

A Bayesian Analysis of Perks Distribution via Markov Chain Monte Carlo Simulation

Arun Kumar Chaudhary¹ and Vijay Kumar²

¹Nepal Commerce Campus, Tribhuvan University, Kathmandu

²Department of Mathematics and Statistics

DDU Gorakhpur University, Gorakhpur, India

e-mail: akchaudhary1@yahoo.com

Abstract

In this paper the Markov chain Monte Carlo (MCMC) method is used to estimate the parameters of Perks distribution based on a complete sample. The procedures are developed to perform full Bayesian analysis of the Perks distributions using MCMC simulation method in OpenBUGS. We obtained the Bayes estimates of the parameters, hazard and reliability functions, and their probability intervals are also presented. We also discussed the issue of model compatibility for the given data set. A real data set is considered for illustration under gamma sets of priors.

Key words: Bayesian estimation, Markov chain Monte Carlo, maximum likelihood estimation, model validation, Perks distribution, predictive simulation, openBUGS

Introduction

Perks (1932) proposed the four-parameter extension of the Gompertz–Makeham distribution that has hazard rate of the form:

$$h(x) = \frac{A + B e^{\lambda x}}{K e^{-\lambda x} + 1 + D e^{\lambda x}} ; x > 0.$$

The choice $K = D = 0$ yields the Gompertz–Makeham hazard rate. It appears that Perks intended the parameters to be nonnegative. Marshall and Olkin (2007) have shown that we cannot take $D=0$. However, Gompertz–Makeham distribution can be obtained by setting $K=0$ and taking limit as $D \rightarrow 0$.

Recently, the modified version of Perks distribution has been introduced by Richards (2008, 2012) by considering the hazard function of Perks distribution as

$$h(x) = \frac{\alpha \beta e^{\beta x}}{(1 + \alpha e^{\beta x})} ; \alpha > 0, \beta > 0, x > 0.$$

Applications of the Perks distribution to actuarial science include: models for pensioner mortality data (Richards 2008), parametric mortality projection

models (Haberman & Renshaw 2011). The moments for this distribution do not appear to be available in closed form, (Nadarajah & Bakar 2012). Another version of Perks distribution has been introduced by (Yee 2012).

It is to be noted that most of the cited literature are confined to classical developments and any systematic development on Bayesian results are rarely seen for the Perks distribution. The importance of the Bayesian method is well known both in the context of reliability studies and otherwise. Among several advantages, the most important is the fact that the Bayesian methods are equally well applicable for small sample sizes and censored data problems; the two common features in reliability data analyses.

In this paper, the Perks distribution has been introduced and maximum likelihood estimation (MLE) is discussed in Section 2. Section 3 deals Bayesian model formulation. A real data set is presented in Section 4 for analysis. The exploratory data analysis and ML estimation and corresponding model validation are performed in section 5. The full Bayesian analysis

under independent gamma of priors of the data set using Markov chain Monte Carlo (MCMC) simulation method in OpenBUGS, an established software, is presented in section 6. We have obtained the Bayes estimates of the parameters, hazard and reliability functions, and their probability intervals are also presented. We have also discussed the issue of model compatibility for the given data set.

Perks distribution

Model analysis

Cumulative distribution function

The distribution function of Perks distribution with two parameters is given by

$$F(x; \alpha, \beta) = 1 - \left(\frac{1 + \alpha}{1 + \alpha e^{\beta x}} \right) ; \alpha > 0, \beta > 0, x \geq 0. \quad (2.1)$$

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

$$f(x; \alpha, \beta) = \alpha \beta e^{\beta x} \frac{(1 + \alpha)}{(1 + \alpha e^{\beta x})^2} ; \alpha > 0, \beta > 0, x \geq 0. \quad (2.2)$$

The reliability/survival function is

$$R(x; \alpha, \beta) = \left(\frac{1 + \alpha}{1 + \alpha e^{\beta x}} \right) ; \alpha > 0, \beta > 0, x \geq 0. \quad (2.3)$$

Some of the typical Perks density functions for different values of β and for $\alpha = 1$ are depicted in Fig. 1. It is clear that the density function of the $PS(\alpha, \beta)$ distribution can take different shapes.

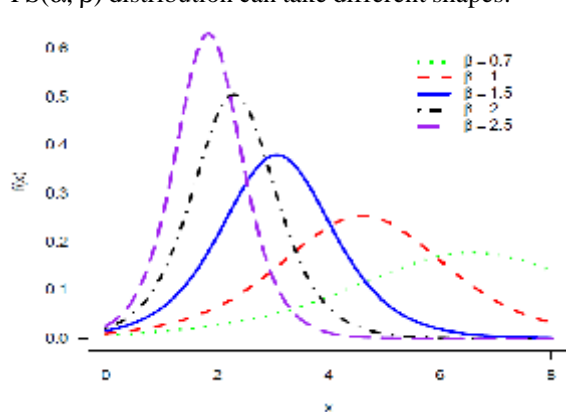


Fig. 1. The probability density function of Perks I distribution for $\alpha=1$ and different value of β

$$\text{Mode} = \frac{1}{\beta} \log \left(\frac{1}{\alpha} \right) ; \alpha \geq 1$$

The hazard rate function is

$$h(x; \alpha, \beta) = \frac{\alpha \beta e^{\beta x}}{(1 + \alpha e^{\beta x})} ; \alpha > 0, \beta > 0, x \geq 0. \quad (2.4)$$

Fig. 2 exhibits the different hazard rate functions of $PS(\alpha, \beta)$ distribution for $\alpha = 1$ and different values of β .

The quantile function is given by

$$x_p = \frac{1}{\beta} \log \left\{ \frac{1}{\alpha} \left(\frac{(1 + \alpha)}{(1 - p)} - 1 \right) \right\} ; 0 < p < 1. \quad (2.5)$$

The cdf has close nice form, using the inverse transform method, the random deviate can be generated from $PS(\alpha, \beta)$ by

$$x = \frac{1}{\beta} \log \left\{ \frac{1}{\alpha} \left(\frac{(1 + \alpha)}{(1 - u)} - 1 \right) \right\} ; 0 < u < 1. \quad (2.6)$$

where u has the $U(0, 1)$ distribution.

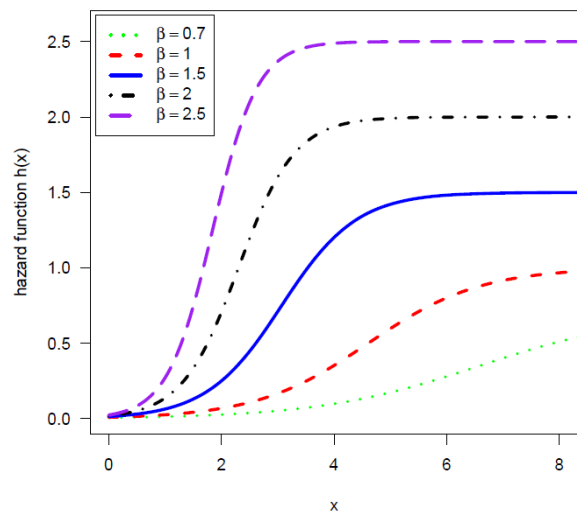


Fig. 2. The hazard function of Perks I distribution for $\alpha=1$ and different values of β

Maximum likelihood estimation (MLE)

For completeness purposes, in this section, we briefly discuss the maximum likelihood estimators (MLE's) of the two-parameter $PS1(\alpha, \beta)$ distribution.

Let $\underline{x}=(x_1, \dots, x_n)$ be a random sample of size n from $PS1(\alpha, \beta)$, then the log-likelihood function $l(\alpha, \beta)$ can be written as;

$$\ell(\alpha, \beta) = n \log(\alpha) + n \log(\beta) + \beta \sum_{i=1}^n x_i + n \log(1 + \alpha) - 2 \sum_{i=1}^n \log(1 + \alpha \exp(\beta x_i)) \quad (2.7)$$

Therefore, to obtain the MLE's of α and β , we can maximize (2.7) directly with respect to α and β or we can solve the following two non-linear equations using Newton-Raphson method.

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} + \frac{n}{1 + \alpha} - 2 \sum_{i=1}^n \frac{\exp(\beta x_i)}{1 + \alpha \exp(\beta x_i)} = 0$$

$$\frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^n x_i - 2 \sum_{i=1}^n \frac{\alpha x_i \exp(\beta x_i)}{1 + \alpha \exp(\beta x_i)} = 0$$

From the asymptotic normality of MLEs, an approximate confidence intervals for α and β can be constructed as

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

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

Bayesian model formulation

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

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

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

$$L(\alpha, \beta | \underline{x}) = (\alpha \beta)^n (1 + \alpha)^n \exp\left(\beta \sum_{i=1}^n x_i\right) \left(\prod_{i=1}^n (1 + \alpha e^{\beta x_i})^{-2}\right)$$

Denote the prior distribution of α and β as $p(\alpha, \beta)$. The joint posterior is

$$p(\alpha, \beta | \underline{x}) = (\alpha \beta)^n (1 + \alpha)^n \exp\left(\beta \sum_{i=1}^n x_i\right) \left(\prod_{i=1}^n (1 + \alpha e^{\beta x_i})^{-2}\right) p(\alpha, \beta)$$

Prior distributions

We assume the independent gamma priors for $\alpha \sim G(a_1, b_1)$ and $\beta \sim G(a_2, b_2)$ as

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

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

Posterior distribution

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

$$p(\alpha, \beta | \underline{x}) = (\alpha \beta)^n (1 + \alpha)^n \exp\left(\beta \sum_{i=1}^n x_i\right) \left(\prod_{i=1}^n (1 + \alpha e^{\beta x_i})^{-2}\right) \left(\frac{b_1^{a_1}}{\Gamma(a_1)} e^{-b_1 \alpha} \alpha^{a_1-1}\right) \left(\frac{b_2^{a_2}}{\Gamma(a_2)} e^{-b_2 \beta} \beta^{a_2-1}\right)$$

It can be written, upto proportionality, as

$$p(\alpha, \beta | \underline{x}) \propto \alpha^{n+a_1-1} \beta^{n+a_2-1} (1 + \alpha)^n \left(\prod_{i=1}^n (1 + \alpha e^{\beta x_i})^{-2}\right) \exp\left(-b_1 \alpha - b_2 \beta + \beta \sum_{i=1}^n x_i\right)$$

It can be seen that posterior is quite messy, therefore, we propose to consider MCMC methods to simulate samples from the posterior so that sample-based inferences can be easily drawn.

The Gibbs sampler is as an important Markov Chain Monte Carlo technique, which provides a way for extracting samples from the posteriors, often specified up to proportionality, as a product of likelihood function and the prior. MCMC is essentially Monte Carlo integration using Markov chains. The scheme proceeds iteratively by generating variate values in a cyclic manner from the various full conditionals, which are also specified up to proportionality from the joint posterior.

This sampling scheme was first introduced by Geman and Geman(1984), but the applicability to statistical modelling for Bayesian computation was demonstrated by Gelfand and Smith (1990).

The algorithm starts by assuming some arbitrarily chosen initial values for the concerned variates and then generating the variate values from the various full conditionals in a cyclic order. That is, every time a variate value is generated from a full conditional, it is influenced by the most recent values of all other conditioning variables and, after each cycle of iteration, it is updated by sampling a new value from its full conditional. The entire generating scheme is repeated unless the generating chain achieves a systematic pattern of convergence. It can be shown that after a large number of iterations the generated variates can be regarded as the random samples from the

corresponding posteriors. Gamerman and Lopes (2006), Albert (2009), Hamada *et al.* (2008), Hoff (2009), Ntzoufras (2009) and Robert and Casella (2010) provide the details of the procedure and the related convergence diagnostic issues.

For Gibbs sampler implementation, the full conditionals of α and β upto proportionality can be specified as

(i) full conditional distribution of the parameter α for given β and \underline{x}

$$p(\alpha | \beta, \underline{x}) \propto \alpha^{n+a_1-1} e^{-b_1\alpha} (1+\alpha)^n \left(\prod_{i=1}^n (1+\alpha e^{\beta x_i})^{-2} \right)$$

(ii) full conditional distribution of the parameter β for given α and \underline{x}

$$p(\beta | \alpha, \underline{x}) \propto \beta^{n+a_2-1} \exp\left[-\beta\left(b_2 - \sum_{i=1}^n x_i\right)\right] \left(\prod_{i=1}^n (1+\alpha e^{\beta x_i})^{-2} \right)$$

We shall use OpenBUGS software to obtain posterior samples. As the Perks distribution is not available in OpenBUGS, it requires incorporation of a module in *ReliaBUGS* [Kumar *et al.* (2010)] and [Lunn (2010)], subsystem of *OpenBUGS* for Perks distribution.

A module *dPerks1_T(alpha, beta)* is written in Component Pascal for Perks, the corresponding computer program can be obtained from authors, to perform full Bayesian analysis in OpenBUGS using the method described in Thomas *et al.* (2006), Thomas (2010), Kumar *et al.* (2010) and Lunn *et al.* (2000, 2013). 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 *dPerks1_T(alpha, beta)* Kumar (2010).

Data set

The following real data set is considered for illustration of the proposed methodology. The data below are from an accelerated life test of 59 conductors, failure times are in hours, and there are no censored observations Lawless (2003).

2.997, 4.137, 4.288, 4.531, 4.700, 4.706, 5.009, 5.381, 5.434, 5.459, 5.589, 5.640, 5.807, 5.923, 6.033, 6.071, 6.087, 6.129, 6.352, 6.369, 6.476, 6.492, 6.515, 6.522, 6.538, 6.545, 6.573, 6.725, 6.869, 6.923, 6.948, 6.956, 6.958, 7.024, 7.224, 7.365, 7.398, 7.459, 7.489, 7.495, 7.496, 7.543, 7.683, 7.937, 7.945, 7.974, 8.120, 8.336, 8.532, 8.591, 8.687, 8.799, 9.218, 9.254, 9.289, 9.663, 10.092, 10.491, 11.038

Classical analysis

Exploratory data analysis (EDA)

The goal of data analysis is to gain information from the data. The modern statistical data analysis tools include the exploratory data analysis (Tukey 1977).

Exploratory data analysis is a set of methods to display and summarize the data :

- displaying the data in a graph that shows overall patterns and unusual observations (bar chart, histogram, density curve)
- computing descriptive statistics that summarize specific aspects of the data (center and spread).

Table 1. Summary statistics

Min.	Q ₁	Median	Mean	Q ₃	Max.	Skewness	Kurtosis
2.997	6.052	6.923	6.980	7.941	11.040	0.1882	-0.0164

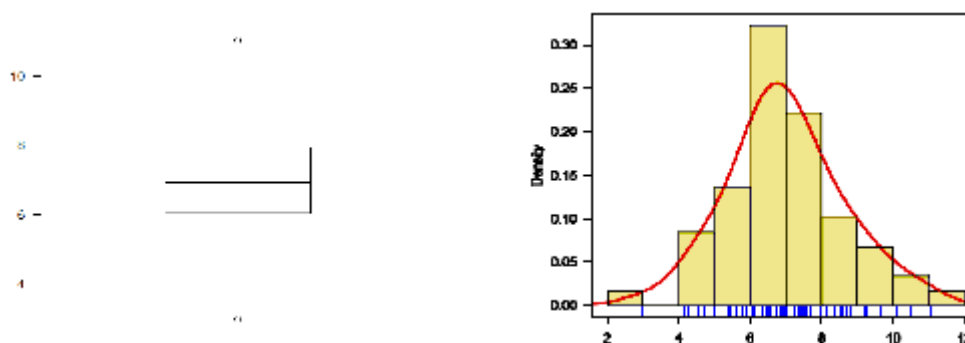


Fig. 3. The boxplot (left panel); Histogram and Kernel density estimate (right panel) along with the data points

The total-time-on-test(TTT) plot is a graphical procedure to get some idea about the shape of the hazard function. We have used the empirical version of the scaled TTT plot, Aarset (1987). We have plotted the empirical version of the scaled TTT transform of the data set in Fig. 4. Since the empirical version of the scaled TTT transform is concave, it indicates that the hazard function is increasing.

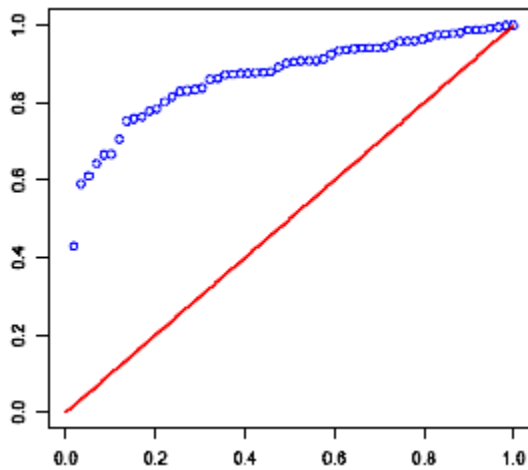


Fig. 4. The empirical scaled TTT transform of the data set

Computation of MLE

The maximum likelihood estimates (MLEs) are obtained by direct maximization of the log-likelihood function $l(\alpha, \beta)$ given in (2.7). The advantage of this procedure is that it runs immediately using existing statistical packages such R [R Development Core Team, (2013)], and [Rizzo(2008)]. We consider the software R through the Quasi-Newton algorithm [Lange (1999)] to compute

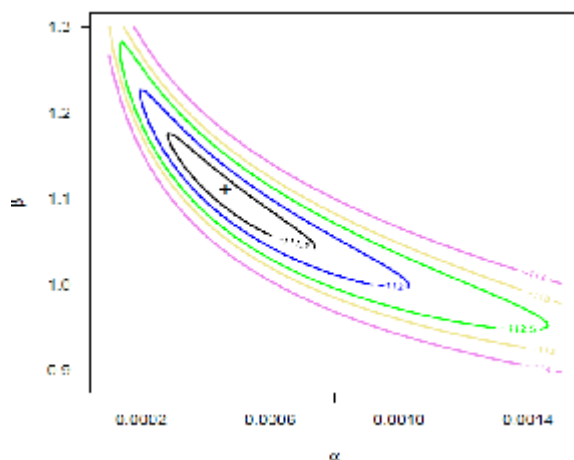


Fig. 5. Contour plot

the MLEs. The MLEs (and the corresponding standard errors in parentheses) and the maximized value of loglikelihood of the Perks distribution parameters are given by

$$\hat{\alpha} = 0.0004597(0.0004142),$$

$$\hat{\beta} = 1.1077(0.1256288)$$

and $l(\hat{\alpha}, \hat{\beta}) = -111.544$

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 0.0576 and the corresponding p-value is 0.9832. We have plotted the empirical distribution function and the fitted distribution function in Fig. 6, (Kumar & Ligges 2011). From the Figure 6, it is clear that the fitted Perks distribution provides good fit to the given data.

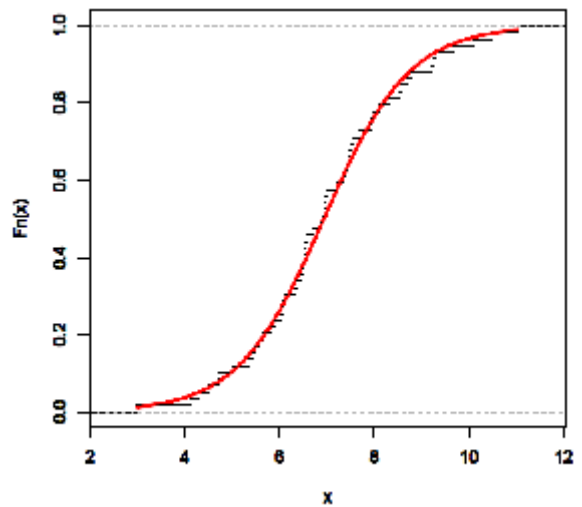


Fig. 6. The graph of empirical and fitted distribution function

A further support for this finding can be obtained by inspecting the probability-probability (P-P) and quantile-quantile (Q-Q) plots. A P-P plot depicts the points:

$$\left(F(x_{(i)}), F(x_{(i)}; \hat{\theta}) \right) \quad ; i = 1, 2, \dots, n$$

Where $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$ and $X_{(i)}$ are the order statistics,

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I(X \leq x)$$

is the empirical distribution function and $I(\cdot)$ is the indicator function.

A Q-Q plot depicts the points:

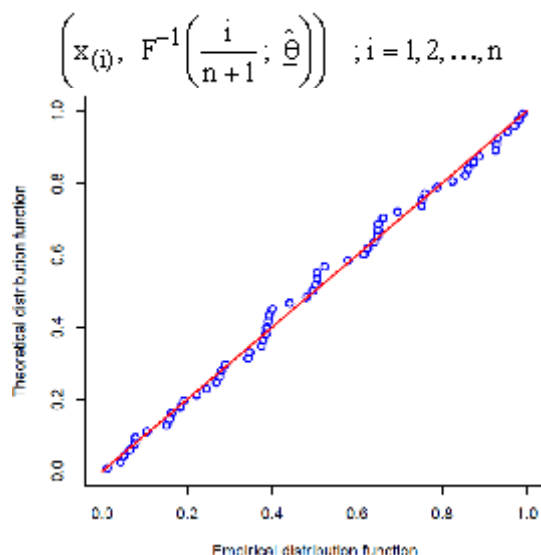


Fig. 7. Probability-Probability(P-P) plot using MLEs as estimate

The P-P and Q-Q plots for the fitted model are shown in Fig. 7 and Fig. 8. It is evident that the fit of the Perks distribution is good.

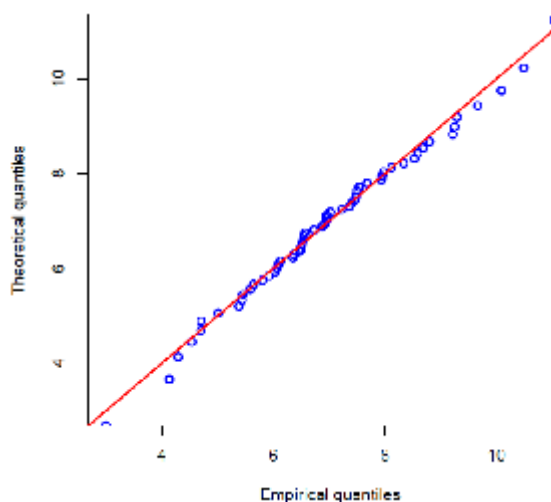


Fig. 8. Quantile-Quantile(Q-Q) plot using MLEs as estimate

Bayesian Analysis

OpenBUGS script for the Bayesian analysis of Perks distribution

```

model
{
  for(i in 1 : N)
  {
    x[i] ~ dperks1_T(alpha, beta)
    reliability[i] <- R(x[i], x[i])           # to estimate reliability
    f[i] <- density(x[i], x[i])              # to estimate density
    hrf[i] <- hrf(x[i], x[i])                # to estimate hazard rate
  }
  # To predict the data set
  ep[i] <- (i - 0.5) / N
  x.new[i] <- (1.0 / beta) * log((1.0 / alpha) * (((1 + alpha) / (1.0 - ep[i])) - 1.0))
}
# Prior distributions of the model parameters
alpha ~ dgamma(0.001, 0.001)
beta ~ dgamma(0.001, 0.001)
}

```

DATA

list(N=59, x=c(2.997, 4.137, 4.288, 4.531, 4.700, 4.706, 5.009, 5.381, 5.434, 5.459, 5.589, 5.640, 5.807, 5.923, 6.033, 6.071, 6.087, 6.129, 6.352, 6.369, 6.476, 6.492, 6.515, 6.522, 6.538, 6.545, 6.573, 6.725, 6.869, 6.923, 6.948, 6.956, 6.958, 7.024, 7.224, 7.365, 7.398, 7.459, 7.489, 7.495, 7.496, 7.543, 7.683, 7.937, 7.945, 7.974, 8.120, 8.336, 8.532, 8.591, 8.687, 8.799, 9.218, 9.254, 9.289, 9.663, 10.092, 10.491, 11.038))

Init1

list(alpha=0.001, beta= 0.5) # Chain 1

Init2

list(alpha=0.0001, beta= 1.0) # Chain 2

We assume the independent gamma priors for $\alpha \sim G(a_1, b_1)$ and $\beta \sim G(a_2, b_2)$ with hyper parameter values $a_1 = b_1 = a_2 = b_2 = 0.001$. We run the model to generate two Markov Chains at the length of 40,000 with different starting points of the parameters. We have chosen initial values ($\alpha = 0.001, \beta = 0.5$) for the first chain and ($\alpha = 0.0001, \beta = 1.0$) 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 7000 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

$$(\alpha_1^{(j)}, \beta_1^{(j)}) ; j = 1, \dots, 7000 \text{ from chain 1 and}$$

$$(\alpha_2^{(j)}, \beta_2^{(j)}) ; j = 1, \dots, 7000 \text{ from chain 2.}$$

The chain 1 is considered for convergence diagnostics plots and the visual summary. The numerical summary is presented for both the chains.

Convergence diagnostics

The first step in making an inference from an MCMC analysis is to ensure that an equilibrium distribution has indeed been reached by the Markov chain, i.e., the chain has converged. There are several ways to check for convergence. Most methods use at least two parallel chains. Visual checks are routinely used to confirm convergence.

The sequential plot of parameters is the plot that most often exhibits difficulties in the Markov chain. Sequential realization of the parameters α and β for iteration number 3001 to 8000 can be observed in Fig. 9. The dotted line represents the mean of all 5000

sample values. In this case Markov chain seems to be mixing well enough and is likely to be sampling from the stationary distribution.

History(Trace) plot

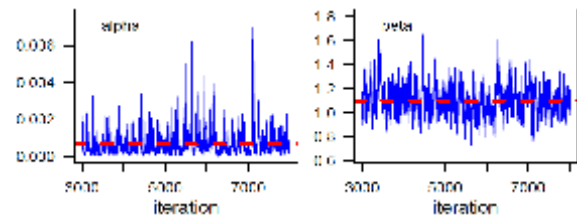


Fig. 9. Sequential realization of the parameters α and β .

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 Fig. 10 indicating the convergence of the chain.

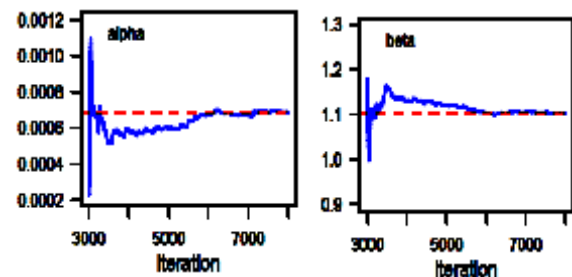


Fig. 10. The Ergodic mean plots for α and β

The dotted line represents the mean of sample values. The plot shows that the ergodic mean stabilizes as chain advances.

Posterior analysis

(a) Numerical summary

In Table 2, we have considered various quantities of interest and their numerical values based on MCMC sample of posterior characteristics for Perks distribution under gamma priors.

Table 2. Numerical summaries based on MCMC sample of posterior characteristics for Perks distribution

Characteristics	Chain 1		Chain 2	
	alpha	beta	alpha	beta
Mean	0.000655	1.1101	0.000679	1.1046
Standard Deviation	0.000636	0.1236	0.000689	0.1213
Minimum	0.000011	0.7322	0.000008	0.6432
2.5th Percentile($P_{2.5}$)	0.000076	0.8798	0.000081	0.8750
First Quartile (Q_1)	0.000253	1.0240	0.000263	1.0210
Median	0.000461	1.1060	0.000476	1.1010
Third Quartile (Q_3)	0.000834	1.1910	0.000850	1.1850
97.5th Percentile($P_{97.5}$)	0.002321	1.3600	0.002475	1.3530
Maximum	0.007013	1.6420	0.011650	1.6500
Mode	0.000232	1.1100	0.000254	1.1017
Skewness	2.812512	0.2051	3.994323	0.1271
95% Credible Interval	(7.620e-05, 0.00232)	(0.8798, 1.36)	(8.098e-05, 0.0025)	(0.875, 1.353)
95% HPD Credible Interval	(1.129e-05, 0.00184)	(0.8691, 1.345)	(2.78e-05, 0.00196)	(0.8687, 1.345)

Highest probability density (HPD): The algorithm described by Chen and Shao (1999) is used to compute the HPD intervals under the assumption of unimodal marginal posterior distribution

(b) Visual summary

Box plots

The boxes represent inter-quartile ranges and the solid

black line at the (approximate) centre of each box is the median; the arms of each box extend to cover the central 95 per cent of the distribution - their ends correspond, therefore, to the 2.5% and 97.5% quantiles. (Note that this representation differs somewhat from the traditional.)

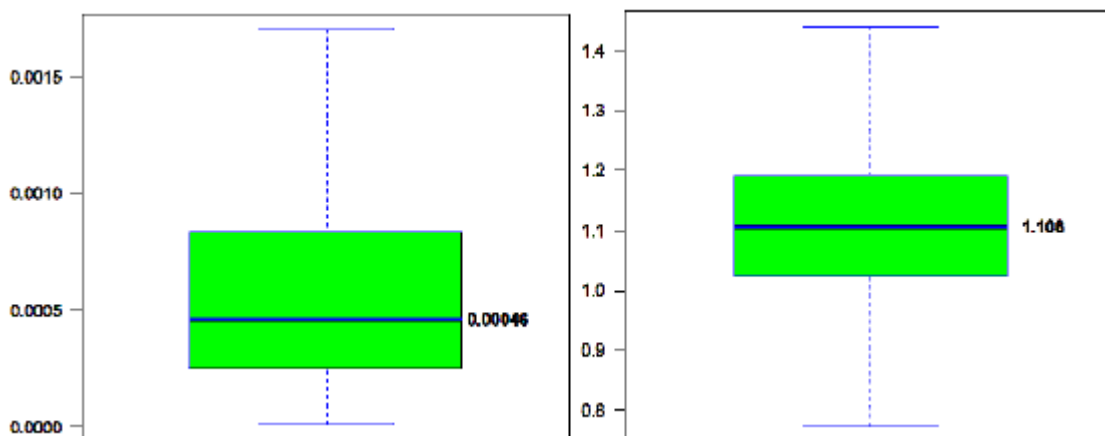


Fig. 11. The boxplots for alpha and beta

Kernel density estimates

Histograms can provide insights on skewness, behaviour in the tails, presence of multi-modal behaviour, and data outliers; histograms can be compared to the fundamental shapes associated with standard analytic distributions.

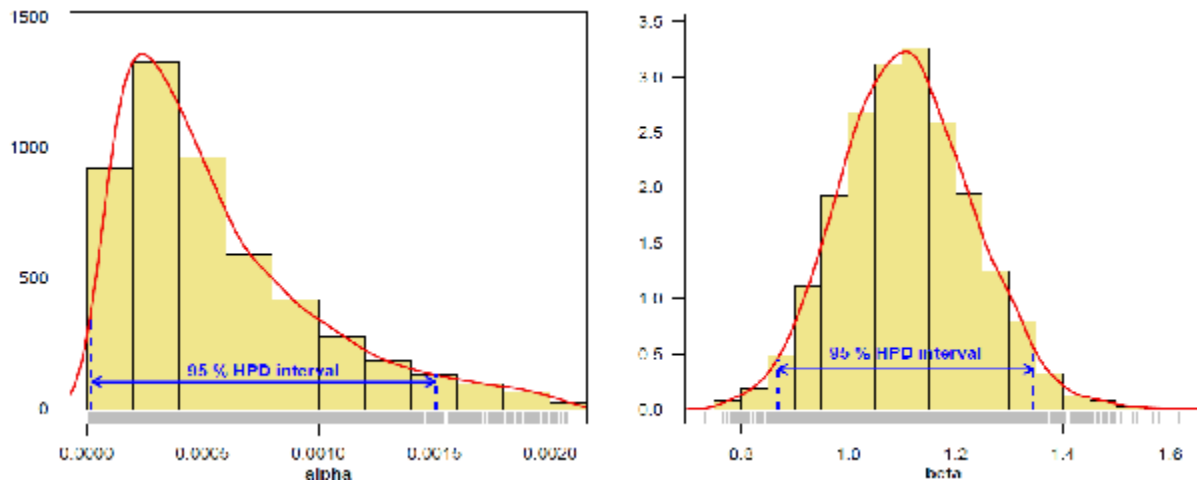


Fig. 12. Kernel density estimate and HPD intervals of α (left panel) and β (right panel)

Figure 12 gives the visual summary based on MCMC output for the parameters. We have plotted histogram, the kernel density estimate, 95% HPD intervals and actual realizations along x-axis (“rug” plot) for α (left panel) and β (right panel). The kernel density estimates have been drawn using R with the assumption of Gaussian kernel and properly chosen values of the bandwidths. It can be seen that β is symmetric whereas α shows positive skewness.

Comparison with MLE

We have used graphical method for the comparison of Bayes estimates with ML estimates. In Fig. 13, the

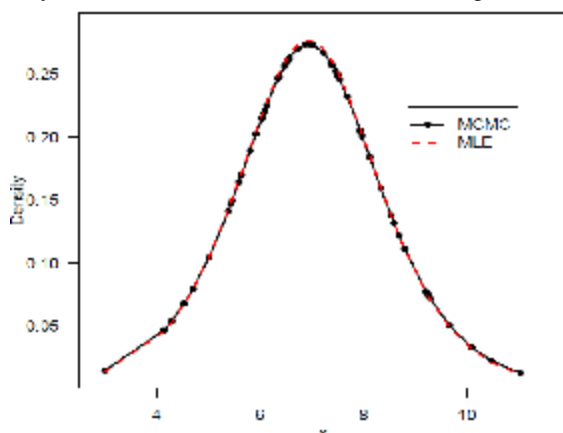


Fig. 13. The density functions using ML and Bayesian estimates

density functions $f(x; \hat{\alpha}, \hat{\beta})$ using MLEs and Bayesian estimates, computed via MCMC samples under gamma priors, are plotted. It is clear from the Fig. 13 that the MLEs and the Bayes estimates with respect to the gamma priors are quite close and fit the data very well.

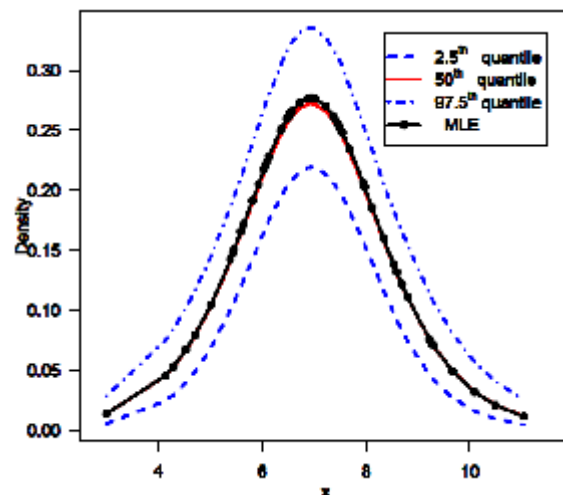


Fig. 14. Density estimates

A further support for this finding can be obtained by inspecting the Fig. 14. In Figure 14, we have plotted 25th, 50th and 97.5th quantiles of the estimated density

based on MCMC sample $(\alpha_1^{(j)}, \beta_1^{(j)}); j=1, \dots, 7000$.

Here, the density is computed at each data point for 7000 posterior samples. 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.

Estimation of reliability function

In this section, our main aim is to demonstrate the effectiveness of proposed methodology. For this we have estimated the reliability function using MCMC samples under gamma priors. 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 Fig. 15 exhibits the estimated reliability function (dashed blue line: 2.5th and 97.5th quantiles; solid red line: 50th quantile) using Bayes estimate based on MCMC output under independent gamma priors for both the parameters and the empirical reliability function (black solid line). The Fig. 15 shows that reliability estimate based on MCMC is very closed to the empirical reliability estimates.

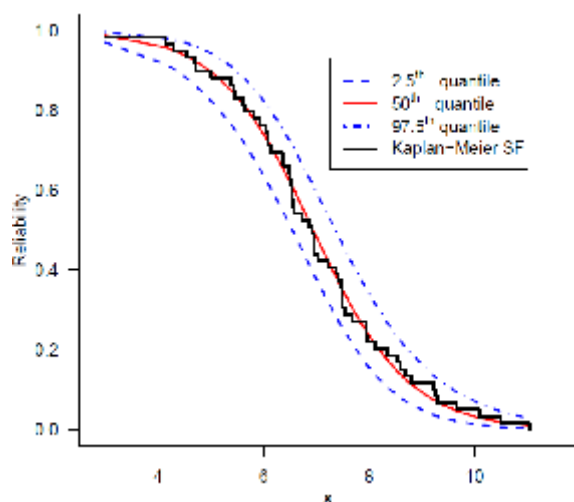


Fig. 15. Reliability function estimate using MCMC and Kaplan-Meier estimate

Estimation of hazard and reliability at $X_{(7)}: t = 5.009$

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. Indeed, the MCMC samples may be used to completely summarize the posterior uncertainty about the parameters α and β 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 $t=5.009$ (at the 7th observed data point).

The marginal posterior density estimates of the reliability and hazard functions and their histograms based on samples of size 7000 are shown in Fig. 16 and

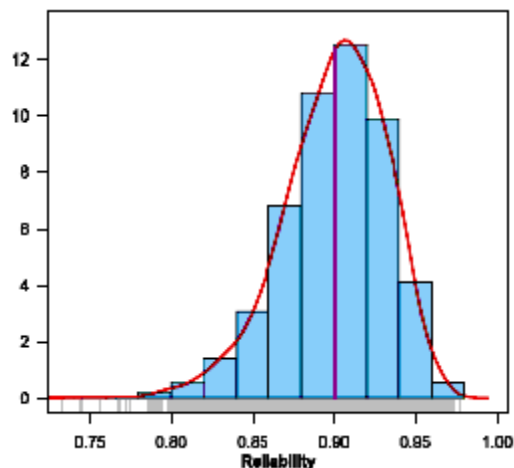


Fig. 16. Visual summary of α

17 using the Gaussian kernel. It is evident from the estimates that the marginal distribution of reliability is negatively skewed whereas hazard is positively skewed.

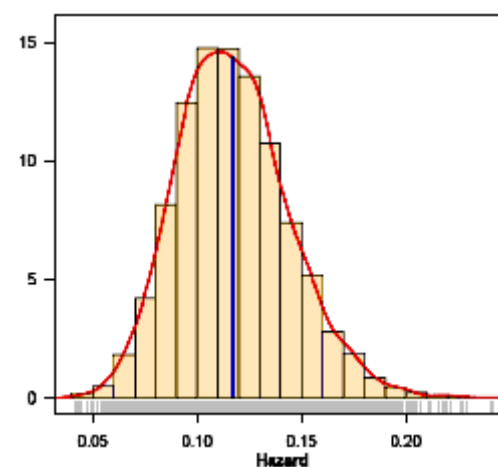


Fig. 17. Visual summary of α

The MCMC results of the posterior mean, median, mode, standard deviation(SD)and skewness of reliability and hazard functions are displayed in Table

3 The ML estimates of reliability and hazard function at $t=5.009$ are computed using invariance property of the MLE.

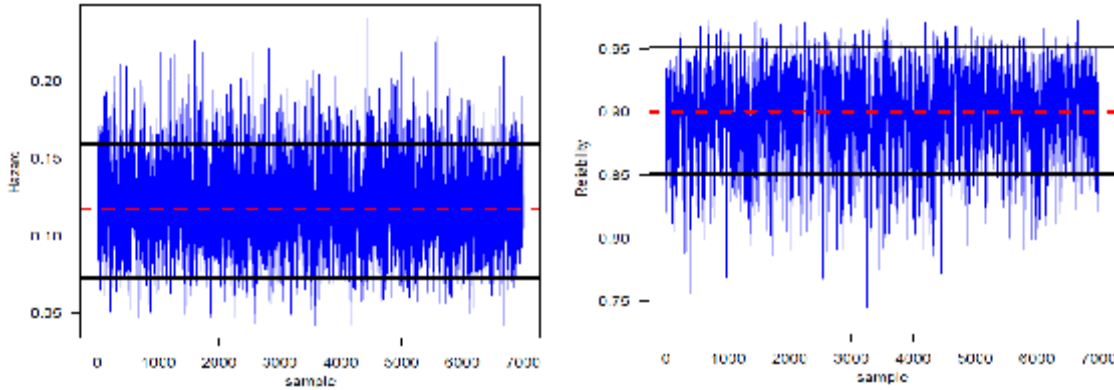


Fig. 18. MCMC output of $h(t = 5.009)$ and $R(t = 5.009)$. Dashed line(...) represents the posterior median and solid lines (-) represent lower and upper bounds of 90% probability intervals (HPD)

Table 3. Posterior summary of hazard and reliability

Parameters	MLE	Based on MCMC output				
		Mean	Median	Mode	Sd	Skewness
$h(t = 5.009)$	0.1162	0.1173	0.1155	0.1113	0.0268	0.4304
$R(t = 5.009)$	0.9027	0.9003	0.9032	0.9069	0.0321	-0.6168

A trace plot is a plot of the iteration number against the value of the draw of the parameter at each iteration. Fig. 18 display 7000 chain values for the reliability $R(t=5.009)$ and hazard $h(t=5.009)$ functions, with their sample median and 90% credible intervals.

Model Compatibility

Posterior Predictive Checks

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 Gelman (2003) and Gelman *et al.* (2004). We do this by comparing the posterior predictive simulations with the data.

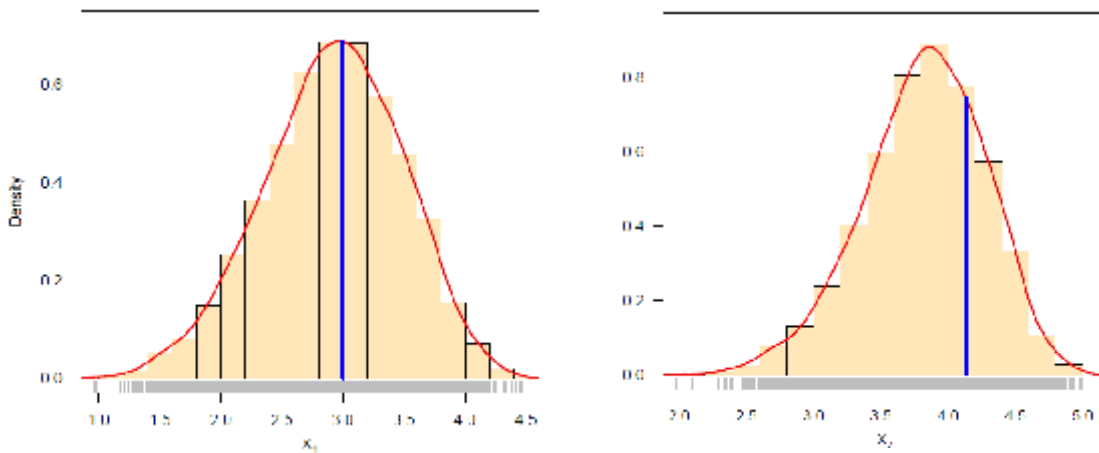


Fig. 19. Density estimates of the two smallest order future observations, vertical lines represent corresponding observed values

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 favourably, the assumed model can be considered to

be an appropriate choice for the data in hand, Gupta *et al.* (2008). 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.

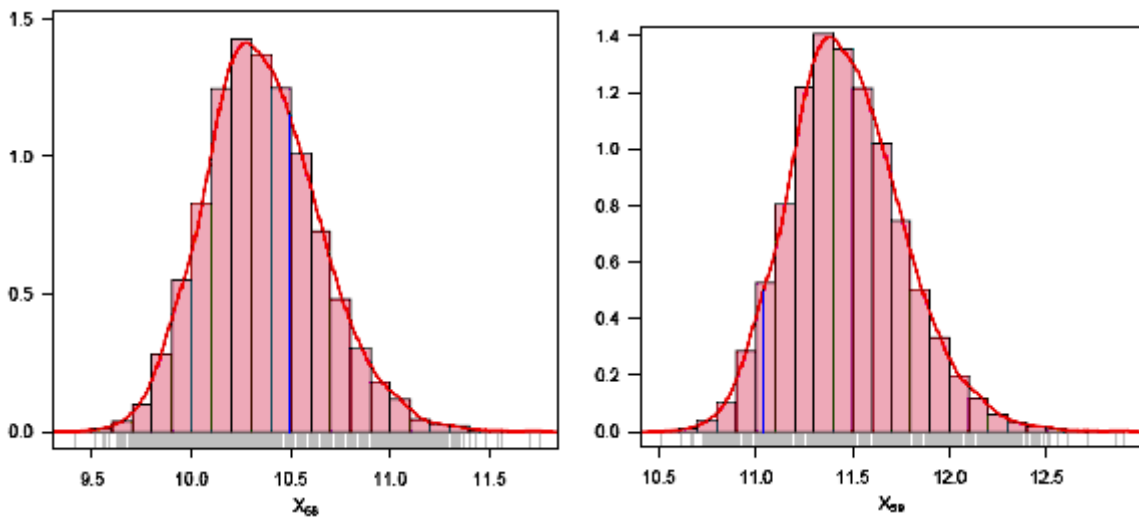


Fig. 20. Density estimates of the two largest order future observations, vertical lines represent corresponding observed values

Comparison of empirical distribution function plots based on the observed and the predictive data may be considered as an informal way to check discrepancies between the data and the model.

To obtain further clarity on our conclusion for the study of model compatibility, we have considered plotting of density estimates of two smallest and two largest replicated future observations from the model with superimposed corresponding observed data. For this purpose, 5000 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.

The MCMC results of the posterior mean, median, mode of two smallest ($X_{(1)}$ and $X_{(2)}$) and two largest

($X_{(58)}$ and $X_{(59)}$) are displayed in Table 4.

Table 4. Posterior characteristics

	Observed	Mode	Mean	Median
$X_{(1)}$	2.997	2.978	2.917	3.325
$X_{(2)}$	4.137	3.868	3.816	4.144
$X_{(58)}$	10.491	10.277	10.362	10.54
$X_{(59)}$	11.038	11.381	11.469	11.65

Density estimates based on replicated future data sets are shown in Figure 19 and Fig. 20. Fig. 20 represents the estimates corresponding to largest two predictive observations, whereas the same for smallest two observations is shown in Fig. 19. The corresponding observed values are also shown by means of vertical lines.

In fact, we have predicated the entire data set. Fig. 21 represents the Q-Q plot of predicted quantiles vs. observed quantiles. We, therefore, conclude that the Perks model is compatible with the given data set.

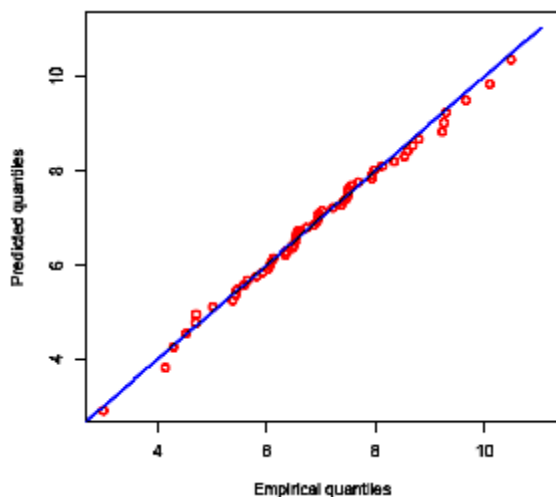


Fig. 21. Q-Q plot of predictive quantiles versus empirical quantiles

We have discussed the Markov chain Monte Carlo (MCMC) method to compute the Bayesian estimates the parameters, hazard and reliability functions of Perks 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. The MCMC method provides an alternative method for parameter estimation of the Perks distribution. It is more flexible when compared with the traditional methods such as MLE method. Moreover, 'exact' probability intervals are available rather than relying on estimates of the asymptotic variances. Indeed, the MCMC sample may be used to completely summarize posterior distribution about the parameters, through kernel estimation. This is also true for any function of the parameters such as reliability and hazard functions.

We have applied the developed techniques on a real data set. Thus, the tools developed can be applied for full Bayesian analysis of Perks distribution.

References

Aarset, M.V. 1987. How to identify a bathtub hazard function. *IEEE Transactions on Reliability* **36**:106-108.
 Albert, J. 2009. *Bayesian computation with R*. 2nd edition. Springer, New York.

Chen, M.H. and Q.M. Shao. 1999. Monte Carlo estimation of Bayesian credible intervals and HPD intervals. *Journal of Computational and Graphical Statistics*. **8**(1):69-92.
 Gamerman, D. and H.F. Lopes. 2006. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, (2nd edition). Taylor & Francis, Boca, Raton, FL.
 Gelfand, A.E. and A.F.M Smith. 1990. Sampling based approach to calculating marginal densities. *Journal of the American Statistical Association* **85**:398-409.
 Gelman, A. 2003. A Bayesian formulation of exploratory data analysis and goodness-of-fit testing. *International Statistical Review* **71**(2):369-382.
 Gelman, A., J. Carlin, H. Stern and D. Rubin. 2004. *Bayesian data analysis*. Second Edition, London, Chapman & Hall.
 Geman, S. and D. Geman. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian Restoration of images. *IEEE Transactions of Pattern Analysis and Machine Intelligence* **6**:721-741.
 Gupta, A., B. Mukherjee and S.K. Upadhyay. 2008. A Bayes study using Markov Chain Monte Carlo simulation. *Reliability Engineering & System Safety* **93**:1434-1443.
 Haberman, S. and A. Renshaw. 2011. A comparative study of parametric mortality projection models. *Insur. Math. Econ.*, **48**:35-55.
 Hamada, M.S., A.G. Wilson, C.S. Reese and H.F. Martz. 2008. *Bayesian Reliability*. Springer, New York.
 Hoff, P.D. 2009. *A first course in Bayesian statistical methods*. Springer, New York.
 Kumar, V. 2010. Bayesian analysis of exponential extension model, *J. Nat. Acad. Math.* **24**:109-128.
 Kumar, V. and U. Ligges. 2011. *reliaR: A package for some probability distributions*. <http://cran.r-project.org/web/packages/reliaR/index.html>.
 Kumar, V., U. Ligges and A. Thomas. 2010. *ReliaBUGS User Manual: A subsystem in OpenBUGS for some statistical models*. version 1.0, OpenBUGS 3.2.1, <http://openbugs.info/w/Downloads/>
 Lange, K. 1999. *Numerical analysis for statisticians*. Springer, New York.
 Lawless, J. F. 2003. *Statistical models and methods for lifetime data*. 2nd ed., John Wiley and Sons, New York.
 Lunn, D. 2010. Recent developments in the BUGS software. *ISBA Bulletin* **17**(3):16-17.
 Lunn, D.J., A. Andrew, N. Best and D. Spiegelhalter. 2000. WinBUGS – A Bayesian modeling framework: Concepts, structure, and extensibility, *Statistics and Computing* **10**:325-337.
 Lunn, D.J., C. Jackson, N. Best, A. Andrew and D. Spiegelhalter. 2013. *The BUGS book: A practical*

- introduction to Bayesian analysis*. Chapman & Hall/CRC, London, UK.
- Marshall, A.W. and I. Olkin. 2007. *life distributions: Structure of nonparametric, semiparametric and parametric families*. Springer, New York.
- Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller. 1953. Equations of state calculations by fast computing machines. *Journal Chemical Physics*, 21, pp. 1087–1091.
- Nadarajah, S. and S.A.A. Bakar. 2012. *ActuDistns : A new R package for actuarial survival models*. <http://cran.r-project.org/web/packages/ActuDistns/index.html>.
- Ntzoufras, I. 2009. *Bayesian modeling using WinBUGS*. John Wiley & Sons, New York.
- Perks, W. 1932. On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries* 43:12–57.
- R. Development Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Richards, S.J. 2012. A handbook of parametric survival models for actuarial use. *Scandinavian Actuarial Journal*. pp. 1-25.
- Richards, S.J. 2008. Applying survival models to pensioner mortality data. *Bra. Actuarial Journal* 14: 257–303
- Rizzo, M.L. 2008. *Statistical computing with R*. Chapman & Hall/CRC.
- Robert, C.P. G and Casella. 2010. *Introducing Monte Carlo Methods with R*. Springer, New York.
- Thomas, A., B. O'Hara, U. Ligges and S. Sturtz. 2006. Making BUGS open. *R. News* 6:12–17.
- Thomas, A. 2010. *OpenBUGS Developer Manual*. ver.3.1.2, <http://www.openbugs.info/>
- Tukey, J.W. 1977. *Exploratory data analysis*, Reading, Mess, Addison-Wesley.
- Yee, T.W. 2012. The VGAM Package, *R News* 8:28-39.