n \log_e \theta + n \alpha \log_e \lambda + \left( \sum_{i=1}^n \log_e x_i \right) (\alpha - 1) + P_3 + P_4$$

$$\text{where } P_3 = n + \sum_{i=1}^n (\lambda x_i)^\alpha - \sum_{i=1}^n e^{(\lambda x_i)^\alpha}$$

$$P_4 = \left\{ \sum_{i=1}^n \log_e \left[ 1 - e^{-\left\{ 1 - e^{-(\lambda x_i)^\alpha} \right\}} \right] \right\} (\theta - 1)$$

(8)

$$\frac{\partial \ell}{\partial \theta} = \sum_{i=1}^n \log_e \left[ 1 - \exp \left\{ 1 - \exp(\lambda x_i)^\alpha \right\} \right] + \frac{n}{\theta} = 0$$

(9)

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

$$+ \left( \sum_{i=1}^n x_i^\alpha \right) (\alpha \lambda^{\alpha - 1}) - \left( \sum_{i=1}^n x_i^\alpha e^{(\lambda x_i)^\alpha} \right) (\alpha \lambda^{\alpha - 1}) = 0$$

(10)

Using MLE's standard asymptotic normality, the  $(1 - \gamma)100\%$  fiducial intervals for  $\theta$ ,  $\lambda$  and  $\alpha$  are found with  $\text{var}(\hat{\theta})$ ,  $\text{var}(\hat{\lambda})$ ,  $\text{var}(\hat{\alpha})$  and determined using the the observed Fisher information matrix's inverse, i.e., the the second-order derivatives of the log-likelihood function's matrix inverse locally at  $\hat{\alpha}$ ,  $\hat{\lambda}$  and  $\hat{\theta}$ . Therefore, using the MLE's asymptotic normality, the formulation for the approximate  $100(1 - \gamma)\%$  fiducial intervals for  $\theta$ ,  $\lambda$ , and  $\alpha$  are shown below

$$\hat{\theta} \pm [\text{var}(\hat{\theta})]^{1/2} z_{\gamma/2}; \hat{\lambda} \pm [\text{var}(\hat{\lambda})]^{1/2} (z_{\gamma/2}) \text{ and} \\ \hat{\alpha} \pm [\text{var}(\hat{\alpha})]^{1/2} (z_{\gamma/2}) \quad (11)$$

Here  $z_{\gamma/2}$  denotes usual normal variate's upper percentile.

### IV. BAYESIAN MODEL FORMULATION

There are three ingredients of the Bayesian model:

- (i) Probability Model :  $f(x | \alpha, \lambda, \theta)$  ;
- (ii) The prior distribution of model parameters :  $p(\alpha, \lambda, \theta)$  ; and
- (iii) Data :  $\underline{x} = (x_1, \dots, x_n)$

The Bayesian model can be created by multiplying the prior distributions of model parameters  $\alpha$ ,  $\lambda$  and  $\theta$ , with the likelihood function  $L(\alpha, \lambda, \theta | \underline{x})$  for the provided data set  $\underline{x} = (x_1, \dots, x_n)$  to get the posterior distribution function using Bayes theorem.

For a provided data set  $\underline{x} = (x_1, \dots, x_n)$  from  $CEP(\alpha, \lambda, \theta)$ , likelihood function can be defined as

$$L(\theta, \lambda, \alpha | \underline{x}) = \lambda^{n\alpha} \theta^n \alpha^n \left( \exp \left( \lambda \sum_{i=1}^n x_i^\alpha \right) \right) \left( \prod_{i=1}^n x_i^{\alpha-1} \right) \\ \left\{ \exp \left( \sum_{i=1}^n \left[ 1 - \exp(\lambda x_i)^\alpha \right] \right) \right\} T_1$$

where

$$T_1 = \prod_{i=1}^n \left[ 1 - e^{-\left\{ 1 - e^{-(\lambda x_i)^\alpha} \right\}} \right]^{\theta-1}$$

The prior distribution of  $\alpha$ ,  $\lambda$  and  $\theta$  is be represented by  $p(\alpha, \lambda, \theta)$ . The expression used for defining the joint posterior is presented below:

$$p(\theta, \lambda, \alpha | \underline{x}) \propto p(\lambda, \alpha, \theta) L(\theta, \lambda, \alpha | \underline{x})$$

#### Prior distributions:

Denote independent uniform priors for  $\alpha \sim U(a_1, b_1)$ ,  $\lambda \sim U(a_2, b_2)$  and gamma prior for  $\theta \sim G(a_3, b_3)$  by

$$p(\alpha) = \frac{1}{b_1 - a_1}; \text{ where } b_1 > a > a_1$$

$$p(\lambda) = \frac{1}{b_2 - a_2}; \text{ where } b_2 > \lambda > a_2$$

and

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

#### Posterior distribution:

The posterior up to proportionality is produced after integrating prior distribution with the likelihood function using the Bayes' theorem, which is given below

$$p(\alpha, \lambda, \theta | \underline{x}) = \alpha^n \theta^n \lambda^{n\alpha} e^{\left( \lambda \sum_{i=1}^n x_i^\alpha \right)} \left\{ \prod_{i=1}^n x_i^{\alpha-1} \right\} T_2 T_3$$

$$\text{where } T_2 = \left\{ \exp \left( \sum_{i=1}^n \left[ 1 - \exp(\lambda x_i)^\alpha \right] \right) \right\}$$

$$T_3 = \left( \frac{1}{b_1 - a_1} \right) \left( \frac{1}{b_2 - a_2} \right) \frac{b_3^{a_3}}{\Gamma a_3} e^{-b_3 \theta} \theta^{a_3-1} T_1$$

or



# Bayesian Estimation and Prediction of Three-Parameter Complementary Exponential Power Distribution using MCMC Technique

$$p(\alpha, \lambda, \theta | \underline{x}) \propto \alpha^n \theta^{\alpha_3 + n - 1} \lambda^{n\alpha} e^{-b_3\theta} e^{\left( \lambda \sum_{i=1}^n x_i^\alpha \right)} \left\{ \prod_{i=1}^n x_i^{\alpha-1} \right\} e^{\left( \sum_{i=1}^n [1 - \exp(-\lambda x_i^\alpha)] \right)} T_1$$

The posterior can be very complex and inferences of close form doesn't seem probable. We suggest using MCMC for simulating samples from the posterior, in order to easily make inferences based on the sample. By operating a Markov chain that ultimately approaches to target distribution (called equilibrium or stationary), the MCMC methods creates samples, which for this situation, can be used as the posterior distribution  $p(\theta, \lambda, \alpha | \underline{x})$ . These chains can be built

in several ways, however every way, including the MCMC method developed by (Geman & Geman, 1984; Gelfand & Smith, 1990), which is known as Gibbs sampler, are particular cases of the Monte Carlo integrations over configuration space (Metropolis et al., 1953; Hastings, 1970), which is the general method.

## CEP probability model in OpenBUGS

Authors have the corresponding code Since OpenBUGS doesn't contain the CEP distribution, the integration of ReliaBUGS module, which is a CEP OpenBUGS subsystem, is necessary. In Component Pascal for Complementary Exponential Power, a module dCEP(alpha, lambda, theta) is developed to conduct complete Bayesian approach in OpenBUGS with the help of the technique given by (Thomas et al., 2006; Thomas, 2010; Kumar et al., 2010 & Lunn et al., 2013). Authors have the corresponding code

## V. REAL DATA ANALYSIS

We have used the set of data, obtained from real-life sample, to demonstrate the Classical and Bayesian analysis for CEP distribution. Data given below shows a sample of  $n=100$ , containing the values of carbon fibre's tensile strength (in GPa) (Nichols & Padgett, 2006) where they used the parametric bootstrap technique for building the limits of control chart to monitor this data.

0.39, 0.81, 0.85, 0.98, 1.08, 1.12, 1.17, 1.18, 1.22, 1.25, 1.36, 1.41, 1.47, 1.57, 1.57, 1.59, 1.59, 1.61, 1.61, 1.69, 1.69, 1.71, 1.73, 1.80, 1.84, 1.84, 1.87, 1.89, 1.92, 2.00, 2.03, 2.03, 2.05, 2.12, 2.17, 2.17, 2.17, 2.35, 2.38, 2.41, 2.43, 2.48, 2.48, 2.50, 2.53, 2.55, 2.55, 2.56, 2.59, 2.67, 2.73, 2.74, 2.76, 2.77, 2.79, 2.81, 2.81, 2.82, 2.83, 2.85, 2.87, 2.88, 2.93, 2.95, 2.96, 2.97, 2.97, 3.09, 3.11, 3.11, 3.15, 3.15, 3.19, 3.19, 3.22, 3.22, 3.27, 3.28, 3.31, 3.31, 3.33, 3.39, 3.39, 3.51, 3.56, 3.60, 3.65, 3.68, 3.68, 3.68, 3.70, 3.75, 4.20, 4.38, 4.42, 4.70, 4.90, 4.91, 5.08, 5.56

### (1) Classical Analysis

#### A. MLE Computation

The ML estimates and standard error are generated directly through the log-likelihood function's maximization, shown in equation (8) from the Newton-Raphson technique using the platform of R. In parentheses, the Maximum Likelihood Estimates and their respective Standard Errors

and the log-likelihood values (after maximization) are given by

$$\hat{\alpha} = 0.9237(0.2305), \hat{\lambda} = 0.3938(0.0716) \quad \hat{\theta} = 3.579(1.6234)$$

and  $\ell(\hat{\alpha}, \hat{\lambda}, \hat{\theta}) = -141.376$

## B. Model Validation

When parameters measured by the MLE method is 0.0646 and their p-value is 0.7983, we determine the Kolmogorov-Smirnov (KS) distance between the empirical and the fitted distribution functions to validate the model. In Fig. 3, we have graphed the fitted and empirical distribution function. From the higher p-value and fig. 3, we can observe the fitted CEP distribution is well fitted to the real data set that we have taken

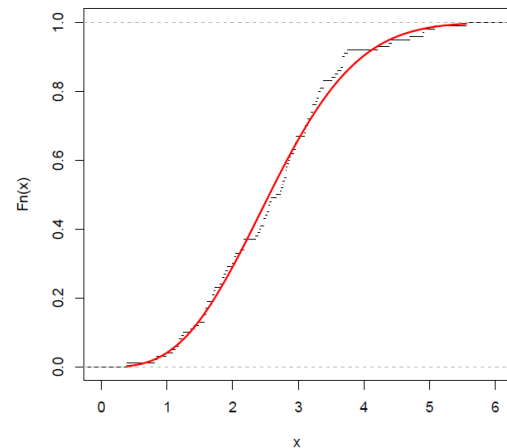


Fig 3. Empirical and Fitted distribution function's graph

Probability-probability (P-P) and Quantile-quantile (Q-Q) diagrams can be studied to further evaluate model validation. For the fitted model, P-P plots and Q-Q are illustrated in Fig. 4 and 5. We can see that, relative to the presented data, CEP distribution is good fitted.

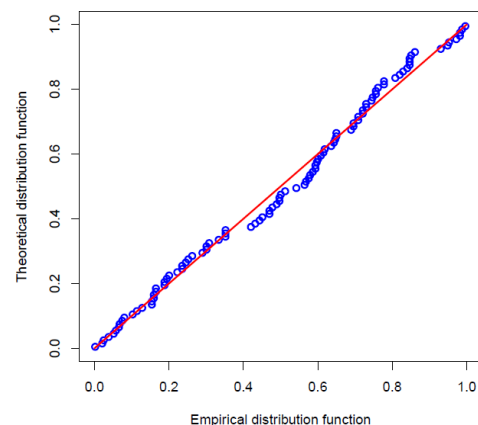


Fig.4. (P-P) plot with the use of MLEs as estimate

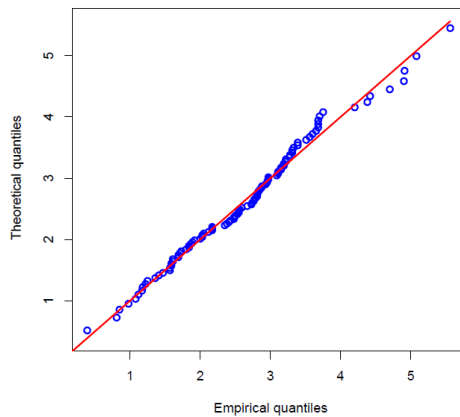


Fig.5. (Q-Q) plot with the use of MLEs as estimate

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

**(2) Bayesian Analysis**

Independent uniform priors for  $\alpha \sim U(a_1, b_1)$ ,  $\lambda \sim U(a_2, b_2)$  and gamma prior for  $\theta \sim G(a_3, b_3)$  are assumed with hyper parameter values

$$(a_1 = 0.0, b_1 = 5.0), (a_2 = 0.0, b_2 = 2.0) \text{ and } (a_3 = 0.001, b_3 = 0.001)$$

The model is operated for the creation of two Markov Chains with a variation in the parameters' starting points, at a length of 40,000. For the first chain, we picked the initial values ( $\alpha = 1.1, \lambda = 0.1, \theta = 1.5$ ) and for the second chain, we pick the initial value ( $\alpha = 0.5, \lambda = 0.5, \theta = 0.5$ ). Through the application of ergodic, as well as, trace mean plots, the convergence is tracked and after about 2000 observations, we observe that the Markov Chain approaches together. So to remove the impact of the initial values, the burn-in of 5000 samples is substantially adequate. Finally, sample sizes of 7000 can be produced from the posterior distribution by taking up every 5th outcome equally gapped for minimizing the auto-correlation between the generated deviates i.e. thin=5, beginning at 5001.

Thus, we get posterior samples  $(\alpha_1^{(j)}, \lambda_1^{(j)}, \theta_1^{(j)}) ; j = 1, \dots, 7000$  using chain 1 and  $(\alpha_2^{(j)}, \lambda_2^{(j)}, \theta_2^{(j)}) ; j = 1, \dots, 7000$  using chain 2.

We have used chain 1 for convergence diagnostics and posterior analysis.

**A. Convergence diagnostics**

Before performing inferences such as the examination of parameter estimates, we need to observe graphs of the parameter estimates' the sequential (dependent) realizations.

**History (Trace) plot**

Fig.6 demonstrates model parameters' sequential realizations. The parameter's graphs are similar to horizontal band, without showing long downward or upward trends, then it has been proven that the chain has converged and hence, samples are expected to draw from the equilibrium distribution.

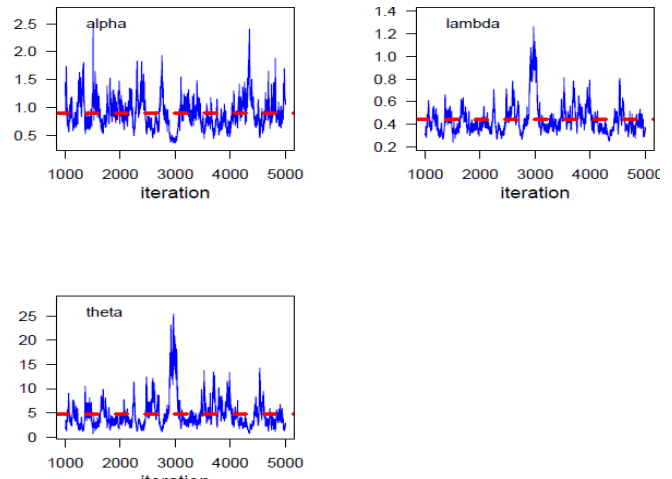


Fig.6  $\alpha, \lambda$  and  $\theta$ 's Sequential realization

**Ergodic mean Plot**

Fig. 7 demonstrates the ergodic mean's time series plot for the chain's every model parameter. The convergence pattern based on running means shows that the chain has converged and hence samples are likely to generate from the stationary distribution.

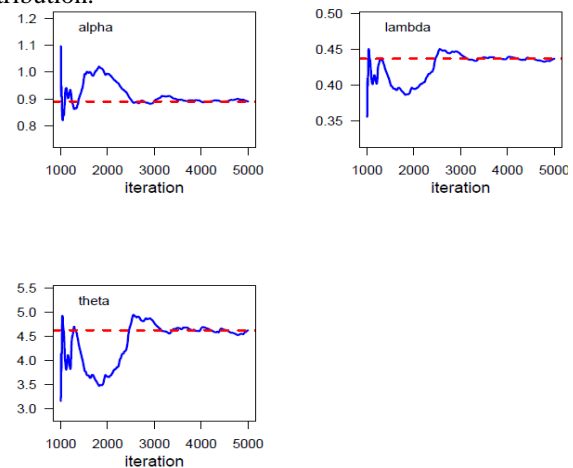


Fig.7.  $\alpha, \lambda$  and  $\theta$ 's plot of Ergodic mean

**Brooks-Gelman-Rubin(BGR) diagnostic:**

Convergence of the model is determined by using the BGR convergence diagnostics. (Gelman & Rubin,1992; Brooks & Gelman,1998). Using their methods, from Fig. 8, we conclude that chain has converged and hence posterior samples are likely to generate from the equilibrium distribution from which we can get posterior summary statistics.

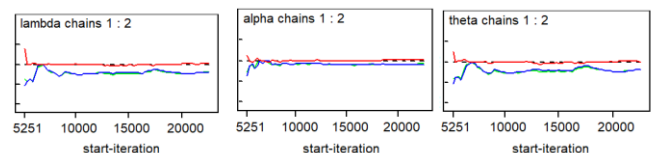


Fig. 8. The BGR plots for  $\alpha, \lambda$  and  $\theta$

Figure 8 clearly demonstrates the achievement of convergence. Thus, the summary statistics (posterior) can now be determined.

# Bayesian Estimation and Prediction of Three-Parameter Complementary Exponential Power Distribution using MCMC Technique

## B. Posterior Analysis

### (a) Numerical Summary

Centered on the MCMC samples of CEP distribution's posterior features, different quantities of interest as well as corresponding numerical values have been observed. Various numerical summaries of the CEP distribution's MCMC posterior samples are shown in Table 1.

**Table 1: Numerical summaries on the basis of MCMC samples under uniform and gamma priors**

Characteristics	alpha	lambda	theta
Mean	0.9345	0.4299	4.4753
Median	0.9089	0.3965	3.627
Mode	0.8691	0.3604	2.7112
Maximum	2.464	1.908	36.17
97.5th Percentile( $P_{97.5}$ )	1.5011	0.7698	12.3908
Third Quartile ( $Q_3$ )	1.101	0.4673	5.2893
First Quartile ( $Q_1$ )	0.7483	0.3491	2.595
2.5th Percentile( $P_{2.5}$ )	0.4863	0.2951	1.558
Standard Deviation	0.2607	0.1355	3.1915
Skewness	0.508	3.3479	3.4112
Minimum	0.303	0.2476	0.7943

**Highest probability density (HPD):** Based on the postulation of the Unimodal-Marginal-Posterior distribution, Chen & Shao(1999) defined the algorithm which is applied to calculate the HPD intervals.

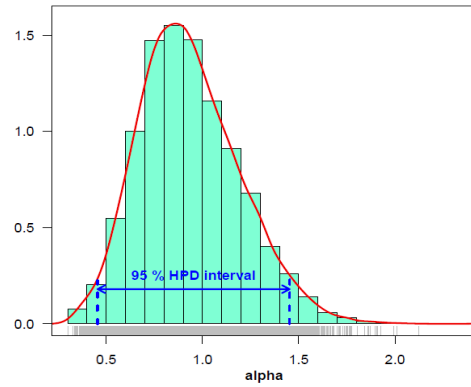
Table 2 given below displays Credible and HPD intervals for model parameters  $\alpha$ ,  $\lambda$  and  $\theta$

**Table 2. Credible and Highest Probability Density Intervals**

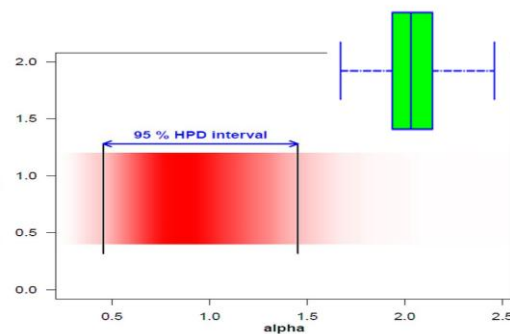
Parameter	HPD C. I.	Credible Interval
$\alpha$	(0.4547, 1.453)	(0.4863, 1.5010)
$\lambda$	(0.2758, 0.6595)	(0.2951, 0.7698)
$\theta$	(1.213, 10.1)	(1.5580, 12.3908)

### (b) Visual summary

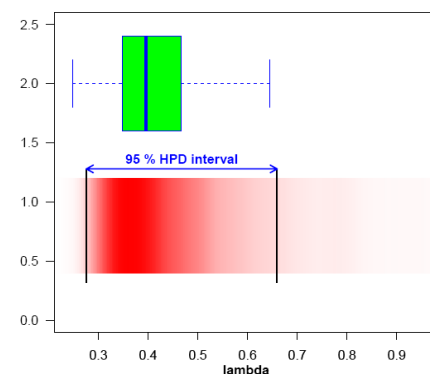
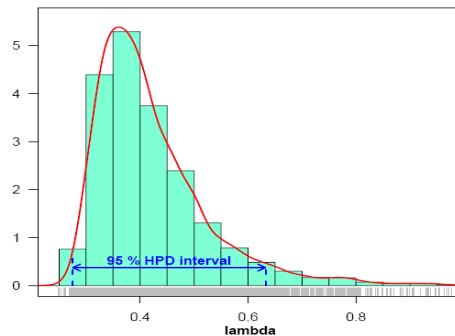
We presented the visual summary of the for the model parameters by constructing histogram, the box plot, rug plots, marginal posterior density estimate, and plot of density strip. The 95% HPD intervals have also been superimposed. These graphs give the nearly entire depiction of the model parameters' posterior uncertainty. The density strip is a shaded monochrome strip which demonstrates univariate distribution and its darkness at a point is in proportion to the PDF of the quantity at that point (Jackson, 2008). Histogram provides information into multi-modal behavior involvement, outliers of data, skewness, tail's behavior and can be contrasted to fundamental shapes related to symmetrical distributions



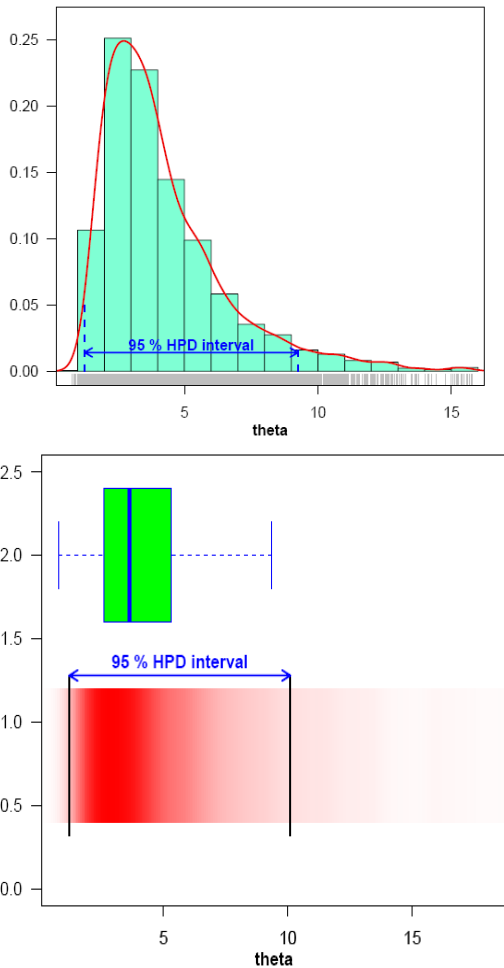
**Fig. 9. Histogram, marginal posterior density and 95% Highest probability density interval for  $\alpha$  centered on posterior sample**



**Fig. 10. density strip and Box plot of  $\alpha$ , centered on the posterior sample**



**Fig. 11 Upper panel: 95% Highest Probability Density interval, marginal posterior density, and Histogram; Lower panel:  $\lambda$ 's Box plot and density strip, centered on the posterior sample**

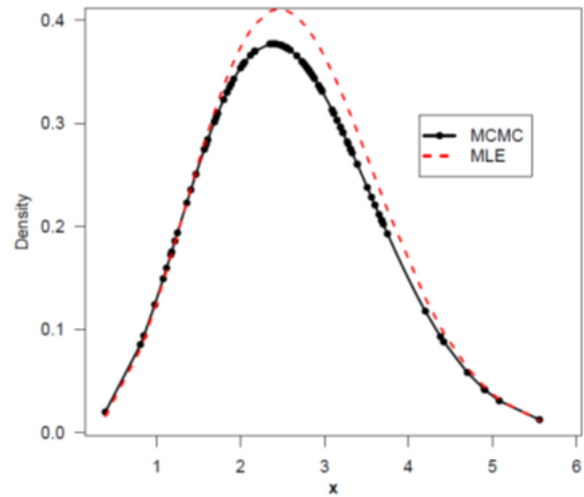


**Fig 12. Upper panel: 95% Highest Probability Density interval, marginal posterior density, Histogram Lower panel: density strip of  $\theta$  (bottom) and box plot (top), centered on the posterior sample**

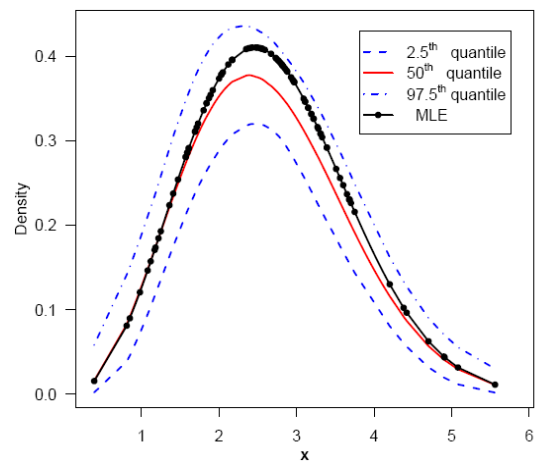
The marginal posterior density, histogram, and 95% percent Highest Probability Density interval are represented in Fig. 9, and for a the box whisker plot and density strip centered on the posterior sample are represented in Figure 10. In Figure 11 and 12, for  $\lambda$  and  $\theta$ , the related graphs have been illustrated. Kernel density estimates were graphed with the help of R, assuming the Gaussian kernel and the bandwidth values chosen properly.  $\alpha$  and  $\lambda$  indicate positive skewed distribution while  $\theta$  is observed to be symmetrical.

**Comparison with Maximum Likelihood Estimation**

We have illustrated three graphs to compare to Maximum likelihood Estimate. The PDF  $f(x; \hat{\alpha}, \hat{\lambda}, \hat{\theta})$ , with the help of Maximum Likelihood Estimates and Bayesian estimates, calculated by MCMC samples are illustrated in Fig. 13.  $\theta$  uses gamma prior, and  $\lambda$  and  $\alpha$  uses the uniform priors The figure shows that the estimates of the Maximum likelihood and Bayes are very similar and show good fit. It is possible to gain more evidence with this observation by examining Fig.14. We have illustrated the 97.5<sup>th</sup>, 50<sup>th</sup>, 2.5<sup>th</sup> quantiles of the approximate density in Figure 14, which can be viewed as a model fit assessment centered on the posterior sample.



**Fig.13 The PDF calculated with MCMC through Bayesian and Maximum-Likelihood estimations, computed via MCMC.**



**Fig 14 Density estimates**

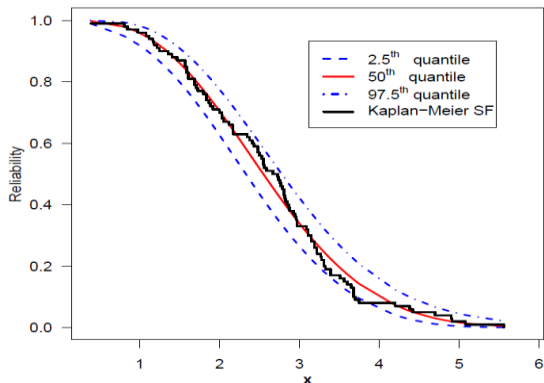
With the help of logical function *density()* in OpenBUGS, the PDF at every observed data point posterior samples with 7000 observations, has been calculated. The PDF associated with Maximum Likelihood Estimate is drawn with the help of "plug-in" parameter estimates. This indicates that the purposed model is reasonably good for the provided set of data.

**Reliability function estimate**

The reliability feature estimation utilizing posterior samples using the MCMC method is conducted to illustrate the proposed methodology's efficacy. Moreover Kaplan-reliability Meier's function calculation is used to add further significance in comparing the effectiveness.. Fig. 15 displays the estimations of reliability functions with the application of Bayes estimate derived from Markov chain Monte Carlo outcome and the empirical reliability function (Kaplan Meier SF).

Fig.15 demonstrates that MCMC estimate of reliability closely matches the empirical reliability estimates

# Bayesian Estimation and Prediction of Three-Parameter Complementary Exponential Power Distribution using MCMC Technique

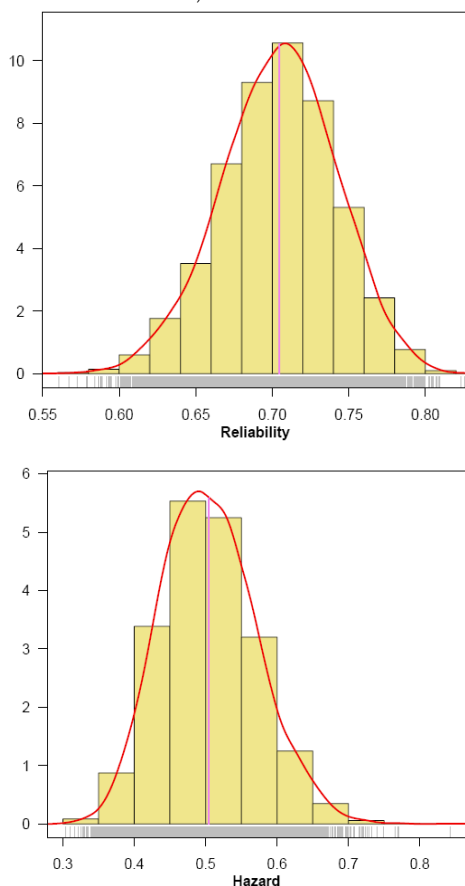


**Fig 15 estimate of Reliability function via Kaplan-Meier and MCMC estimate Reliability and Hazard Estimation at  $X_{(30)} : t = 2.0$**

The posterior samples entirely explains posterior uncertainty of the parameters  $\theta$ ,  $\lambda$  and  $\alpha$  using the posterior distribution's kernel estimate which is valid for any parameter function, such as hazard and reliability functions. Consider giving interval and point estimates at  $t=2.0$  ( $X_{30}$ ) for hazard and reliability functions. In OpenBUGS with the help of reliability() and logical function hrf() as per (Kumar et al., 2010), at mission time  $t=2.0$ , we calculated the hazard and reliability functions for 7000 posterior samples. We can calculate with the help of the equations in (4) and (5) at  $t=2.0$ .

$$R(x = 2.0; \alpha_1^{(j)}, \lambda_1^{(j)}, \theta_1^{(j)}); j = 1, \dots, 7000 \text{ And,}$$

$$h(x = 2.0; \alpha_1^{(j)}, \lambda_1^{(j)}, \theta_1^{(j)}); j = 1, \dots, 7000$$



**Fig. 16 Visual presentation of hazard(left side) and reliability(right side) at  $t=2$**

Fig. 16 illustrates the hazard (left side) and reliability functions (right side) marginal posterior density estimates as well as their histograms centered on sample sizes of 7000 with the help of the Gaussian Kernel. In Fig. 16 The estimates show that the marginal reliability distribution shows negative skewness while the hazard distribution shows positive skewness.

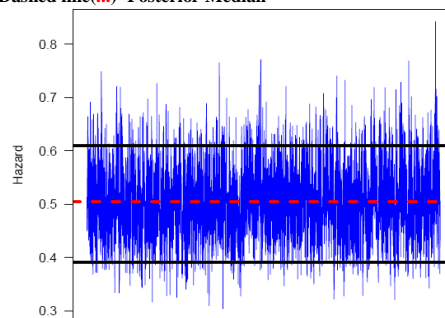
In Table 3, the posterior summary for Reliability and Hazard function at mission time  $t=2.0$  using MCMC method are presented. The MLEs can be calculated with the help of the invariance feature of the Maximum Likelihood estimates. Hence, MLEs  $\hat{R}(t = 2.0) = 0.7097$  and  $\hat{h}(t = 2.0) = 0.5264$

Figure 17 shows trash plot of 7000 chain values, with respective 90% Highest probability density and posterior median, for the reliability functions and hazard functions

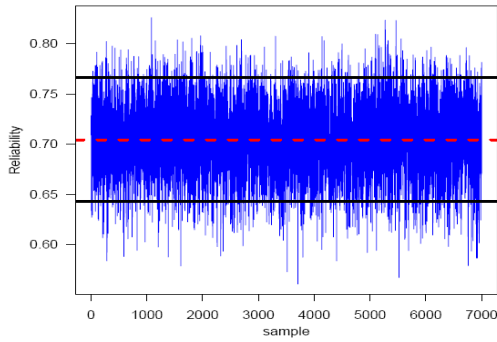
**Table 3: Hazard and Reliability functions' posterior summary when  $t=2.0$**

Statistics	Hazard	Reliability
Mean	0.5047	0.7043
Median	0.5013	0.7054
Mode	0.4893	0.7083
Standard Deviation	0.0669	0.0376
Skewness	0.2992	-0.1501
Maximum	0.8423	0.8259
97.5th Percentile( $P_{97.5}$ )	0.6451	0.7755
Third Quartile ( $Q_3$ )	0.5484	0.7299
First Quartile ( $Q_1$ )	0.4567	0.6791
2.5th Percentile( $P_{2.5}$ )	0.3846	0.6272
Minimum	0.3041	0.5607
95% HPD Credible Interval	(0.3770, 0.6356)	(0.6313, 0.7788)
95% Credible Interval	(0.3846, 0.6451)	(0.6272, 0.7755)

Solid lines(-)-90% Highest Probability density  
Dashed line(...)=Posterior Median





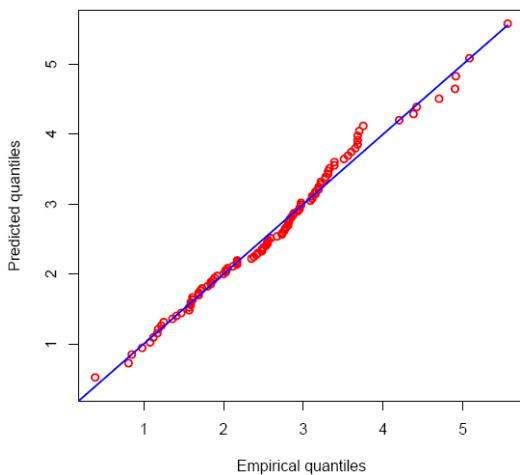


**Fig 17. Hazard function(upper panel) and Reliability function(Lower panel) at t=2 sec Markov Chain Monte Carlo Outcomes**

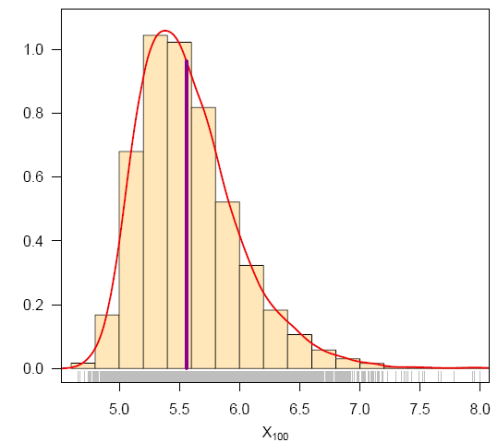
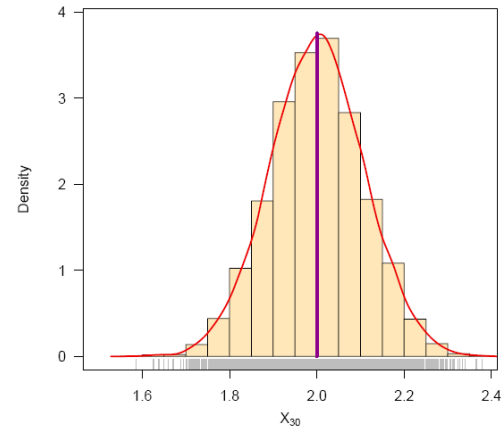
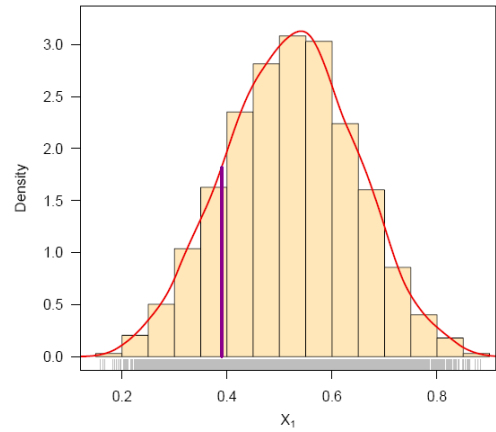
**VI. MODEL COMPATIBILITY**

**Posterior Predictive Checks:**

For examining model fit, (Gelman et al.,1996; Gelman et al.,2004)'s posterior predictive checks has proven to be one of the superior approach where the model is considered to have a good fit, when the data obtained by simulation from posterior predictive distribution in comparison is close to observed data. Figure 18 displays predicted vs. observed quantiles Q-Q plot of 1000 replication of each MCMC generated samples sizes 2000 from the posterior distribution for the model parameter  $\delta = (\alpha, \lambda, \theta)$ . In Fig.18 as simulated data is very close to the observed data of the predicted vs. observed Q-Q plot that, the conclusion that, complementary exponential power model is compatible with the given real dataset where equation of replication is  $x_i^{rep}; i = 1, \dots, 100$  for each, can be derived.



**Fig. 18. predictive vs empirical Q-Q plot based on 1000 MCMC generated samples**



**Fig 19. Estimates of density of largest  $(X_{(100)})$ , smallest  $(X_{(1)})$  and the  $X_{(30)}$  order future observation**

We can get more clarification on model compatibility by plotting density estimates of replicated future data. With the help of MCMC methodology, 1000 samples have been plotted from the posterior and then from simulation predictive future data of the same size is obtained. The posterior model's replicated future observation can be superimposed with the corresponding data obtained through observation. We have plotted density estimates of  $(X_{(1)}, X_{(100)} \text{ and } X_{(30)})$  in Figure 19. Corresponding observed values is represented by vertical lines. As shown in Fig 19, the posterior predictive distributions are centered over the observed data, which signifies that the model has a good fit. The Markov Chain Monte Carlo results of posterior characteristics are displayed in Table 4



# Bayesian Estimation and Prediction of Three-Parameter Complementary Exponential Power Distribution using MCMC Technique

Table 4: Posterior characteristic

	$X_{(1)}$	$X_{(30)}$	$X_{(100)}$
Observed	0.39	2.00	5.56
Mode	0.545	2.007	5.378
Mean	0.521	2.001	5.580
Median	0.606	2.072	5.803

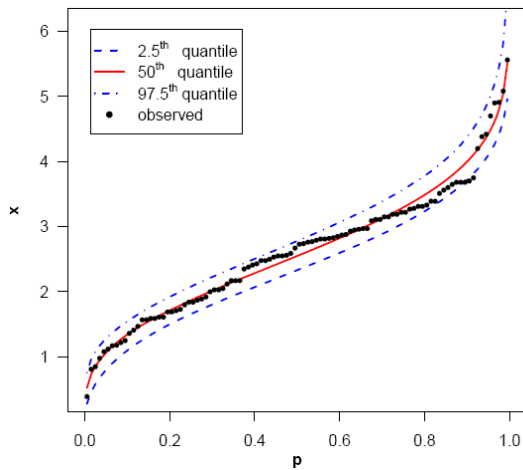


Fig. 20 Graphical posterior predictive check for the model adequacy

Figure 20 shows a model adequacy's posterior predictive check in a graphical manner, where, data observed is superimposed. As a whole, the simulation outcomes of the posterior predictive model show that the data demonstrates good fit with the model. Model fit evaluations are far too liberal only centered on posterior predictive checks, thus posterior predictive checks alone is not suitable for selection of model (Ntzoufras 2009)

## VII. CONCLUSION

In classical approach, we have plotted P-P and Q-Q diagrams for validation of model and found that CEP distribution is good fitted to the given real data set but this approach does not give entire pictures of posterior distribution. Bayesian approach uses prior knowledge about model parameters as well as for the available real data set and it gives posterior distribution's entire summaries for the model parameters on the basis of the posterior MCMC samples technique. We have applied exploratory data analysis procedures for the posterior analysis. We have performed convergence diagnostic using trace plot, ergodic mean plot and BGR diagnostic and found that Markov chain has converged and posterior MCMC samples are likely to draw from stationary posterior distribution. It has been shown that MCMC procedure is true for model parameters every function including hazard and reliability function. Credible intervals for model parameters, hazard and reliability function, etc. can be estimated. The posterior predictive check method has been applied to present model compatibility. The developed techniques are applied on a real data set under uniform and gamma priors. We have that Bayesian estimators are more efficient than classical estimators for any real data set. Therefore, the MCMC

method is used for performing complex Bayesian modeling relating to the CEP model

## REFERENCES

- Barriga, G.D.C., Neto, F.L. & Cancho, V.G. (2011). The complementary exponential power lifetime model, *Computational Statistics and Data Analysis*, 55(3), 1250-1259
- Brooks, S. & Gelman, A. (1998). General Methods for Monitoring Convergence of Iterative Simulations, *Journal of Computational and Graphical Statistics*, 7(4), 434-455.
- Carrasco, J.M., Ortega, E.M.M. & Cordeiro, G.M. (2008). A generalized modified Weibull distribution for lifetime modeling, *Computational Statistics and Data Analysis*, 53 (2), 450-462.
- Chen, M. H. & Shao, Q. M. (1999). Monte Carlo estimation of Bayesian credible intervals and HPD intervals, *Journal of Computational and Graphical Statistics*, 8(1), 69-92.
- 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(2), 155-161.
- Gelfand, A.E. and Smith, A.F.M.(1990). Sampling-based approach to calculating marginal densities. *Journal Statistical Association of the American*, 85, 398-409.
- Gelman, A., Carlin, J., Stern, H., & Rubin, D. (2004). *Bayesian Data Analysis*, Second Edition, London, Chapman & Hall.
- Gelman, A., Meng, X.L. & Stern, H.S. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Stat. Sin.*, 6, 733-807.
- Gelman, A. & Rubin, D.B. (1992). Inference from Iterative Simulation using Multiple Sequences, *Statistical Science*, 7(4), 457-472
- Geman, S., & Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images, *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 6, 721-741.
- Gupta, R. D. & Kundu, D. (1999). Generalized exponential distributions, *Australian and New Zealand Journal of Statistics*, 41(2), 173 - 188.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, 57, 97 -109.
- Jackson, C.H. (2008). Displaying uncertainty with shading, *The American Statistician*, 62(4), 340-347.
- Kumar, V., Ligges, U. & Thomas, A. (2010). ReliaBUGS User Manual: A subsystem in *OpenBUGS for some statistical models*, Version 1.0, OpenBUGS 3.2.1, <http://openbugs.info/w/Downloads/>
- Leemis, L. M. (1986). Lifetime distribution identities, *IEEE Transactions on Reliability*, 35, 170-174.
- Lunn, D.J., Jackson, C., Best, N., Andrew, A., & Spiegelhalter, D. (2013). *The BUGS Book: A Practical Introduction to Bayesian Analysis*, Chapman & Hall/CRC, London, UK.
- Marshall, A. W., & Olkin, I.(2007). *Life Distributions: Structure of Nonparametric, Semiparametric, and Parametric Families*, Springer, New York.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. & Teller, E.(1953). Equations of state calculations by fast computing machines, *Journal Chemical Physics*, 21, 1087-1091.
- Mudholkar, G.S. & Srivastava, D.K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data, *IEEE Transactions on Reliability*, 42(2), 299-302.
- Mudholkar, G.S., Srivastava, D.K., & Freimer, M. (1995). The exponentiated Weibull family—a reanalysis of the bus-motor-failure data, *Technometrics*, 37(4), 436-445.
- Mudholkar, G.S., Srivastava, D.K., & Kollia, G.D. (1996). A generalization of the Weibull distribution with application to the analysis of survival data, *Journal of the American Statistical Association*, 91(436), 1575-1583.
- Nadarajah, S. & Kotz, S. (2006). The exponentiated type distributions. *Acta Applicandae Mathematicae*, 92, 97-111.

24. Nichols, M.D. & Padgett, W.J. (2006), A bootstrap control chart for Weibull percentiles, *Quality and Reliability Engineering International*, 22, 141-151.
25. Ntzoufras, I. (2009). *Bayesian Modeling Using WinBUGS*, John Wiley & Sons, New York.
26. Pham, H. & Lai, C.D. (2007). On recent generalizations of the Weibull distribution, *IEEE Transactions on Reliability*, 56 (1), 454-458.
27. Rajarshi, S. & Rajarshi, M.B. (1988). Bathtub distributions: a review. *Communications in Statistics—Theory and Methods*, 17, 2597-2621.
28. Smith, R.M. & Bain, L.J. (1975). An exponential power life-testing distribution. *Communications in Statistics-Theory and Methods*, 4, 469-481.
29. Thomas, A. (2010). *OpenBUGS Developer Manual*, version 3.1.2, <http://www.openbugs.info/>
30. Thomas, A., O'Hara, B., Ligges, U. & Sturtz, S. (2006). Making BUGS Open, *R News*, 6, 12–17, URL <http://mathstat.helsinki.fi/openbugs/>.
31. Xie, M., Tang, Y. & Goh, T.N. (2002). A modified Weibull extension with bathtub-shaped failure rate function, *Reliability Engineering & System Safety*, 76(3), 279-285.

### AUTHORS PROFILE



**Arun Kumar Chaudhary** received his M.Sc in Statistics from Central Department of Statistics, Tribhuvan University and PhD in Statistics from D.D.U. Gorakhpur University, India. Currently working as Associate Professor in Department of Management Science(Statistics), Nepal Commerce Campus, Tribhuvan University, Nepal. He has got 26 years of teaching experience. He is associated with many other reputed colleges as visiting faculty. He is a life member of Nepal Statistical Association (NEPSA) and vice-chairperson of Nepal Statistical Society (NESS). He has authored more than 30 textbooks on statistics and mathematics.



**Vijay Kumar** received his M.Sc and Ph.D. in Statistics from D.D.U. Gorakhpur University. Currently working as Professor in Department of Mathematics and Statistics in DDU Gorakhpur University, Gorakhpur U.P. He has got 27 years of teaching/research experience. He is visiting Faculty of MaxPlanck-Institute, Germany. His main research interests are Reliability Engineering, Bayesian Statistics, and Actuarial Science. 9 PhD scholars have been awarded under his supervision and currently he is supervising 4 PhD scholars. He has published more than 40 national and international articles.