International Journal of Latest Trends in Engineering and Technology
Vol.(17)Issue(4), pp.028-046
DOI: http://dx.doi.org/10.21172/1.174.06
e-ISSN:2278-621X

THE ESTIMATION AND PREDICTION OF THE KWHC DISTRIBUTION UNDER BAYESIAN FRAMEWORK

Ramesh Kumar Joshi[1], Vijay Kumar[2]

Abstract- In this paper, we consider the estimation and prediction of the Kumaraswamy-Half-Cauchy(KwHC) distribution under Bayesian framework. The computation of MLEs, approximate variances, and confidence intervals of the parameters for the KwHC distribution based on a complete sample are performed. For the Bayesian analysis we have assumed the uniform prior for the scale parameter and gamma priors for shape parameters. Then the Gibbs sampling technique is applied to obtain the posterior samples using OpenBUGS software. Bayesian estimators of the parameters, posterior variances, and credible intervals are obtained using posterior samples. We have obtained the Bayes estimates of the hazard and reliability functions, and their probability intervals are also presented. We have applied the predictive check method to discuss the issue of model compatibility. All computational tools are developed in OpenBUGS and R software. A real data set is considered for illustration of the proposed Bayesian approach.

Keywords – Kumaraswamy-Half-Cauchy(KwHC) distribution, Bayesian framework, OpenBUGS

I. INTRODUCTION

Recently, some attempts have been made to define new families of distributions to extend well known models and at the same time provide great flexibility in modeling data in practice. Several techniques have been employed to form a larger family from an existing distribution by incorporating extra parameters.

Half-Cauchy(HC) distribution provides an alternative to inverse-Gamma distribution as a default prior for a scale parameter in Bayesian hierarchical models, in particular, when a proper prior is necessary [1]. It is obtained from the standard Cauchy distribution by folding the curve on the origin so that only positive values can be observed. As a heavy tailed distribution, the HC distribution has been used as an alternative to the exponential distribution.

The distribution introduced by Kumaraswamy [2] is quite flexible and has been little explored in the literature. It's cumulative distribution function (c.d.f.) has a simple form

$$F(x;\alpha,\beta) = 1-\left(1-x^{\alpha}\right)^{\beta} \;;0<x<1,$$

where $\alpha > 0$ and $\beta > 0$ are the two shape parameters. The corresponding density function is given by

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

which can be unimodal, increasing, decreasing or constant, depending on the parameter values. We consider the term "Kw" distribution to denote the Kumaraswamy distribution advocated the KW distribution as a generator since its quantile function takes a simple form. In his paper they highlighted several advantages over beta distribution: simple normalizing constant, simple explicit formula for the distribution and quantile functions[3]. It does not involve any special functions for quantile function and random variate generation.

In recent years, new classes of models have been proposed based on modifications of the existing one or two parameter models. Adding one or more parameters to a distribution makes it richer and more flexible for

[1] *Trichandra Multiple Campus, Tribhuvan University, Kathmandu, Nepal*
[2] *Department of Mathematics and Statistics, DDU Gorakhpur University, Gorakhpur-273009, U.P., INDIA*

modeling data. There are different ways for adding parameter(s) to a distribution. Such an addition of parameters makes the resulting distribution richer and more flexible for modeling data.

If G denotes the c.d.f. of a random variable, [4] defined the $KW-G$ distribution given by

$$F(x;\alpha,\beta)=1-\left(1-G(x)^{\alpha}\right)^{\beta},$$

where $\alpha>0$ and $\beta>0$ are two additional parameters whose role is to introduce skewness and to vary tail weights. Because of its tractable distribution function the KW-G distribution can be used quite effectively even if the data are censored. Correspondingly, the density function of this family of distributions has a very simple form

$$f(x;\alpha,\beta)=\alpha\beta\, g(x)G(x)^{\alpha-1}\left(1-G(x)^{\alpha}\right)^{\beta-1}.$$

So, the KW-G distribution is obtained by adding two shape parameters α and β to the G distribution. It contains distributions with unimodal and bathtub shaped hazard rate functions. Clearly, the KW density function is a particular case with $G(x)=x$. Let X follows a Half-Cauchy(HC) distribution with parameter θ with the c.d.f

$$G(x;\theta)=\frac{2}{\pi}\tan^{-1}\left(\frac{x}{\theta}\right) \quad ;\theta>0,x>0,$$

the corresponding p.d.f. is given by

$$g(x;\theta)=\frac{2\beta(x/\theta)}{\pi x\left(1+(x/\theta)^2\right)} \quad ;\theta>0,x>0.$$

A generalization of the HC distribution, named Beta Half-Cauchy distribution obtained through beta transformation was introduced by [5]. Few more generalizations of the HC distribution exist in the literature, namely, KwHC by [6] and Marshall-Olkin half-Cauchy (MOHC) by [7]. The Gamma-Half-Cauchy distribution was introduced by[8], [9]. The exponentiated half-Cauchy distribution may be considered as a sub-model of Beta Half-Cauchy or Gamma-Half-Cauchy distributions.

Ghosh defined and studied KwHC distribution as special case of the Kumaraswamy-G family of distributions by taking G as Half-Cauchy distribution [6].

The c.d.f of the KwHCcan be written as

$$F(x;\alpha,\beta,\theta)=1-\left[1-\left\{\frac{2}{\pi}\arctan\left(\frac{x}{\theta}\right)\right\}^{\alpha}\right]^{\beta}; x\geq 0$$

where $\alpha>0$ and $\beta>0$ are shape parameters and $\theta>0$ is the scale parameter. When $\alpha=1$ and $\beta=1$, then the KW-Half-Cauchy distribution reduces to the Half-Cauchy distribution with parameter θ. It also the given the characterization of the KW-Half-Cauchy distribution which establishes the relation between KW-Half-Cauchy and uniform distribution.

If a random variable U follows a uniform $(0,1)$ distribution with parameters then

$$x=\theta\ tan\left(\frac{\pi}{2}\left\{1-\left(1-(1-u)^{1/\beta}\right)^{1/\alpha}\right\}\right)$$

follows the KW-Half-Cauchy with parameters θ,α and β .

The KwHC distribution can be approximately symmetric, right-skewed or left skewed. The KwHC distribution has a thicker right tail than the other well-known distributions and is thus appropriate for positively skewed data. Also, the KwHC hazard function can be a decreasing failure rate or upside down bathtub shapes.

We do not find any systematic study classical as well as Bayesian on KwHC distribution in the literature. In recent decades, the Bayesian viewpoint has received frequent attention for analyzing failure data and other time-to-event data, and has been often proposed as a valid alternative to traditional statistical perspectives.

In this paper, we consider the estimation and prediction of the KwHC distribution under Bayesian framework. The computation of MLEs, approximate variances, and confidence intervals of the parameters for the KwHC distribution based on a complete sample are performed. For the Bayesian analysis we have assumed the uniform prior for the scale parameter and gamma priors for shape parameters[10]. Then the Gibbs sampling technique is applied to obtain the posterior samples using OpenBUGS software. Bayesian estimators of the parameters, posterior variances, and credible intervals are obtained using posterior sampls. We have obtained the Bayes estimates of the hazard and reliability functions, and their probability intervals are also presented. We have applied the predictive check method to discuss the issue of model compatibility[11]. All computational tools are developed in OpenBUGS and R software A real data set is considered for illustration of the proposed Bayesian approach.

II. Model Analysis of KwHC Distribution

The Cumulative distribution function of KwHC distribution with three parameters is given by

$$F\left(x;\alpha,\beta,\theta\right)=1-\left[1-\left\{\frac{2}{\pi}\arctan\left(\frac{x}{\theta}\right)\right\}^{\alpha}\right]^{\beta};\ x\geq 0, \tag{1}$$

where $\alpha>0$, $\beta>0$ and $\theta>0$ are the parameters. The KwHC distribution will be denoted by $KwHC(\alpha,\beta,\theta)$. The probability density function is given by

$$f(x;\alpha,\beta,\theta)=\frac{\alpha\beta}{\theta}\left(\frac{2}{\pi}\right)^{\alpha}\left\{\arctan\left(\frac{x}{\theta}\right)\right\}^{\alpha-1}\left[1-\left\{\frac{2}{\pi}\arctan\left(\frac{x}{\theta}\right)\right\}^{\alpha}\right]^{\beta-1}\left[1+\left(\frac{x}{\theta}\right)^{2}\right]^{-1}, \tag{2}$$

where $(\alpha,\beta,\theta)>0$ and $x\geq 0$.

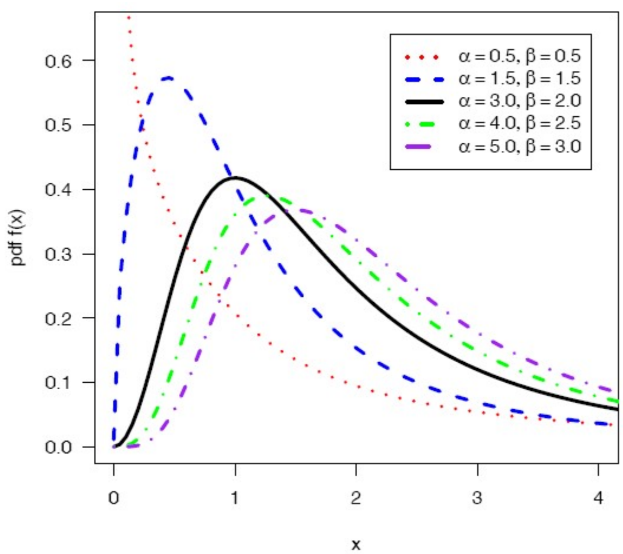


Figure 1 The probability density function of KwHC distribution for $\theta=1$ and different values of α and β.

The R functions dkw.halfCauchy() and pkw.halfCauchy() given in [12]. It can be used for the computation of pdf and cdf respectively. Some of the typical KwHC density functions for different values of α and β for $\theta=1$ are depicted in Figure1. It is clear from this figure that the density function of the KwHC distribution can take different shapes.

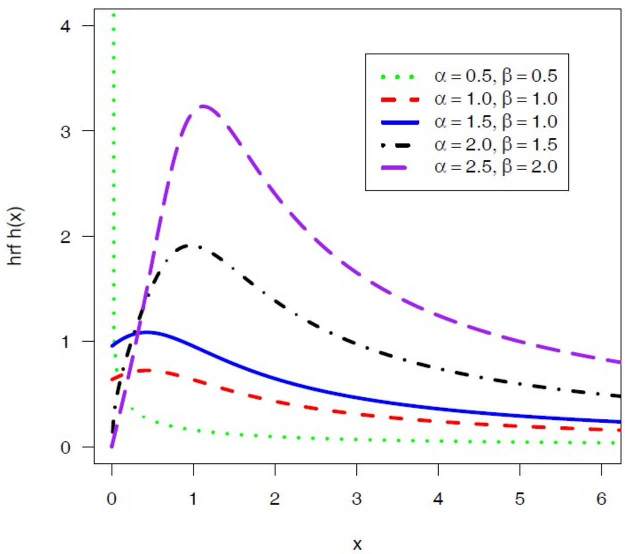


Figure 2 The hazard rate function of KwHC distribution for $\theta = 1$ and different values of α and β.

The reliability/survival function of KwHC distribution is

$$R(x;\alpha,\beta,\theta) = \left[1 - \left\{\frac{2}{\pi}\arctan\left(\frac{x}{\theta}\right)\right\}^{\alpha}\right]^{\beta}. \qquad (.3)$$

The R function skw.halfCauchy() computes the reliability/ survival function[12].

The hazard rate function of KwHC distribution is

$$h(x;\alpha,\beta,\theta) = \frac{\frac{\alpha\beta}{\theta}\frac{2^{\alpha}}{\pi^{\alpha}}\left\{\arctan\left(\frac{x}{\theta}\right)\right\}^{\alpha-1}\left[1+\left(\frac{x}{\theta}\right)^{2}\right]^{-1}}{\left[1 - \left\{\frac{2}{\pi}\arctan\left(\frac{x}{\theta}\right)\right\}^{\alpha}\right]}. \qquad (4)$$

Some of the typical KwHC hazard functions for different values of α and β for $\theta = 1$ are depicted in Figure 2. It is clear from this Figure that the hazard function of the KwHC distribution can take different shapes including the unimodel (upside down bathtub). The associated R function hkw.halfCauchy() [12]. Figure 2 exhibits the different hazard rate functions for KwHC distribution.

The quantile function of KwHC distribution is given by

$$x_p = \theta\tan\left[\frac{\pi}{2}\left\{1-(1-p)^{1/\beta}\right\}^{1/\alpha}\right] \quad ; 0 < p < 1. \qquad (5)$$

For the computation of quantiles the R function qkw.halfCauchy() can be used[12].

The random deviate can be generated from $KwHC(\alpha,\beta,\theta)$ by

$$x = \theta\tan\left[\frac{\pi}{2}\left\{1-(1-u)^{1/\beta}\right\}^{1/\alpha}\right] \quad ; 0 < u < 1, \qquad (6)$$

where u has the $U(0,1)$ distribution. The R function rkw.halfCauchy() generates the random deviate from $KwHC(\alpha,\beta,\theta)$ [12].

For model choice based on information criterion, the values of AIC and BIC can be computed using the R function abic.kw.halfCauchy() [12].

III. Maximum Likelihood Estimation (MLE) and Information Matrix

In this section, we discuss the maximum likelihood estimators (MLE's) of the KwHC distribution and their asymptotic properties to obtain approximate confidence intervals based on MLE's.

Let $\underline{x} = (x_1, \ldots, x_n)$ be a random sample of size *n* from $KwHC(\alpha,\beta,\theta)$, then the log-likelihood function $\ell(\alpha,\beta,\theta \mid \underline{x})$ can be written as

$$\begin{aligned}\ell(\alpha,\beta,\theta \mid \underline{x}) &= n\log(\alpha\beta) + n\alpha\log\left(\frac{2}{\pi}\right) - n\log\theta + (\alpha-1)\sum_{i=1}^{n}\log\left\{arctan\left(\frac{x_i}{\theta}\right)\right\} \\ &+ (\beta-1)\sum_{i=1}^{n}\log\left[1-\left\{\frac{2}{\pi}arctan\left(\frac{x_i}{\theta}\right)\right\}^{\alpha}\right] - \sum_{i=1}^{n}\log\left[1+\left(\frac{x_i}{\theta}\right)^2\right]\end{aligned} \tag{7}$$

Differentiating with respect to α, β and θ, we have

$$\begin{aligned}\frac{\partial\ell}{\partial\alpha} &= \frac{n}{\alpha} + n\log\left(\frac{2}{\pi}\right) + \sum_{i=1}^{n}\log\left\{arctan(x_i/\theta)\right\} \\ &-(\beta-1)\left(\frac{2}{\pi}\right)^{\alpha}\sum_{i=1}^{n}\frac{\left\{arctan(x_i/\theta)\right\}^{\alpha}\log\left\{(2/\pi)arctan(x_i/\theta)\right\}}{\left[1-\left\{(2/\pi)arctan(x_i/\theta)\right\}^{\alpha}\right]} = 0\end{aligned}$$

$$\frac{\partial\ell}{\partial\beta} = \frac{n}{\beta} + \sum_{i=1}^{n}\log\left[1-\left\{\frac{2}{\pi}arctan\left(\frac{x_i}{\theta}\right)\right\}^{\alpha}\right] = 0$$

and

$$\begin{aligned}\frac{\partial\ell}{\partial\theta} &= -\frac{n}{\theta} - \frac{(\alpha-1)}{\theta^2}\sum_{i=1}^{n}x_i\left[arctan(x_i/\theta)\right]^{-1}\left[1+(x_i/\theta)^2\right]^{-1} + \frac{2}{\theta^3}\sum_{i=1}^{n}x_i^2\left[1+(x_i/\theta)^2\right]^{-1} \\ &+\frac{\alpha(\beta-1)}{\theta^2}\left(\frac{2}{\pi}\right)^{\alpha}\sum_{i=1}^{n}\frac{x_i\left\{arctan(x_i/\theta)\right\}^{\alpha-1}\left\{1+(x_i/\theta)^2\right\}^{-1}}{\left[1-\left\{(2/\pi)arctan(x_i/\theta)\right\}^{\alpha}\right]} = 0\end{aligned}$$

Therefore, to obtain the MLE's of α, β and θ, we can maximize (7) directly with respect to α, β and θ or setting these equations to zero and solving them simultaneously yield the maximum likelihood estimates (MLEs) of the model parameters. Numerical methods can be used to obtain the ML estimates of the parameters. For example, the Newton-Raphson iterative technique could be applied to solve the likelihood equations numerically.

Let us denote the parameter vector by $\underline{\delta} = (\alpha,\beta,\theta)$ and the corresponding MLE of $\underline{\delta}$ as $\hat{\underline{\delta}} = (\hat{\alpha},\hat{\beta},\hat{\theta})$, then the asymptotic normality results in

$$\left(\hat{\underline{\delta}} - \underline{\delta}\right) \rightarrow N_3\left(0, \left(I(\underline{\delta})\right)^{-1}\right) \tag{8}$$

where $I(\underline{\delta})$ is the Fisher's information matrix given by

$$I(\underline{\delta}) = -\begin{pmatrix} E\left(\frac{\partial^2 \ell}{\partial \alpha^2}\right) & E\left(\frac{\partial^2 \ell}{\partial \alpha\, \partial \beta}\right) & E\left(\frac{\partial^2 \ell}{\partial \alpha\, \partial \theta}\right) \\ E\left(\frac{\partial^2 \ell}{\partial \beta\, \partial \alpha}\right) & E\left(\frac{\partial^2 \ell}{\partial \beta^2}\right) & E\left(\frac{\partial^2 \ell}{\partial \beta\, \partial \theta}\right) \\ E\left(\frac{\partial^2 \ell}{\partial \theta\, \partial \alpha}\right) & E\left(\frac{\partial^2 \ell}{\partial \theta\, \partial \beta}\right) & E\left(\frac{\partial^2 \ell}{\partial \theta^2}\right) \end{pmatrix} \tag{9}$$

In practice, it is useless that the MLE has asymptotic variance $\left(I(\underline{\delta})\right)^{-1}$ because we do not know $\underline{\delta}$. Hence, we approximate the asymptotic variance by "plugging in" the estimated value of the parameters. The common procedure is to use observed Fisher information matrix $O(\hat{\underline{\delta}})$ (as an estimate of the information matrix $I(\underline{\delta})$) given by

$$O(\hat{\underline{\delta}}) = -\begin{pmatrix} \frac{\partial^2 \ell}{\partial \alpha^2} & \frac{\partial^2 \ell}{\partial \alpha\, \partial \beta} & \frac{\partial^2 \ell}{\partial \alpha\, \partial \theta} \\ \frac{\partial^2 \ell}{\partial \beta\, \partial \alpha} & \frac{\partial^2 \ell}{\partial \beta^2} & \frac{\partial^2 \ell}{\partial \beta\, \partial \theta} \\ \frac{\partial^2 \ell}{\partial \theta\, \partial \alpha} & \frac{\partial^2 \ell}{\partial \theta\, \partial \beta} & \frac{\partial^2 \ell}{\partial \theta^2} \end{pmatrix}\Bigg|_{(\hat{\alpha}, \hat{\beta}, \hat{\theta})} = -H(\underline{\delta})\Big|_{\underline{\delta}=\hat{\underline{\delta}}} \tag{10}$$

where H is the Hessian matrix, $\underline{\delta} = (\alpha, \beta, \theta)$ and $\hat{\underline{\delta}} = \left(\hat{\alpha}, \hat{\beta}, \hat{\theta}\right)$. The Newton-Raphson algorithm to maximize the likelihood produces the observed information matrix. Therefore, the variance-covariance matrix is given by

$$\left(-H(\underline{\delta})\Big|_{\underline{\delta}=\hat{\underline{\delta}}}\right)^{-1} = \begin{pmatrix} var(\hat{\alpha}) & \operatorname{cov}(\hat{\alpha}, \hat{\beta}) & \operatorname{cov}(\hat{\alpha}, \hat{\theta}) \\ \operatorname{cov}(\hat{\beta}, \hat{\alpha}) & var(\hat{\beta}) & \operatorname{cov}(\hat{\beta}, \hat{\theta}) \\ \operatorname{cov}(\hat{\theta}, \hat{\alpha}) & \operatorname{cov}(\hat{\theta}, \hat{\beta}) & var(\hat{\theta}) \end{pmatrix} \tag{11}$$

Hence, from the asymptotic normality of MLEs, approximate $100(1-\gamma)\%$ confidence intervals for α, β and θ can be constructed as

$$\hat{\alpha} \pm z_{\gamma/2}\sqrt{var(\hat{\alpha})} \quad \hat{\beta} \pm z_{\gamma/2}\sqrt{var(\hat{\beta})} \text{ and } \hat{\theta} \pm z_{\gamma/2}\sqrt{var(\hat{\theta})} \tag{12}$$

where $z_{\gamma/2}$ is the upper percentile of standard normal variate.

IV. Bayesian Model Formulation

The Bayesian model is constructed by specifying the prior distributions for the model parameters α, β and θ, and then multiplying with the likelihood function to obtain the posterior distribution function.

- Probability Model : $f(x \mid \alpha, \beta, \theta)$
- Prior distribution : $p(\alpha, \beta, \theta)$
- Data : $\underline{x} = (x_1, \ldots, x_n)$

Given a set of data $\underline{x} = (x_1, \ldots, x_n)$, the likelihood function is

$$L(\alpha, \beta, \theta \mid \underline{x}) = \left(\frac{\alpha\beta}{\theta}\right)^n \left(\frac{2}{\pi}\right)^{n\alpha} \left(\prod_{i=1}^{n}\left[1-\left\{(2/\pi)\arctan\left(x_i/\theta\right)\right\}^{\alpha}\right]^{\beta-1}\right)$$
$$\left(\prod_{i=1}^{n}\left\{\arctan\left(x_i/\theta\right)\right\}^{\alpha-1}\right)\left(\prod_{i=1}^{n}\left\{1+\left(x_i/\theta\right)^2\right\}^{-1}\right)$$

Let $p(\alpha,\beta,\theta)$ denotes the joint prior distribution of α , β and θ . The joint posterior is

$$p(\alpha,\beta,\theta \mid \underline{x}) = L(\alpha,\beta,\theta \mid \underline{x})\; p(\alpha)\; p(\beta)\; p(\theta)$$

Prior distributions:

We assume the independent gamma priors for $\alpha \sim G(a_1, b_1)$ and $\beta \sim G(a_2, b_2)$, and uniform prior for $\theta \sim U(a_3, b_3)$ as

$$p(\alpha) = \frac{b_1^{a1}}{\Gamma(a_1)}\; \alpha^{a_1-1}\, e^{-b_1\alpha} \quad ; \alpha > 0, (a_1, b_1) > 0 .,$$

$$p(\beta) = \frac{b_2^{a2}}{\Gamma(a_2)}\; \beta^{a_2-1}\, e^{-b_2\beta} \quad ; \beta > 0, (a_2, b_2) > 0 ,$$

and

$$p(\theta) = \frac{1}{b_3 - a_3} \quad ; \; a_3 < \theta < b_3 .$$

Posterior distribution:

Combining the likelihood function with the prior via Bayes' theorem yields the posterior up to proportionality as

$$p(\alpha,\beta,\theta \mid \underline{x}) \propto \alpha^{a_1+n-1} \beta^{a_2+n-1} \theta^{-n} \left(2/\pi\right)^{n\alpha} \exp\left(-b_1\alpha - b_2\beta\right) T_1\, T_2\, T_3$$

where

$$T_1 = \prod_{i=1}^{n} \left[1 - \left\{(2/\pi)\, arctan\left(x_i/\theta\right)\right\}^{\alpha}\right]^{\beta-1} \quad T_2 = \prod_{i=1}^{n} \left\{arctan\left(x_i/\theta\right)\right\}^{\alpha-1} \text{ and } \quad T_3 = \prod_{i=1}^{n} \left\{1 + \left(x_i/\theta\right)^2\right\}^{-1} .$$

The posterior is obviously complicated and no close form inferences appear possible. We, therefore, propose to consider MCMC methods to simulate samples from the posterior so that sample-based inferences can be easily drawn.

Markov chain Monte Carlo draws samples by running a cleverly constructed Markov chain that eventually converges to the target distribution (called stationary or equilibrium) which, in our case, is the posterior distribution $p(\alpha,\beta,\theta \mid \underline{x})$.

There are many ways of constructing these chains, but all of them, including the Gibbs sampler .

Gibbs Sampler : Algorithm

For Gibbs sampler implementation, the full conditionals for α, β and θ up to proportionality can be specified as

(i) Full conditional distribution of the parameter α for given β , θ and $\underline{x}$

$$p(\alpha \mid \beta,\theta,\underline{x}) \propto \alpha^{a_1+n-1} \left(2/\pi\right)^{n\alpha} \exp\left(-b_1\alpha\right) T_1\, T_2$$

(ii) Full conditional distribution of the parameter β for given α , θ and $\underline{x}$

$$p(\beta \mid \alpha,\theta,\underline{x}) \propto \beta^{a_2+n-1} \; \exp\left(-b_2\beta\right) T_1$$

(iii) Full conditional distribution of the parameter θ for given α , β and $\underline{x}$

$$p(\theta \mid \alpha,\beta,\underline{x}) \propto \theta^{-n}\, T_1\, T_2\, T_3$$

We shall use OpenBUGS software to obtain posterior samples. As the KwHC distribution is not available in OpenBUGS, it requires incorporation of a module in *ReliaBUGS*[13] and [14] subsystem of OpenBUGS for KwHC distribution. A module dkwh.cauchy(alpha, beta, theta) is written in Component Pascal for KwHC to perform full Bayesian analysis in OpenBUGS using the method described in [15]. It is important to note that this module can be used for any set of suitable priors of the model parameters. Almost all aspects of the model in Bayesian framework can be studied using the developed module dkwh.cauchy(alpha, beta, theta)

Gibbs Sampler : Implementation

1. Select an initial value $\underline{\delta}^{(0)} = \left(\alpha^{(0)}, \beta^{(0)}, \theta^{(0)}\right)$ to start the chain.

2. Suppose at the i^{th}-step, $\underline{\delta} = (\alpha, \beta, \theta)$ takes the value $\underline{\delta}^{(i)} = \left(\alpha^{(i)}, \beta^{(i)}, \theta^{(i)}\right)$ then from full conditionals, we generate

   $\alpha^{(i+1)}$ from $p\left(\alpha \mid \beta^{(i)}, \theta^{(i)}, \underline{x}\right)$

   $\beta^{(i+1)}$ from $p\left(\beta \mid \alpha^{(i+1)}, \theta^{(i)}, \underline{x}\right)$ and

   $\theta^{(i+1)}$ from $p\left(\theta \mid \alpha^{(i+1)}, \beta^{(i+1)}, \underline{x}\right)$.

3. This completes a transition from $\underline{\delta}^{(i)}$ to $\underline{\delta}^{(i+1)}$
4. Repeat Step 2, N times.

MCMC output : Posterior sample

It is well known that rapid convergence is facilitated by choosing appropriate starting values. In order to guarantee the convergence and to remove the effect of the selection of initial value, the first 'B' simulated variates are discarded. Also to reduce the effect of autocorrelation, select a sampling lag $L > 1$ after which the corresponding autocorrelation are low. Consider $\left(\underline{\delta}^{(1)}, \ldots, \underline{\delta}^{(j)}, \ldots, \underline{\delta}^{(M)}\right)$ as the MCMC output (posterior sample) for the posterior analysis

$$\underline{\delta}^{(j)} = \left(\alpha^{(j)}, \beta^{(j)}, \theta^{(j)}\right); \; j = 1, 2, \ldots, M \, .$$

Thus, MCMC output is referred as the sample after removing the initial iterations (produced during the burn-in period) and considering the appropriate *lag,* which can be used to develop the Bayesian inference..

The Bayes estimates of $\underline{\delta} = (\alpha, \beta, \theta)$, under the square error loss (SEL) function, are given by

$$\hat{\alpha} = \frac{1}{M}\sum_{j=1}^{M}\alpha^{(j)} \; ; \quad \hat{\beta} = \frac{1}{M}\sum_{j=1}^{M}\beta^{(j)} \; ; \quad \hat{\theta} = \frac{1}{M}\sum_{j=1}^{M}\theta^{(j)}$$

The Bayes estimates under absolute and zero-one loss functions are posterior median and mode, respectively.

V. Data Analysis

Data Set : The real data set considered for illustration of the proposed methodology on KwHC distribution. The real data set represents the remission times (in months) of a random sample of 128 bladder cancer patients, [16]:

I. Classical Analysis

The estimation of the parameter of proposed model is obtained by the method of maximum likelihood(ML) estimation. To check the validity of the model, we compute the Kolmogorov-Smirnov (KS) distance between the empirical distribution function and the fitted distribution function when the parameters are obtained by method of maximum likelihood. The following graphical methods are also used for suitability of the model under consideration: (a) Quantile-Quantile(QQ) plot and (b) Probability–Probability (PP) plot.

Computation of MLE

The maximum likelihood estimates (MLEs) are obtained by direct maximization of the log-likelihood function $\ell(\alpha, \beta, \theta)$ given in (3.2.1). The advantage of this procedure is that it runs immediately using existing statistical packages such as R[17]. We consider the software R through the Quasi-Newton algorithm to compute the MLEs.

Figure 3 indicates that the likelihood equations have a unique solution. The Table 1 shows the ML estimates, standard error(SE) and 95 % confidence Intervals for parameters α, β and θ. The maximized value of loglikelihood is $\ell(\hat{\alpha}, \hat{\beta}, \hat{\theta}) = -409.6779$.

Parameter	MLE	Std. Error	95% Confidence Interval
alpha	1.3340	0.1580	(1.0243, 1.6437)
beta	2.4351	0.4721	(1.5098, 3.3604)

theta	10.2867	2.3471	(5.6865, 14.8869)

Table 1 MLE, standard error and 95% confidence interval

The Akaike information criterion(AIC) and Bayesian information criterion(BIC) can be used to determine which model is most appropriate for the given data. For the given data set we have computed AIC=825.3557 and BIC=833.9118, [12].

B. Model Validation

To check the validity of the model we compute the Kolmogorov-Smirnov (KS) distance between the empirical distribution function and the fitted distribution function when the parameters are obtained by method of maximum likelihood is 0.0336 and the corresponding *p-value* is 0.9987. We have plotted the empirical distribution function and the fitted distribution function in Figure 3.4. From the Figure 3.4, it is clear that the KwHC distribution provides a nice fit to the given real data.

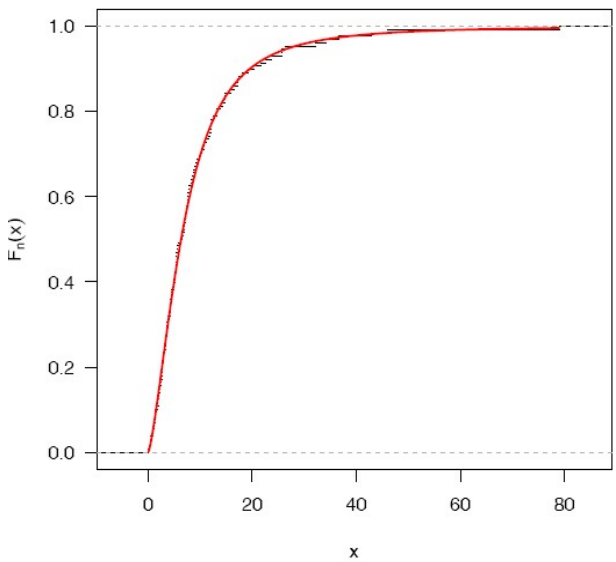


Figure 4 The graph of empirical and fitted distribution function.

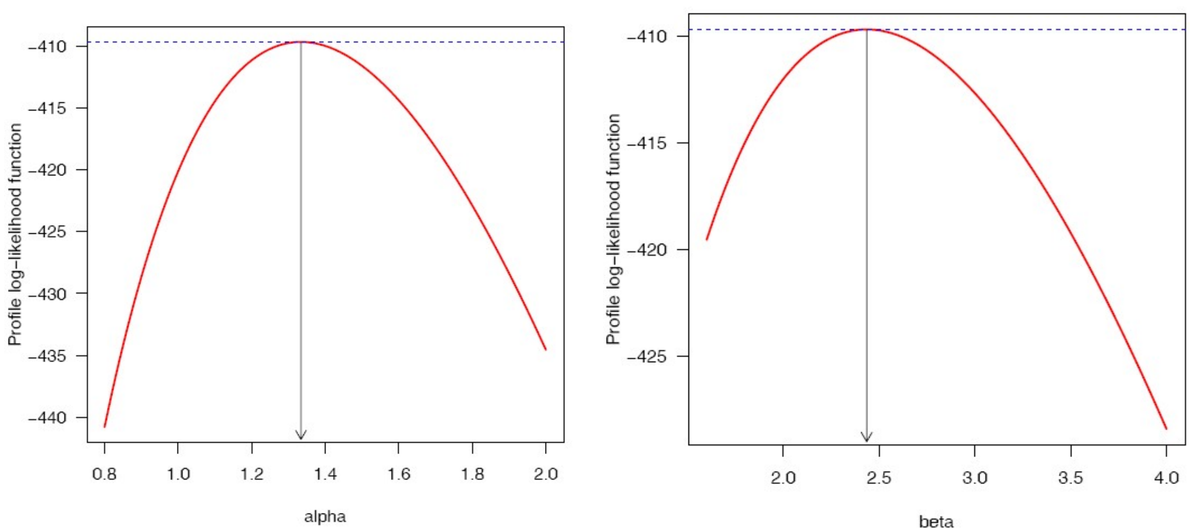
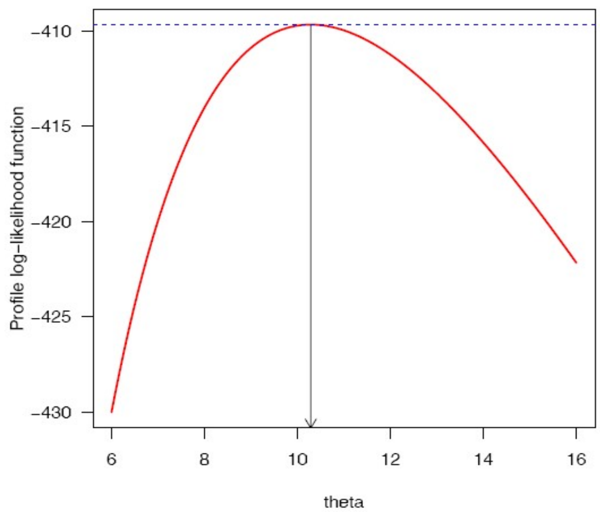


Figure 3 The profile negative log-likelihhod plots of alpha, beta and theta.



II. Bayesian Analysis

OpenBUGS script for the Bayesian analysis of Kumaraswamy-Half-Cauchy distribution

```
model
        {
                for( i in 1 : N )
                { x[i] ~ dkwh.cauchy(alpha, beta, theta)
                      reliability[i] <- R(x[i], x[i])
                      f[i] <- density(x[i], x[i])   }
                alpha ~ dgamma(0.001, 0.001)
                beta ~ dgamma(0.001, 0.001)
                theta~    dunif(0, 50)
          }
Data
              list(N=128, c(0.08, 0.20,...,46.12, 79.05))
Initial values
                list(alpha=1.0, beta=1.0, theta=5.0)
                list(alpha=5.0, Beta= 5.0, theta=20.0)
```

We assume the independent uniform prior for $\theta \sim U(a_3, b_3)$ and gamma priors for $\alpha \sim G(a_1, b_1)$ and $\beta \sim G(a_2, b_2)$ with hyperparameter values

$$(a_1 = 0.001,\ b_1 = 0.001), (a_2 = 0.001,\ b_2 = 0.001) \text{ and } (a_3 = 0, b_3 = 50.0).$$

We run the model to generate two Markov Chains at the length of 30,000 with different starting points of the parameters. We have chosen initial values $(\alpha = 1.0, \beta = 1.0, \theta = 5.0)$ for the first chain and $(\alpha = 5.0, \beta = 5.0, \theta = 20.0)$ for the second chain. The convergence is monitored using trace and ergodic mean plots, we find that the Markov Chain converge together after approximately 2000 observations. Therefore, burn-in of 5000 samples is more than enough to erase the effect of starting point(initial values). Finally, samples of size 5000 are formed from the posterior by picking up equally spaced every fifth outcome (to minimize the auto correlation among the generated deviates.), i.e. thin=5, starting from 5001.

Therefore, we have the posterior sample

$$\left(\alpha_1^{(j)}, \beta_1^{(j)}, \theta_1^{(j)}\right);\ j = 1,\ldots,5000 \text{ from chain 1} \quad \text{and} \quad \left(\alpha_2^{(j)}, \beta_2^{(j)}, \theta_2^{(j)}\right);\ j = 1,\ldots,5000 \text{ from chain 2.}$$

The chain 1 is considered for convergence diagnostics plots. The visual summary is based on posterior sample obtained from chain 1 whereas the numerical summary is presented for both the chains.

A. Convergence diagnostics

Before examining the parameter estimates or performing other inference, it is a good idea to look at plots of the sequential(dependent) realizations of the parameter estimates and plots thereof. The sequential plot of parameters is the plot that most often exhibits difficulties in the Markov chain. Figure 3.6 shows the sequential realizations of the parameters of the model.

History(Trace) plot:

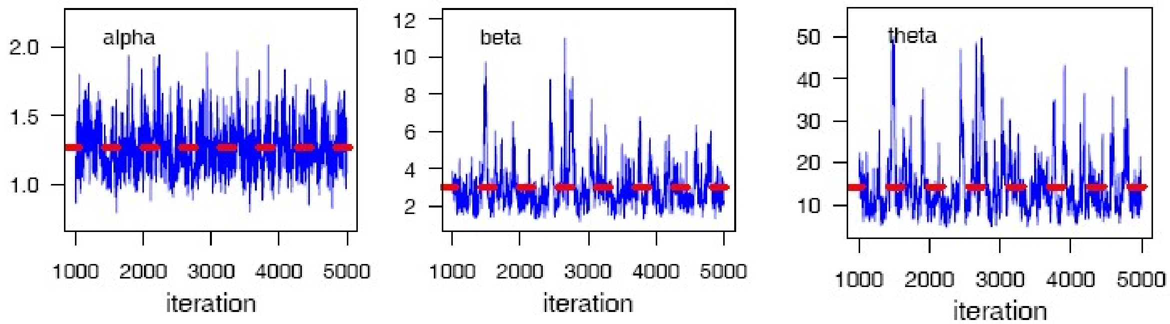


Figure 5 Sequential realization of the parameters α, β and θ .

It looks like nice oscillograms around a horizontal line without any trend. The Markov chain is most likely to be sampling from the stationary distribution and is mixing well.

Running Mean (Ergodic mean) Plot:

Generate a time series(iteration number) plot of the running mean for each parameter in the chain. The running mean is computed as the mean of all sampled values up to and including that at a given iteration. The convergence pattern based on ergodic averages is shown in Figure 3.7 indicating the convergence of the chain.

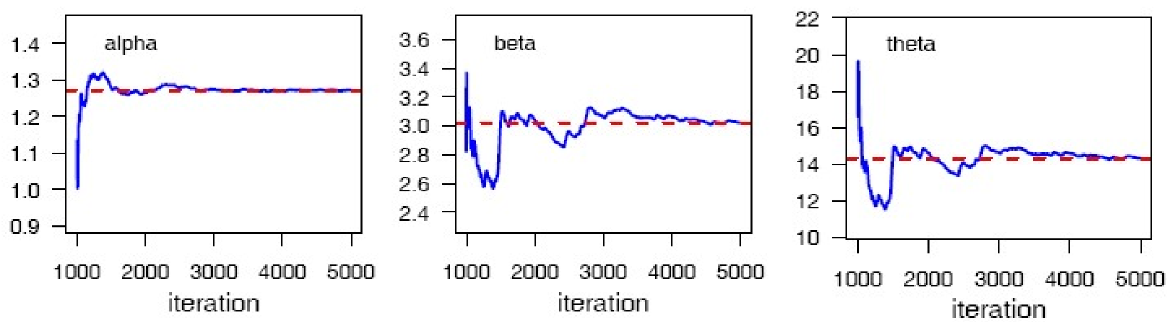


Figure 6 The ergodic mean plots for α, β and θ .

III. Posterior Analysis

A. Numerical Summary

The numerical summary is presented for $\left(\alpha_1^{(j)}, \beta_1^{(j)}, \theta_1^{(j)}\right)$; $j = 1,\ldots,5000$ from chain 1 We have considered various quantities of interest and their numerical values based on MCMC sample of posterior characteristics for KwHC distribution. The MCMC results of the posterior mean, mode, standard deviation(SD), first quartile, median, third quartile, skewness and kurtosis of parameters α, β and θ are displayed in Table 2

Characteristics	alpha	lambda	theta
Mean	1.2755	3.0564	14.4484
Standard Deviation	0.1769	1.2118	7.3137
First Quartile (Q_1)	1.1500	2.2628	9.6853
Median	1.2580	2.7490	12.4350

Third Quartile (Q_3)	1.3793	3.4733	16.9200
Mode	1.2091	2.4111	10.2710
Skewness	0.6303	2.0197	1.8875
Kurtosis	0.6646	6.4145	4.3969

Table 2 Numerical summaries for KwHC distribution

The advantage of using the MCMC method over the MLE method is that we can always obtain a reasonable interval estimate of the parameters by constructing the probability intervals based on the empirical posterior distribution. This is often unavailable in maximum likelihood estimation. The algorithm described by [18] is used to compute the HPD intervals.. The width of the HPD is another way of measuring uncertainty of beliefs. If the HPD is wide, then beliefs are uncertain. If the HPD is narrow, then beliefs are reasonable.

Parameter	Symmetric Credible Interval	HPD Credible Interval
alpha	(0.9801, 1.670)	(0.9436, 1.618)
beta	(1.6550, 6.3381)	(1.417, 5.471)
theta	(6.227, 35.700)	(4.809, 29.48)

Table 3 95% symmetric and HPD credible intervals

B. Visual summary

The visual graphs include the boxplot, density strip plot, histogram, marginal posterior density estimate and rug plots for the parameters. We have also superimposed the 95% HPD intervals.

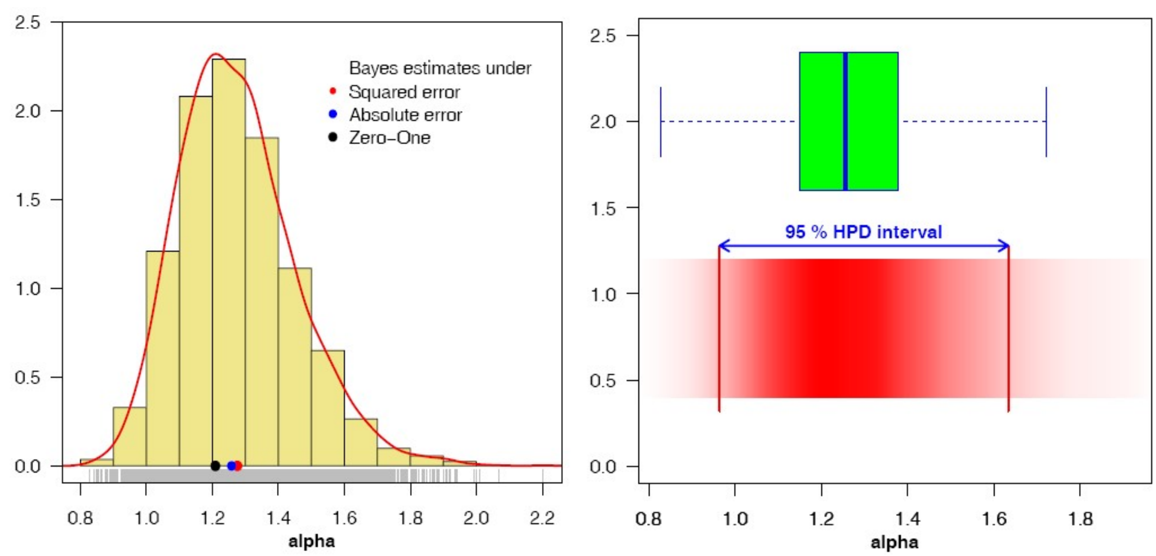


Figure 7 Left panel : Histogram and marginal posterior density,
Right panel : boxplot, density strip and 95% HPD interval of α .

and zero-one loss functions loss, respectively. Figure 7-(right panel) shows the boxplot and density strip plot. The 95% HPD interval is also superimposed.

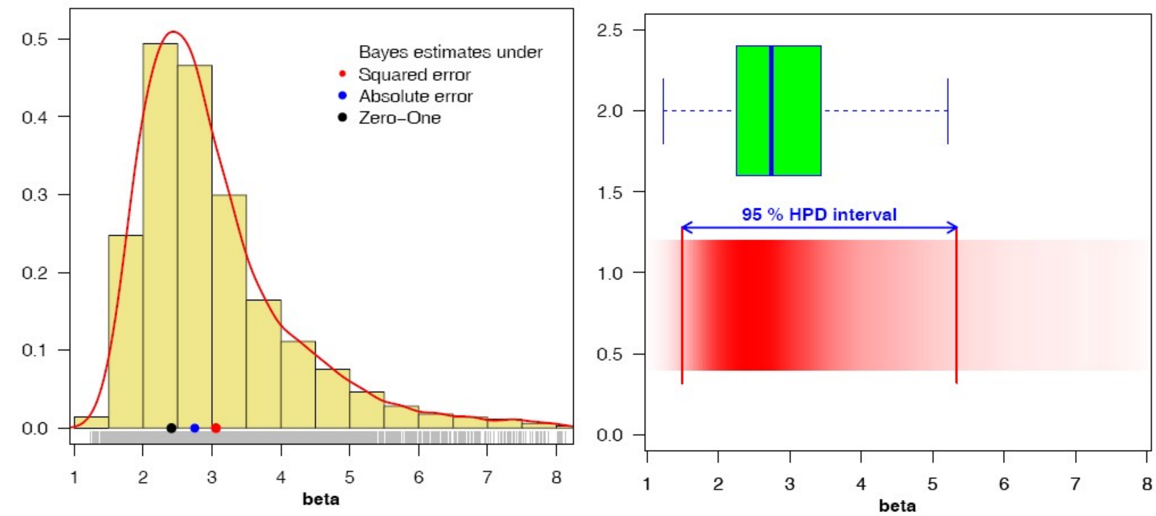


Figure 8 Left panel : Histogram and marginal posterior density,
Right panel : boxplot, density strip and 95% HPD interval of β .

The density strip shows a univariate distribution as a shaded rectangular, whose darkness at a point is proportional to the probability density. We have plotted the similar graphs for β and θ displayed in Figure 8 and 9. It can be seen that α , β and θ show positive skewness.

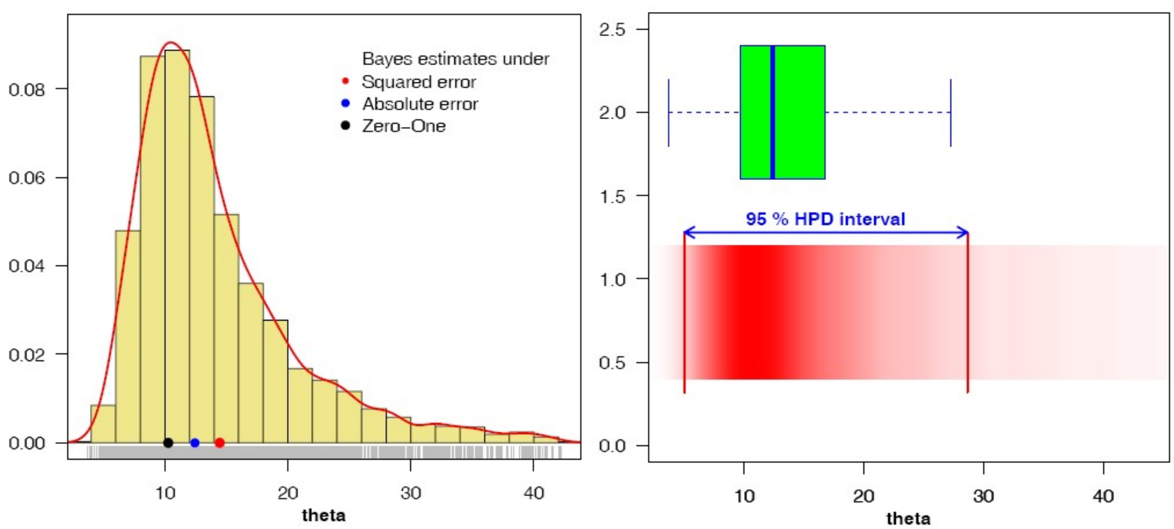


Figure 9 Left panel : Histogram and marginal posterior density,
Right panel : boxplot, density strip and 95% HPD interval of θ.

IV. Comparison with MLE

We have used graphical method for the comparison of Bayes estimates with ML estimates. In Figure 10, the density functions $f(x;\hat{\alpha},\hat{\beta},\hat{\theta})$ using MLEs and Bayesian estimates (the posterior means), computed via MCMC samples, are plotted.It is evident from the Figure 10 that the MLEs and the Bayes estimates are quite close and fit the data very well.

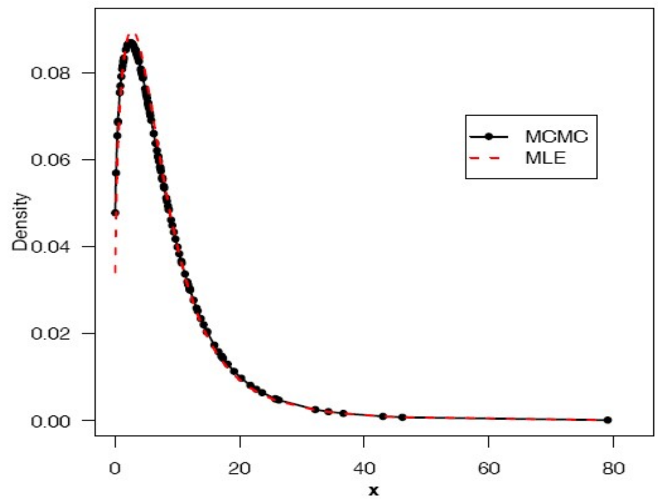


Figure 10 The density functions using ML and Bayesian estimates

A further support for this finding can be obtained by inspecting the Figure 11. where we have plotted 2.5^{th}, 50^{th} and 97.5^{th} quantiles of the estimated density, it can be considered as evaluation of model fit, based on posterior sample, $\left(\alpha_1^{(j)}, \beta_1^{(j)}, \theta_1^{(j)}\right)$; $j = 1, \ldots, 5000$.

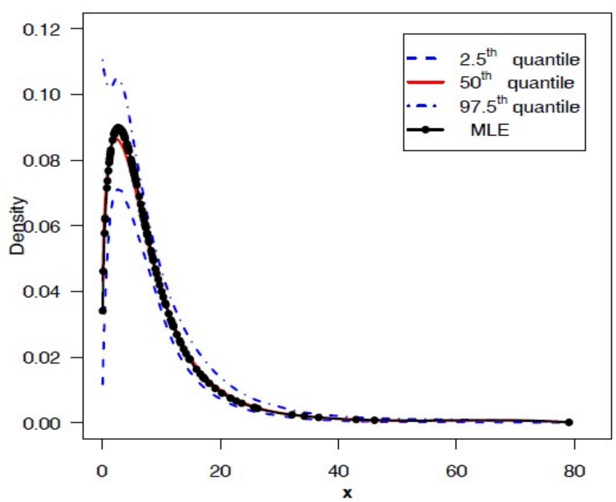


Figure 11 Density estimates

We have computed the density function at each observed data point for 5000 posterior samples, using logical function *density( )* in OpenBUGS $f\left(x_i; \alpha_1^{(j)}, \beta_1^{(j)}, \theta_1^{(j)}\right)$; $j = 1, \ldots, 5000$; $i = 1, \ldots, 128$.

The density corresponding to MLE has been plotted using the "plug-in" estimates of the parameters. It shows that we have a fairly good model for the given data set.

IV. Estimation of Hazard and Reliability functions:

In this section, our main aim is to demonstrate the effectiveness of proposed methodology. For this, we have estimated the reliability function using posterior samples. Since we have an effective MCMC technique, we can estimate any function of the parameters. We have used the Kaplan-Meier estimate of the reliability function to make the comparison more meaningful. The Figure 12 (left panel), exhibits the estimated reliability function (dashed blue line: 2.5^{th} and 97.5^{th} quantiles; solid red line: 50^{th} quantile) using Bayes estimate based on MCMC output and the empirical reliability function (black solid line). The Figure 12 (left panel) shows that reliability estimate based on

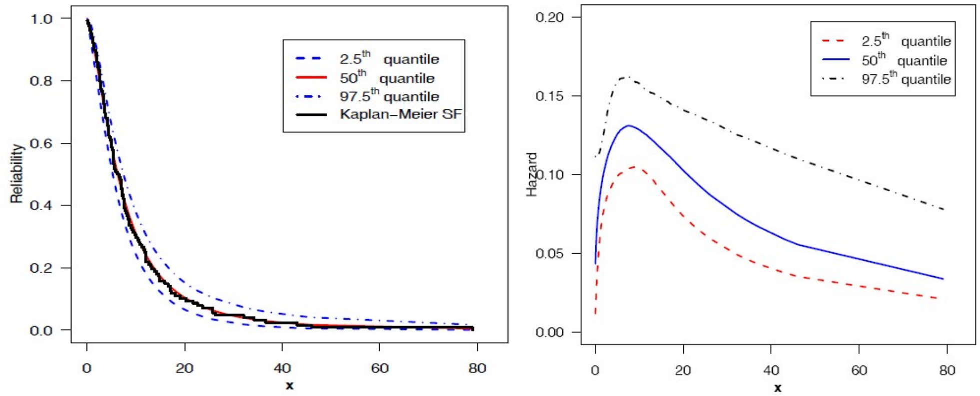


Figure 12 Reliability function(left panel) and hazard function(right panel) estimate using MCMC

MCMC is very close to the empirical reliability estimates. The estimated hazard function (dashed blue line: 2.5^{th} and 97.5^{th} quantiles; solid red line: 50^{th} quantile) using Bayes estimate based on MCMC output has been displayed in the Figure 12 (right panel).

V. Estimation of Hazard and Reliability at $\mathbf{X}_{(30)}$: **t** = 3.02

Indeed, the MCMC samples may be used to completely summarize the posterior uncertainty about the parameters α, β and θ through a kernel estimate of the posterior distribution. This is also true of any function of the parameters e.g. reliability and hazard functions. Suppose we wish to give point and interval estimates for reliability and hazard functions at the mission time *t*=3.02 (at the 30^{th} observed data point).

We have computed the hazard and reliability functions at mission time t=3.02(at the 30^{th} observed data point) for 5000 posterior samples, using logical functionhrf( ) and reliability( ), [12] in OpenBUGS. It can be computed directly using hazard and reliability functions given in (4) and (3) respectively.

$$h\left(x = 3.02\,; \alpha_1^{(j)}, \beta_1^{(j)}, \theta_1^{(j)}\right);\ j = 1, \ldots, 5000 \text{ and } R\left(x = 3.02\,; \alpha_1^{(j)}, \beta_1^{(j)}, \theta_1^{(j)}\right);\ j = 1, \ldots, 5000$$

Alternatively, we can use R functions hkw.halfCauchy() and skw.halfCauchy() [12].

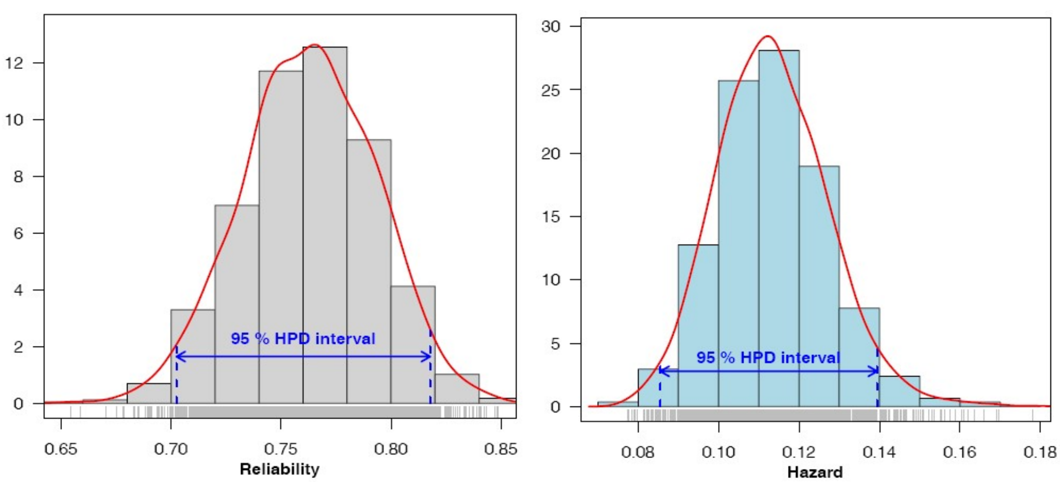


Figure 13 Visual summary of reliability(left panel) and hazard(right panel) at t=3.02

The marginal posterior density estimates of the reliability (left panel) and hazard functions(right panel) and their histograms based on samples of size 5000 are shown in Figure 13 using the Gaussian kernel. The 95% HPD intervals are superimposed. It is evident from the estimates that the marginal distribution of reliability is negatively skewed whereas hazard is positively skewed.

The MCMC results of the posterior mean, mode, standard deviation(SD), first quartile, median, third quartile, skewness, kurtosis, 95% symmetric and HPD credible intervals of reliability and hazard functions are displayed in Table 4. The ML estimates of reliability and hazard function at t=3.02 are computed using invariance property of the MLE. ML estimates $\hat{h}(t=3.02)=0.1167$ and $\hat{R}(t=3.02)=0.7677$.

Characteristics	Reliability	Hazard
Mean	0.7625	0.1134
Standard Deviation	0.0301	0.0138
First Quartile (Q_1)	0.7425	0.1038
Median	0.7630	0.1129
Third Quartile (Q_3)	0.7841	0.1223
Mode	0.7661	0.1125
Skewness	-0.0852	0.4029
Kurtosis	-0.1385	0.5601
95% Credible Interval	(0.7029, 0.8189)	(0.0881, 0.1422)
95% HPD Credible Interval	(0.7024, 0.8178)	(0.0855, 0.1394)

Table 4 Posterior summary for Reliability and Hazard functions at t=3.02

VI. Posterior Predictive Checks forModel compatibility

A natural way to assess the fit of a Bayesian model is to look at how well the predictions from the model agree with the observed data [19]. We do this by comparing the posterior predictive simulations with the data.

There are several approaches available for the study of model compatibility in Bayesian framework. Predictive simulation is an easiest and flexible one. The basic idea of studying the model compatibility through predictive simulation is to compare the observed data or some function of it with the data that would have been anticipated from the assumed model called the predictive data. If the two data sets compare favorably, the assumed model can be considered to be an appropriate choice for the data in hand, [20].

.

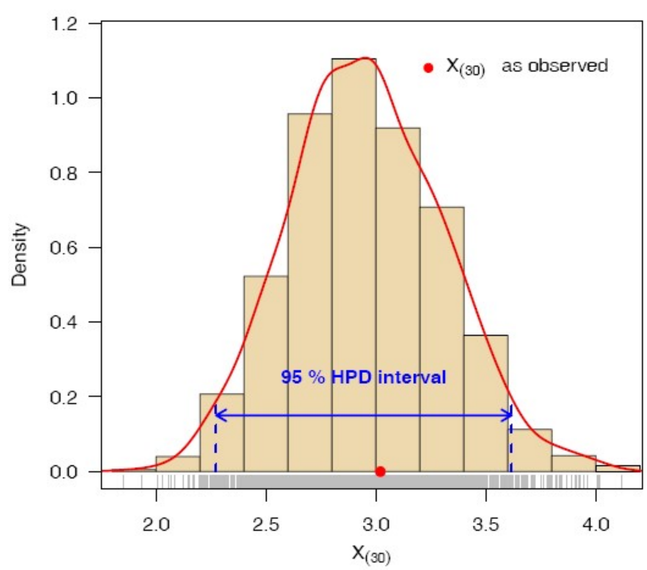


Figure 14 Posterior predictive distribution of $X_{(30)}$ with corresponding observed value.

Modern Bayesian computational tools however provide straightforward solutions as one can easily simulate predictive samples if MCMC outputs are available from the posterior corresponding to the assumed model. Most of the standard numerical and graphical methods based on predictive distribution can then be easily implemented to study the compatibility of the model.

One of the best ways to assess model adequacy based on posterior predictive distributions is graphically. To obtain further clarity on our conclusion for the study of model compatibility, we have considered plotting of density estimates of. $\left(X_{(1)}, X_{(2)}, X_{(30)}, X_{(127)} \text{ and } X_{(128)}\right)$ replicated future observations from the model with superimposed corresponding observed data. For this purpose, 2000 samples have been drawn from the posterior using MCMC procedure and then obtained predictive samples from the model under consideration using each simulated posterior sample. The size of predictive samples is same as that of observed data.

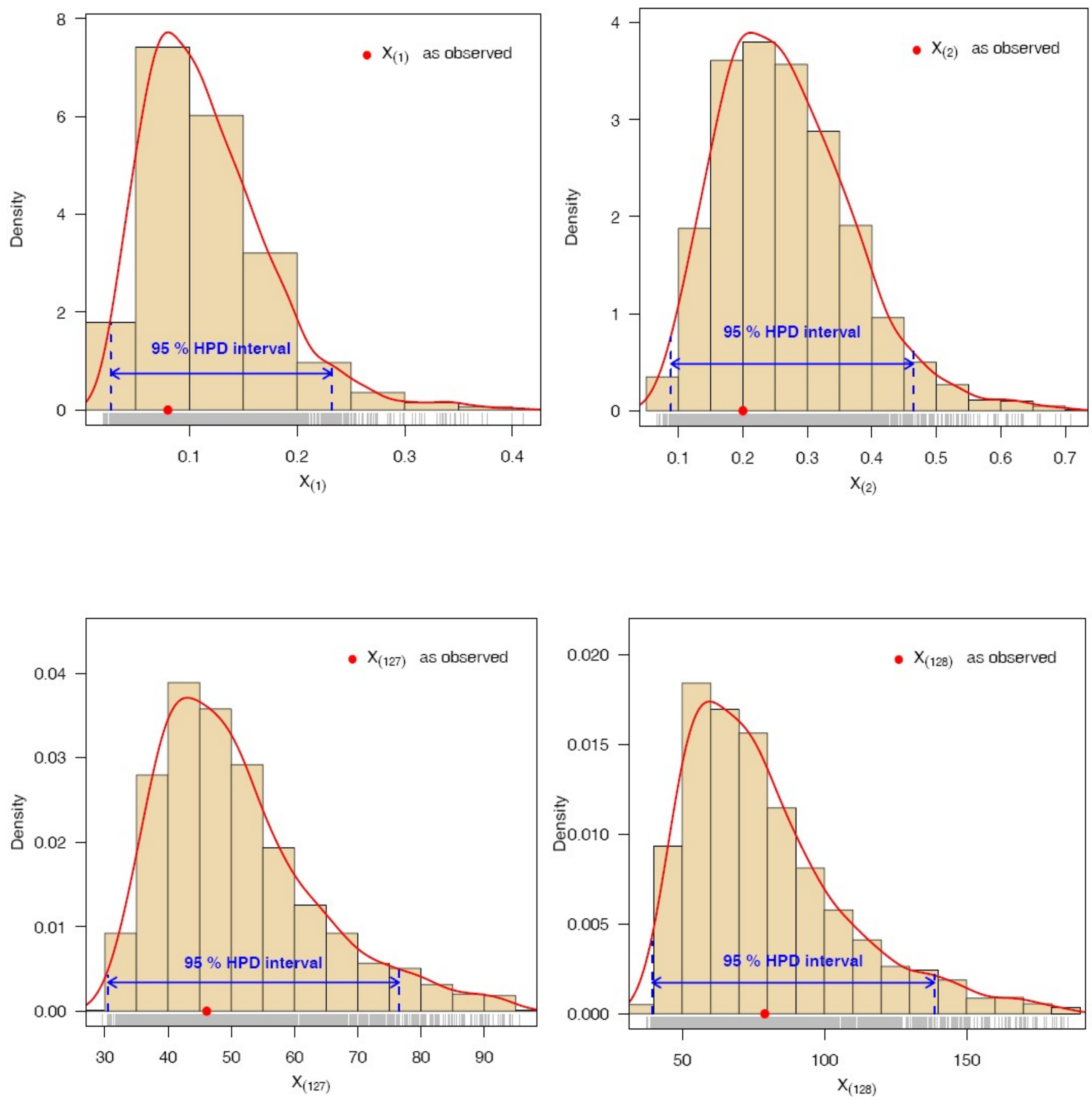


Figure 15 Posterior predictive distribution, 95% HPD interval of the $\left(X_{(1)}\right)$, $\left(X_{(2)}\right)$, $\left(X_{(127)}\right)$ and $\left(X_{(128)}\right)$, corresponding observed values are marked as(●) on the axis.

The posterior predictive distributions based on replicated future data sets are shown in Figures 14 and 15. Here Figure 15 represents the estimates corresponding to smallest, second smallest, second largest and largest predictive observations, whereas the same for 30^{th} smallest observations is shown in Figure 14. The corresponding observed values are also shown.

The MCMC results of the posterior mean, median, mode of smallest and largest $\left(X_{(1)}, X_{(2)}, X_{(127)} \text{ and } X_{(128)}\right)$ and $X_{(30)}$ are displayed in Table 3.5.

	Observed	Mode	Mean	Median	95% HPD
$X_{(1)}$	0.08	0.08	0.12	0.11	(0.027, 0.233)
$X_{(2)}$	0.20	0.21	0.27	0.25	(0.089, 0.465)
$X_{(30)}$	3.02	2.97	2.96	2.95	(2.273, 3.617)
$X_{(127)}$	46.12	42.71	51.72	48.67	(31.62, 80.54)
$X_{(128)}$	79.05	58.92	82.13	73.61	(39.08, 146.7)

Table 5 Posterior characteristics of KwCH distribution

As the Figures 14 and 15 shows, the posterior predictive distributions are centered over the observed values, which indicate good fit. In general, the distribution of replicated data appears to match that of the observed data fairly well. Overall, the results of the posterior predictive simulation indicate that model fits these data particularly well.

VI. Conclusion

We have proposed KwHC distribution and discussed some of its properties using R software. We have obtained the MLE of the parameters and their asymptotic probability intervals. Then, we have discussed the Markov chain Monte Carlo (MCMC) method to compute the Bayesian estimates of the parameters, hazard and reliability functions of KwHC distribution based on a complete sample. We have obtained the probability intervals for parameters, hazard and reliability functions. We have presented the model compatibility via the posterior predictive check method. We have applied the developed techniques on a real data set. Thus, the tools developed can be applied for full Bayesian analysis of KwHC distribution.

To obtain further clarity on our conclusion for the study of model compatibility, we have considered plotting of density estimates of. $\left(X_{(1)}, X_{(2)}, X_{(30)}, X_{(127)} \text{ and } X_{(128)}\right)$ replicated future observations from the model with superimposed corresponding observed data. For this purpose, 2000 samples have been drawn from the posterior using MCMC procedure and then obtained predictive samples from the model under consideration using each simulated posterior sample. The size of predictive samples is same as that of observed data.

REFERENCES

[1] Polson, N. G., & Scott, J. G. (2012). On the half-Cauchy prior for a global scale parameter. Bayesian Analysis, 7(4), 887-902.

[2] Kumaraswamy (1980). A generalized probability density function for double-bounded random processes. J Hydrol. 1980;46:79–88. doi: 10.1016/0022-1694(80)90036-0

[3] Jones MC. (2009). Kumaraswamy's distribution: a beta-type distribution with some tractability advantages. Stat Methodol. 2009;6:70–81. doi: 10.1016/j.stamet.2008.04.001

[4] Cordeiro, G. M., & de Castro, M. (2011). A new family of generalized distributions. Journal of statistical computation and simulation, 81(7), 883-898.

[5] Cordeiro, G.M., & Lemonte, A.J. (2011). The beta-half- Cauchy distribution. Journal of Probability and Statistics, Vol. 2011, 1-18, doi: 10.1155/2011/904705.

[6] Ghosh, I.; Nadarajah, S.(2017). On the bayesian inference of kumaraswamy distributions based on censored samples. Commun. Stat. 2017, 46, 8760–8777

[7] Jacob, E., & Jayakumar, K. (2012). On half-Cauchy distribution and process. International Journal of Statistika and Mathematika, 3(2), 77-81.
[8] Alzaatreh, A., Lee, C., Famoye, F., & Ghosh, I. (2016). The generalized Cauchy family of distributions with applications, Journal of Statistical Distributions and Applications, 3-12
[9] Alzaatreh, A., Mansoor, M., Tahir, M. H., Zubair, M., & Ghazali, A. S. (2016). The Gamma-Half-Cauchy distribution : Properties and Applications, Hacettepe Journal of Mathematics and Statistics, 45 (4), 1143-1159.
[10] Srivastava, A.K. (2019). "Bayesian Approach in Estimation of Shape and Scale Parameter of Log-Weibull model". International Journal of Innovative Science, Engineering & Technology, Vol. 6 Issue 12, December 2019, 185-196.
[11] Srivastava, A.K. and Kumar, V.(2015). " A study on several issues of Reliability Modelling for a Real Data Set using different Software Reliability Models", International Journal of Emerging Technology and Advanced Engineering, Vol.05(12), 49-57.
[12] Kumar, V. and Ligges, U. (2011). reliaR : A package for some probability distributions. http://cran.r-project.org/web/packages/reliaR/index.html.
[13] 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, Retrieved from http://openbugs.info/w/Downloads/
[14] Lunn, D.J., Jackson, C., Best, N., Andrew, A., & Spiegelhalter, D. (2013). The BUGS Book :A Practical Introduction to Bayesian Analysis. London, UK: Chapman & Hall/CRC.
[15] Thomas, A.(2010). OpenBUGS Developer Manual, version 3.1.2, Retrieved from http://www.openbugs.info/
[16] Lee, E.T. and Wang, T.W.(2003). Statistical Methods for Survival Data Analysis. 3rd ed., John Wiley & Sons.
[17] R Development Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Retrieved from URL http://www.R-project.org.
[18] 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.
[19] Gelman, A. (2003). A Bayesian Formulation of Exploratory Data Analysis and Goodnessof-fit Testing. International Statistical Review, 71(2), 369-382.
[20] Gupta, A., Mukherjee, B., & Upadhyay, S.K. (2008). A Bayes study using Markov Chain Monte Carlo simulation. Reliability Engineering & System Safety, 93, 1434-1443.