Journal of Data Science **12**(2014), 157-173

Estimation for the Parameter of Poisson-Exponential Distribution under Bayesian Paradigm

Sanjay Kumar Singh*, Umesh Singh and Manoj Kumar
Banaras Hindu University

Abstract: The present paper deals with the maximum likelihood and Bayes estimation procedure for the shape and scale parameter of Poisson-exponential distribution for complete sample. Bayes estimators under symmetric and asymmetric loss function are obtained using Markov Chain Monte Carlo (MCMC) technique. Performances of the proposed Bayes estimators have been studied and compared with their maximum likelihood estimators on the basis of Monte Carlo study of simulated samples in terms of their risks. The methodology is also illustrated on a real data set.

Key words: Bayes estimators, complementary risk, entropy loss function, Gibbs sampling, loss function, Metropolis-Hasting algorithm, prior distribution.

1. Introduction

In life testing, exponential distribution is one of the most discussed distributions due to its simplicity and easy mathematical manipulations. However, its use is inappropriate in those situations where associated hazard rate is not constant. A number of life time distributions having non constant hazard rate are available in the literature e.g., gamma, Weibull, exponentiated exponential, etc. These distributions are generalization of exponential distribution and possess increasing, decreasing or constant hazard rate depending on the value of the shape parameters and reduce to exponential distribution for their specific choices. A modification in exponential distribution is proposed by Kuş (2007) to get a decreasing failure rate distribution. Barreto-Souza and Cribari-Neto (2009) generalized the distribution proposed by Kuş (2007) by including a power parameter. Cancho *et al.* (2011) proposed a new family of distribution, called Poisson-exponential (PE) distribution based having increasing failure rate. The

*Corresponding author.

motivation for the proposed family of distribution is related to the study of competing risk (CR) problems in presence of latent risks (see, Louzada-Neto, 1999) i.e., for those situations when only life-time values are observed but no information is available about the factors responsible for component failures.

Louzada-Neto *et al.* (2011) studied the statistical properties of PE distribution and discussed about the Bayes estimators of its parameters under squared error loss function (SELF), but paid no attention to the maximum likelihood estimators. Further in life testing problems, over estimation and under estimation of equal magnitude cannot be of equal consequence and hence, an asymmetric loss seems to be more justified for life testing and reliability problems as compared to SELF. Thus, our aim in this paper is to obtain the maximum likelihood estimators and Bayes estimators of the parameters under symmetric and asymmetric loss function for Poisson-exponential distribution and to compare the proposed estimators with maximum likelihood estimators in terms of their risks.

The paper is organized as follows. In Section 2, we have discussed briefly the PE distribution, giving its density along with associated failure rate and survival functions. We have obtained the maximum likelihood estimators of the parameters in Section 3. The prior distribution, loss functions and Bayes estimators of the parameters using MCMC technique are presented in Section 4. Comparisons of the estimators and analysis are given in Section 5. In this section the proposed methodology is illustrated through a real data set. The Bayesian prediction is discussed in Section 6 and finally the conclusions are presented in Section 7.

2. The Model

Let X be a non negative random variable denoting the life time of a component/system. The random variable X is said to have a PE distribution with parameter θ and λ, if its probability density function (pdf) is given by

$$f(x) = \frac{\theta\lambda e^{-\lambda x-\theta e^{-\lambda x}}}{1-e^{-\theta}}, \quad x > 0, \tag{1}$$

where λ is the scale parameter, while θ is shape parameter of the distribution. As θ approaches zero, the PE distribution converges to an exponential distribution with parameter λ. Its pdf is decreasing if $0 < \theta < 1$ and unimodal for $\theta \geq 1$. The modal value λe^{-1} is obtained at $x = \log(\theta/\lambda)/\lambda$. As pointed out by Louzada-Neto *et al.* (2011), the parameters θ and λ of the distribution have direct interpretation in terms of CR. In fact θ represents the mean of the number of CR, whereas λ denotes the lifetime failure rate.

The survival (or reliability) function of the PE distribution is given by

$$S(x) = \frac{1-e^{-\theta e^{-\lambda x}}}{1-e^{-\theta}}, \quad x > 0, \tag{2}$$

and the hazard function is

$$h(x) = \frac{\theta\lambda e^{(-\lambda x - \theta e^{-\lambda x})}}{1 - e^{-\theta e^{-\lambda x}}}, \quad x > 0. \tag{3}$$

The hazard function (3) is increasing. The initial and long term hazard values are finite and are given by $h(0) = \lambda\theta/(e^{\theta} - 1)$ and $h(\infty) = \lambda$. For other details about PE distribution see Ristić and Nadarajah (2010).

3. Maximum Likelihood Estimators

Suppose that $X_1, X_2, \cdots, X_n$ be a random sample of size n drawn from a population having pdf (1). Then the likelihood function can easily be obtained as follow:

$$L(\theta, \lambda \,|\, X) = \exp\left\{ n\log(\theta\lambda) - \lambda\sum_{i=1}^{n} x_i - \theta\sum_{i=1}^{n} e^{-\lambda x_i} - n\log(1 - e^{-\theta}) \right\}. \tag{4}$$

The log of likelihood (4) is

$$\log L = n\log(\theta\lambda) - \lambda\sum_{i=1}^{n} x_i - \theta\sum_{i=1}^{n} e^{-\lambda x_i} - n\log(1 - e^{-\theta}). \tag{5}$$

Differentiating (5) w.r.t. (with respect to) to θ and λ and equating the derivatives to zero, we get the following normal equations:

$$\frac{n}{\theta} - \sum_{i=1}^{n} e^{-\lambda x_i} - \frac{ne^{-\theta}}{(1 - e^{-\theta})} = 0, \tag{6}$$

$$\frac{n}{\lambda} - \sum_{i=1}^{n} x_i + \theta\sum_{i=1}^{n} x_i e^{-\lambda x_i} = 0. \tag{7}$$

The normal equations (6) and (7) are implicit system of equations in θ and λ. It can not be solved analytically. Therefore, we propose to use fixed point iteration method for solving these equation numerically, for maximum likelihood estimate (mle) of θ and λ. For details about the proposed method readers may refer Jain *et al.* (1985).

4. Loss Function, Prior and Posterior Distribution

SELF is frequently used for the estimation of parameters in classical as well as in Bayesian estimation procedures. No doubt, its use can easily be justified to some extent on the ground of uniformly minimum variance unbiased estimation

(Berger, 1985). However, in practical situation the real loss is often not symmetric i.e., overestimation of a parameter may lead to harder (weaker) consequences than under estimation and vice-a-versa. Therefore, in order to cope up with such situations, the use of asymmetric loss function has been suggested by many authors (c.f., Zellner, 1986; Varian, 1975; Berger, 1985; Singh *et al.*, 2011, etc.). An asymmetric loss function known as the general entropy loss function (GELF), proposed by Calabria and Pulcini (1994) is defined as

$$L_E(\bar{\alpha}, \alpha) \propto \left(\frac{\bar{\alpha}}{\alpha}\right)^c - c \log\left(\frac{\bar{\alpha}}{\alpha}\right) - 1, \tag{8}$$

where $\bar{\alpha}$ is the estimate of parameter α.

This loss function is a generalization of the entropy loss function used by several authors where the shape parameter c is taken equal to 1. The general version (8) allows different shapes of the loss function to meet the practical needs. It may be noted that when $c > 0$, a positive error causes more serious consequence than a negative error and reverse is the situation when $c < 0$. It may easily be seen that the Bayes estimate $\bar{\alpha}_E$ of α under GELF is given as

$$\bar{\alpha}_E = \left[E_\alpha(\alpha^{-c})\right]^{-\frac{1}{c}}, \tag{9}$$

where $E_\alpha(\cdot)$ denotes the posterior expectation of $(\cdot)$. The estimator $\bar{\alpha}_E$ exists only if the expectation in r.h.s. of (9) exists and is non negative.

Another important point of Bayesian point estimation problem is specification of prior distribution for the parameters. Louzada-Neto *et al.* (2011) suggested the use of Jeffrey's prior for λ for given θ and gamma prior for θ, i.e., $g_1(\theta) \propto \theta^{a-1}\exp(-b\theta)$ and $g_2(\lambda\,|\,\theta) \propto 1/\lambda$. Thus, the joint prior for θ and λ as suggested by them is

$$g(\theta, \lambda) \propto \frac{\theta^{a-1}}{\lambda}\exp(-b\theta), \quad a > 0, \quad b > 0. \tag{10}$$

The above choice of prior distributions can be justified for the situation when there is no information or very little information about λ. The prior distribution for θ is informative prior and the hyper parameter a and b can be chosen easily if we have prior guess about θ (say, m) with specified confidence expressed as prior variance (say, v). A large prior variance indicates less confidence in prior guess and the resulting prior distribution is relatively flat. On the other hand, small prior variance shows greater confidence in prior guess and gives a peaked prior. However, if we have negligible prior information about θ, a and b, and can be taken very small (close to zero) which will result into the choice of a non informative prior for θ. Once prior guess (m) and prior variance (v) is specified, the values of a and b are obtained by solving $m = a/b$ and $v = a/b^2$, i.e., $a = m^2/v$ and

$b = m/v$. It may also be noted that besides having the flexibility as mentioned above, the prior distribution provides to computational ease also.

Combining the likelihood function, $L(\theta, \lambda \,|\, X)$ from (4) and the prior from (10) through Bayes theorem, we get the joint posterior distribution for θ and λ as follows:

$$\begin{aligned} p(\theta, \lambda \,|\, X) &\propto L(\theta, \lambda \,|\, X) g(\theta, \lambda) \\ &= k^{-1} \frac{\theta^{(n+a-1)} \lambda^{(n-1)}}{(1-e^{-\theta})^n} \exp\left\{ -\lambda \sum_{i=1}^{n} x_i - \theta \left(b + \sum_{i=1}^{n} e^{-\lambda x_i} \right) \right\}, \end{aligned} \tag{11}$$

where

$$k = \int_0^\infty \int_0^\infty \frac{\theta^{(n+a-1)} \lambda^{(n-1)}}{(1-e^{-\theta})^n} \exp\left\{ -\lambda \sum_{i=1}^{n} x_i - \theta \left(b + \sum_{i=1}^{n} e^{-\lambda x_i} \right) \right\} d\theta \, d\lambda.$$

It may be noted here that posterior distribution in (11) is proper see for proof Louzada-Neto *et al.* (2011), although the prior distribution considered for λ is improper non-informative prior. It may also be noted that the constant of proportionality k involves double integral and it is not reducible in nice closed form. For computation of above expression, one way would be to use numerical techniques. An alternative method would be to use simulation techniques. Although generating samples directly from joint posterior density is not possible, a full conditional can be easily written from (11). Therefore, MCMC can be implemented in this case quite routinely, for more details, readers may refer to Chen *et al.* (2000), Upadhyay *et al.* (2009) and Mukherjee *et al.* (2010). We propose the use of Gibbs sampler with Metropolis algorithm to simulate samples from the posterior distribution so that sample-based inferences can be deduced.

The full conditional distribution for θ and λ obtained from (11), are given below:

$$p_1(\theta \,|\, \lambda, X) \propto \theta^{n+a-1} \exp\left\{ -\theta \left(b + \sum_{i=1}^{n} e^{-\lambda x_i} \right) - n \log\left(1 - e^{-\theta}\right) \right\}, \tag{12}$$

$$p_2(\lambda \,|\, \theta, X) \propto \lambda^{n-1} \exp\left\{ -\lambda \sum_{i=1}^{n} x_i - \theta \sum_{i=1}^{n} e^{-\lambda x_i} \right\}. \tag{13}$$

The conditional distribution (12) and (13) do not belong to any known parametric distribution family, therefore we propose the implementation of Metropolis-Hasting algorithm with Gibbs iterations and hence, the following MCMC procedure is to be used for the computation of Bayes estimators of θ and λ:

Algorithm to compute Bayes estimates and confidence interval estimate of θ and λ:

Step 1: Set the initial guess of λ and θ say λ_0 and θ_0.

Step 2: Set $i = 1$.

Step 3: Generate λ_i from $p_2(\lambda \,|\, \theta, X)$ and θ_i from $p_1(\theta \,|\, \lambda, X)$.

Step 4: Repeat Steps 2-3, N times.

Step 5: Obtain the Bayes estimates of λ and θ with under GELF as

$$\hat{\lambda}_G = [E(\lambda^{-c_1}|\text{data})]^{-1/c_1} = \left[\frac{1}{N - N_0} \sum_{i=1}^{N-N_0} \lambda_i^{-c_1}\right]^{-1/c_1} \quad \text{and,}$$

$$\hat{\theta}_G = [E(\theta^{-c_1}|\text{data})]^{-1/c_1} = \left[\frac{1}{N - N_0} \sum_{i=1}^{N-N_0} \theta_i^{-c_1}\right]^{-1/c_1} \quad \text{respectively,}$$

where N_0 is the burn-in-period of Markov Chain. Substituting c_1 equal to -1 in Step 5, we get Bayes estimates of λ and θ under SELF.

Step 6: To compute the HPD (highest posterior density) credible interval of θ, order the MCMC sample of θ (say θ_1, θ_2, θ_3, $\cdots$, θ_N as $\theta_{[1]} < \theta_{[2]} < \theta_{[3]} < \cdots < \theta_{[N]}$). Then construct all the $100(1-\alpha)\%$ credible intervals of θ say $(\theta_{[1]}, \theta_{[N(1-\alpha)-1)]})$, $\cdots$, $(\theta_{[N\alpha]}, \theta_{[N]})$. Here $[X]$ denotes the largest integer less than or equal to X. Then the HPD credible interval of θ is that interval which has the shortest length. Similarly, the HPD credible interval of λ can also be constructed.

Step 7: Using the asymptotic normality property of mles, we can construct approximate $100(1-\alpha)\%$ confidence intervals for λ and θ as

$$\hat{\lambda} \pm z_{\alpha/2}\left(\sqrt{\widehat{\text{Var}}(\hat{\lambda})}\right) \quad \text{and} \quad \hat{\theta} \pm z_{\alpha/2}\left(\sqrt{\widehat{\text{Var}}(\hat{\theta})}\right),$$

where $z_{\alpha/2}$ is the $100(1-\alpha/2)\%$ upper percentile of standard normal variate.

5. Simulation Study

In this section, we shall compare Bayes estimator under GELF with the corresponding Bayes estimator under SELF and mles. The estimators $\hat{\theta}_M$ and $\hat{\lambda}_M$ denotes the mle of the parameters θ and λ respectively while $\hat{\theta}_S$ and $\hat{\lambda}_S$ are corresponding Bayes estimators under SELF and $\hat{\theta}_G$ and $\hat{\lambda}_G$ are corresponding Bayes estimators under GELF. Comparisons are based on simulated risks (average loss over sample space) under GELF. And also we obtained the 95% confidence interval and HPD interval of the parameters θ and λ for different sample sizes. Where

$\hat{\theta}_L^C$, $\hat{\lambda}_L^C$ and $\hat{\theta}_U^C$, $\hat{\lambda}_U^C$ represents the lower and upper limit of confidence interval while $\hat{\theta}_L^H$, $\hat{\lambda}_L^H$ and $\hat{\theta}_U^H$, $\hat{\lambda}_U^H$ represents the lower and upper limit of HPD interval of θ and λ respectively. It may be mentioned here that the expressions for the risks cannot be obtained in closed form. Therefore, the risks of the estimators are estimated on the basis of Monte-carlo simulation study of 1000 samples. It may be noted that the risks of the estimators will depend on value of n, θ, λ, a, b and c. In order to consider variation in the values of these parameters, we have obtained the simulated risk for $n = 30, 60, 90, 100$, $c = -2.5, -1.5, -1, 1, 1.5, 2.5$, $m = 4, 5, 6$ for $v = 1, 8$. The entries in brackets in Tables 2 and 3 denote the risks of the estimators when c is negative and the other non-bracket entries are the risks when c is positive, respectively.

We have generated 1000 samples from (1) for arbitrarily chosen value of θ and λ as 5 and 2 respectively. From Table 1 we observed that the risks of estimators $\hat{\theta}_G$ and $\hat{\lambda}_G$ are found to be smallest that their competing estimators $(\hat{\theta}_S, \hat{\theta}_M)$ and $(\hat{\lambda}_S, \hat{\lambda}_M)$ for all considered the values as c. The performance of estimator estimators $\hat{\lambda}_G$ and for all considered values as c, while in this case $\hat{\theta}_G$ perform well for those negative values of c where magnitude is small. Thus for knowing the estimators θ and λ for different values of other parameters, we fixed c as moderate values ± 1.5. In order to study the effect of variation of sample size n (see Table 2) on the performance of estimator estimators θ and λ have taken prior mean as true value of the parameters θ. When $v = 1$ (showing more confidence in m). From the table we observed that the risks of all the estimators of θ and λ decreases as n increases. It is true for both negative and positive values of c. Furthermore, for $c > 0$, the risks of the proposed estimators $\hat{\theta}_G$ and $\hat{\lambda}_G$ are smallest in comparison to their competing estimators. But for $c < 0$ the estimator $\hat{\theta}_S$ and $\hat{\lambda}_S$ have perform well under both losses, in comparison to their rival estimator.

Table 1: Risks of estimators of θ and λ under GELF for fixed, $\theta = 5$, $\lambda = 2$, $n = 30$, $a = 25$, $b = 5$

c	GELF					
	$R_G(\hat{\theta}_M)$	$R_G(\hat{\theta}_S)$	$R_G(\hat{\theta}_G)$	$R_G(\hat{\lambda}_M)$	$R_G(\hat{\lambda}_S)$	$R_G(\hat{\lambda}_G)$
-2.5	0.341757	0.102691	0.103734	0.083392	0.060644	0.060063
-1.5	0.140006	0.036678	0.036647	0.029918	0.02134	0.021257
-1	0.056947	0.016212	0.016212	0.012135	0.008665	0.008665
1	0.048791	0.016958	0.01626	0.012315	0.008917	0.008844
1.5	0.112657	0.038643	0.036218	0.029871	0.021523	0.021021
2.5	0.324095	0.107173	0.097647	0.082524	0.058721	0.05707

From Table 3, when $c > 0$, the risks of the estimators of λ increases, under both the loss functions with increase in prior mean of θ when prior variance $v = 1$ (small). While, in this situation, in case of estimation of θ, it is observed that

Table 2: Risks of estimators of θ and λ under GELF and SELF for fixed $\theta = 5$, $\lambda = 2$, $a = 25$, $b = 5$, $c = \pm 1.5$

n	GELF					
	$R_G(\hat{\theta}_M)$	$R_G(\hat{\theta}_S)$	$R_G(\hat{\theta}_G)$	$R_G(\hat{\lambda}_M)$	$R_G(\hat{\lambda}_S)$	$R_G(\hat{\lambda}_G)$
30	0.105433 (0.121773)	0.019226 (0.019368)	0.019215 (0.01977)	0.027657 (0.029646)	0.017421 (0.018019)	0.017051 (0.018028)
60	0.059165 (0.058232)	0.018815 (0.018573)	0.019203 (0.018864)	0.013766 (0.015073)	0.009598 (0.010341)	0.009416 (0.01036)
90	0.038623 (0.037964)	0.018606 (0.017112)	0.017083 (0.017366)	0.010001 (0.009719)	0.007691 (0.007248)	0.007558 (0.007268)
100	0.034675 (0.031989)	0.018496 (0.015344)	0.017002 (0.015566)	0.009116 (0.008252)	0.007211 (0.006328)	0.007051 (0.006347)
n	SELF					
	$R_S(\hat{\theta}_M)$	$R_S(\hat{\theta}_S)$	$R_S(\hat{\theta}_G)$	$R_S(\hat{\lambda}_M)$	$R_S(\hat{\lambda}_S)$	$R_S(\hat{\lambda}_G)$
30	2.608063 (2.830213)	0.551406 (0.487099)	0.376561 (0.505589)	0.103177 (0.111107)	0.064418 (0.068513)	0.062047 (0.069215)
60	1.436244 (1.49614)	0.521304 (0.486913)	0.360041 (0.50401)	0.050044 (0.056685)	0.035053 (0.040079)	0.034111 (0.040341)
90	0.91176 (0.948819)	0.451359 (0.482361)	0.310604 (0.498171)	0.035995 (0.035631)	0.028098 (0.027848)	0.027467 (0.028012)
100	0.819651 (0.796515)	0.449559 (0.438331)	0.41009 (0.446692)	0.033087 (0.030546)	0.026494 (0.024469)	0.025788 (0.024608)

the risks of the estimators of θ increases under SELF, but under GELF, it is decreasing and almost all reverse trend is noted for prior variance $v = 8$ (large). For $c < 0$, the risks of the estimator of λ under both losses namely GELF and SELF, when we increase prior mean of θ for either lower or higher prior variance of θ. However, in case of estimation of θ, no definite trend of the magnitude of the risk is found under both losses and for each prior variance also. Furthermore, for $c > 0$, $(\hat{\theta}_G, \hat{\lambda}_G)$ perform well (in sense of having smaller risk), under both losses both prior variances (either small or large), while for reverse situation, $(\hat{\theta}_S, \hat{\lambda}_S)$ perform well.

The 95% HPD intervals are also calculated using MCMC samples for the parameters along with classical 95% confidence interval. The intervals catching the true value of the parameter is also calculated and the results are presented in Table 4, for different sample sizes.

From the above table, it is observed that the average length of the confidence interval and HPD interval decrease when sample size increases. It is also noted from the table that the average length of the HPD interval is smaller than that

Table 3: Risks of estimators of θ and λ under GELF and SELF for fixed, $n = 30$, $c = \pm 1.5$

Prior mean m and var v	GELF					
	$R_G(\hat{\theta}_M)$	$R_G(\hat{\theta}_S)$	$R_G(\hat{\theta}_G)$	$R_G(\hat{\lambda}_M)$	$R_G(\hat{\lambda}_S)$	$R_G(\hat{\lambda}_G)$
$m = 4$, $v = 1$	0.102657 (0.098045)	0.035809 (0.03701)	0.034527 (0.037033)	0.026563 (0.026614)	0.019401 (0.019838)	0.019361 (0.019845)
$m = 5$, $v = 1$	0.102545 (0.097155)	0.03577 (0.035556)	0.033605 (0.035741)	0.026778 (0.028062)	0.019659 (0.020257)	0.019447 (0.020257)
$m = 6$, $v = 1$	0.1022548 (0.113981)	0.034721 (0.038394)	0.033447 (0.038799)	0.02677 (0.028195)	0.020357 (0.020648)	0.019736 (0.021867)
$m = 4$, $v = 8$	0.116567 (0.106712)	0.081378 (0.064518)	0.076 (0.06474)	0.029847 (0.026959)	0.025823 (0.023113)	0.025417 (0.023151)
$m = 5$, $v = 8$	0.114907 (0.106648)	0.081348 (0.063153)	0.073593 (0.063306)	0.029263 (0.027355)	0.02554 (0.023127)	0.025127 (0.023981)
$m = 6$, $v = 8$	0.119997 (0.115041)	0.078151 (0.063813)	0.073796 (0.064225)	0.02729 (0.028884)	0.023988 (0.024351)	0.023488 (0.024356)
Prior mean m and var v	SELF					
	$R_S(\hat{\theta}_M)$	$R_S(\hat{\theta}_S)$	$R_S(\hat{\theta}_G)$	$R_S(\hat{\lambda}_M)$	$R_S(\hat{\lambda}_S)$	$R_S(\hat{\lambda}_G)$
$m = 4$, $v = 1$	2.350557 (2.535097)	0.761138 (0.778047)	0.728847 (0.788698)	0.096196 (0.105062)	0.06898505 (0.07372506)	0.068374 (0.074106)
$m = 5$, $v = 1$	2.515071 (2.61888)	0.840863 (0.871337)	0.773042 (0.887255)	0.098303 (0.1111)	0.07249318 (0.07971482)	0.070591 (0.080399)
$m = 6$, $v = 1$	2.583887 (2.658552)	1.17893 (1.147803)	1.071587 (1.169253)	0.098310 (0.198291)	0.0760667 (0.07546214)	0.072728 (0.0762)
$m = 4$, $v = 8$	2.904274 (2.712018)	2.088657 (1.946832)	1.890934 (1.983546)	0.111935 (0.106461)	0.09727108 (0.08988114)	0.094287 (0.090493)
$m = 5$, $v = 8$	2.867835 (2.7679)	2.087391 (1.959926)	1.889961 (1.944097)	0.108339 (0.115189)	0.09498336 (0.09172668)	0.092033 (0.092398)
$m = 6$, $v = 8$	2.733583 (2.683322)	2.047891 (2.024053)	1.863488 (2.064109)	0.100497 (0.198689)	0.08913714 (0.09695368)	0.085974 (0.097798)

of the confidence interval. The decrease in the average length of the interval due to the use of HPD interval against confidence interval is more for small sample than those for large samples. However we may also observed that when the sample size increases confidence interval, HPD interval and upper limit of both intervals are decreasing but lower limit of these intervals increase as sample size

Table 4: Average length and confidence coefficient of HPD and confidence interval for different sample size

n	95% Confidence interval							
	$\hat{\theta}_L^C$	$\hat{\theta}_U^C$	Length	C.P.	$\hat{\lambda}_L^C$	$\hat{\lambda}_U^C$	Length	C.P.
30	2.136318	9.1592	7.0229	0.97	1.43853	2.73994	1.3014	0.96
60	2.987244	7.5194	4.5321	0.95	1.58737	2.48668	0.8993	0.93
90	3.393616	7.022	3.6284	0.97	1.66415	2.39194	0.7278	0.94
100	3.451435	6.8659	3.4144	0.95	1.6777	2.36884	0.69114	0.93
n	95% HPD interval							
	$\hat{\theta}_L^H$	$\hat{\theta}_U^H$	Length	C.P.	$\hat{\lambda}_L^H$	$\hat{\lambda}_U^H$	Length	C.P.
30	3.9831	6.70515	2.7221	0.98	1.68817	2.4421	0.754	0.96
60	4.2619	6.40527	2.1434	0.97	1.77901	2.3152	0.5362	0.95
90	4.4508	6.2891	1.8383	0.95	1.82609	2.2652	0.4391	0.97
100	4.4691	6.22095	1.7518	0.97	1.83424	2.2519	0.4176	0.96

increases. While the coverage probability (C.P.) does not shows any specific trend with variation of sample.

6. Data Analysis

In this section we re-analyze the data extracted from Lawless (1982) to illustrate our proposed work. The data presented below are the numbers of million revolutions before failure for each of the 23 ball bearing put on a life test: 17.88, 28.92, 33.00, 41.52, 42.12, 45.60, 48.80, 51.84, 51.96, 54.12, 55.56, 67.80, 68.64, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, 173.40.

The sample mean and variance of the data are 72.24 and 1343.64 respectively. Francisco *et al.* (2011) has used this data for fitting PE and exponential distribution and on the basis of deviance information criterion concluded that PE distribution fits better than the exponential distribution. Since we have no prior information available about the parameters, a non-informative Jeffrey's prior for λ and gamma distribution with large variance as prior distribution for θ seems to be most justified. There are many way of choosing the value of hyper parameters of the gamma distribution. One way would be to choose hyper parameter $a = 1$ and $b = 0.0001$ which gives the prior variance to be 106. But in this arbitrary choice the prior mean $m = 1000$ is too large. Therefore, we suggest to obtain the value of hyper parameters a and b such that the prior mean is equal to mle of θ and prior variance moderately large (here we have taken $v = 8$. The mle and Bayes estimators under general entropy loss function for $c_1 = 1.5$ and $c_1 = -1.5$, HPD intervals and confidence intervals for the parameters θ and λ are obtained for both type of choices of hyper parameters and these are presented in Tables 5-7.

Table 5: Estimates of the parameter under general entropy loss function for different prior mean, variance

Prior m & variance (v)	MLE		SELF		GELF			
					$c=-1.5$		$c=1.5$	
	$\hat{\theta}_M$	$\hat{\lambda}_M$	$\hat{\theta}_S$	$\hat{\lambda}_S$	$\hat{\theta}_G$	$\hat{\lambda}_G$	$\hat{\theta}_G$	$\hat{\lambda}_G$
$m=\hat{\theta}_M$, $v=8$	7.330451	0.035827	7.232101	0.035574	7.266457	0.035659	7.055779	0.035142
$m=1000$, $v=1E+6$	7.330451	0.03577	7.37826	0.035773	7.417151	0.035856	7.178585	0.035337

Table 6: Lower, upper and length of 95% HPD and confidence interval for the parameter θ

Prior m & variance (v)	95% HPD interval of θ			95% confidence Inter of θ		
	$\hat{\theta}_L^H$	$\hat{\theta}_U^H$	Length	$\hat{\theta}_L^C$	$\hat{\theta}_U^C$	Length
$m=\hat{\theta}_M$, $v=8$	4.726236	9.9961	5.269864	2.241163	12.42088	10.17972
$m=1000$, $v=1E+6$	4.655145	10.72255	6.067405	2.240973	12.41993	10.17896

Table 7: Lower, upper and length of 95% HPD and confidence interval for the parameter λ

Prior m & variance (v)	95% HPD interval of λ			95% confidence Inter of λ		
	$\hat{\lambda}_L^H$	$\hat{\lambda}_U^H$	Length	$\hat{\lambda}_L^C$	$\hat{\lambda}_U^C$	Length
$m=\hat{\lambda}_M$, $v=8$	0.028409	0.042846	0.014436	0.023703	0.047956	0.024253
$m=1000$, $v=1E+6$	0.028565	0.043494	0.024252	0.023701	0.047954	0.024252

It may be mentioned here that Bayes estimators and HPD intervals have been obtained above using the MCMC procedures described in Section 4. It is well known that MCMC analysis provides reliable results only when the chains have run sufficiently large number of times and reached to the stationary distribution. In the existing literature of MCMC, a number of tools to assess the convergence of chain like auto correlation, mixing of chain and normalizing density are mentioned. The following sample of Figures drawn by using R software, is enough to show that the chains in the present analysis have converged.

(a) (b)

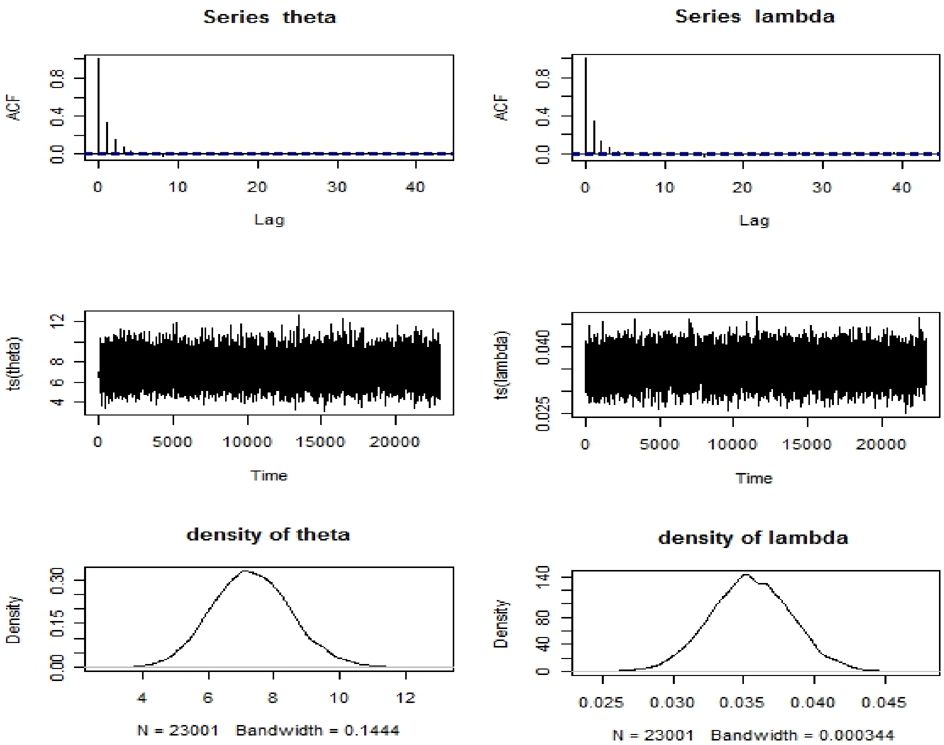

Figure 1: (a) Series of theta, (b) Series of lambda

7. Bayes Prediction

The prediction of future observation on the basis of available information is one of the important topics and it comes up quite naturally in several real life situations. For details see, Geisser (1971), Aitchison and Dunsmore (1975), Al-Hussaini (1999), Smith (1997, 1999), etc. Many of the researchers have discussed the Bayes prediction of future sample based on informative sample, see Ren *et al.* (2006), and Al-Jarallah and Al-Hussaini (2007), etc. A numerical approach to Bayesian prediction for two parameter of Weibull distribution has been discussed by Dellaportas and Wright (1991). Recently, Pradhan and Kundu (2011) have proposed the procedure of estimation of posterior predicting density of future observation, based on the current sample and observed that Gibbs sampling technique can be used quite effectively.

Suppose that we are interested in the predictive density of the r^{th} order statistic $y_{(r)}$ from future sample $\{y_1, y_2, \cdots, y_m\}$ of size m, independent of the informative data $X = \{x_1, x_2, \cdots, x_n\}$. We know that the probability density function of the r^{th} order statistic in the future sample denoted $g_{(r)}(\cdot\,|\,\theta, \lambda)$ is given as

$$g_{(r)}(y\,|\,\theta,\lambda)=\frac{m!}{(r-1)!(m-r)!}[F(y\,|\,\theta,\lambda)]^{r-1}[1-F(y\,|\,\theta,\lambda)]^{m-r}f(y\,|\,\theta,\lambda). \quad (14)$$

Here $f(\cdot\,|\,\theta,\lambda)$ is same as (1) and $F(\cdot\,|\,\theta,\lambda)$ is the corresponding cumulative distribution function.

If we denote the predictive density $y(r)$ as $g^*_{(r)}(y\,|\,X)$ then it can be given as

$$g^*_{(r)}(y\,|\,X)=\int_0^\infty\int_0^\infty g_{(r)}(y\,|\,\theta,\lambda)p(\theta,\lambda\,|\,\text{data})d\theta d\lambda, \quad (15)$$

where $p(\theta,\lambda|X)$ is the joint posterior density of (θ,λ) as given in (11). It is evident that $g^*_{(r)}(y\,|\,X)$ cannot be expressed in nice closed form. However, a simulation consistent estimator of $g^*_{(r)}(y\,|\,X)$ can be obtained using Gibbs sampling procedure described in Section 4.

Suppose (θ_i,λ_i), $i=\{1,2,3,\cdots,M\}$ is an MCMC sample obtained from $p(\theta,\lambda\,|\,X)$ using the Gibbs sampling technique then as suggested by Kundu and Pradhan (2011), a simulation consistent estimator of $g^*_{(r)}(y\,|\,X)$ can be obtained as;

$$g^*_{(r)}(y\,|\,X)=\frac{1}{M}\sum_{i=1}^{M}g_{(r)}(y\,|\,\theta_i,\lambda_i). \quad (16)$$

Similarly, if we want to estimate the predictive distribution of $y_{(r)}$, say $G^*_{(r)}(y\,|\,\text{data})$, a simulation consistent estimator of $G^*_{(r)}(y\,|\,X)$ can be obtained as

$$G^*_{(r)}(y\,|\,X)=\frac{1}{M}\sum_{i=1}^{M}G_{(r)}(y\,|\,\theta_i,\lambda_i), \quad (17)$$

where $G_{(r)}(y\,|\,\theta,\lambda)$ denotes the distribution function of the density function $g_{(r)}(y\,|\,\theta,\lambda)$, i.e.,

$$\begin{aligned}G_{(r)}&(y\,|\,\theta,\lambda)\\&=\frac{m!}{(r-1)!(m-r)!}\int_0^y[F(y\,|\,\theta,\lambda)]^{(r-1)}[1-F(y\,|\,\theta,\lambda)]^{m-r}f(y\,|\,\theta,\lambda)dy\\&=\frac{m!}{(r-1)!(m-r)!}\int_0^{F(y\,|\,\theta,\lambda)}u^{(1-r)}(1-u)^{(m-r)}du. \quad (18)\end{aligned}$$

It should be noted that the same MCMC sample $\{(\theta_i,\lambda_i), i=1,\cdots M\}$ can be used to compute $g^*_{(r)}(y\,|\,X)$ or $G^*_{(r)}(y\,|\,X)$ for all y. For illustration, we would like to estimate the predictive density and distribution for the first order statistic of future sample based on given sample of size 30, generated from PE distribution

with 2 and 5 as the Scale and shape parameter respectively. The generated sample is given below:
1.414, 0.948, 1.686, 0.967, 2.183, 1.229, 0.325, 0.51, 1.229, 0.556, 1.552, 0.191, 1.067, 0.65, 0.88, 0.237, 1.571, 0.199, 0.773, 2.57, 0.909, 0.015, 2.578, 0.7, 0.893, 0.867, 0.76, 2.156, 1.375, 2.024.

Using Gibbs sampler procedure, we obtained 10,000 values of (θ, λ). Based on these, we estimated the predictive density and distribution function for the first order statistics following the procedure described above. The estimated density function and distribution function are presented graphically in Figures 2(a) and 2(b).

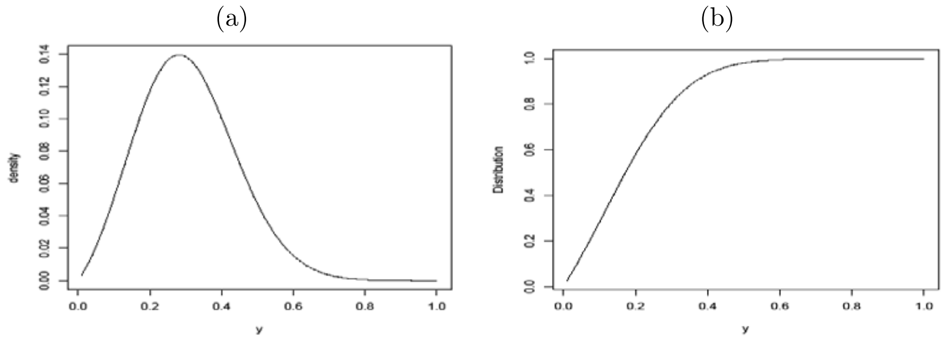

Figure 2: (a) graph of predictive density, (b) graph of predictive distribution

8. Conclusion

In this paper we have considered the problem of estimation of parameters of PE distribution. Procedure for the maximum likelihood estimation and Bayesian estimation has been discussed. On the basis of comparison of risk of the estimators, it is found that Bayes estimator performs better than the maximum likelihood estimator in most of the situations under symmetric and asymmetric loss function. The paper also discusses the classical interval estimation and Bayesian HPD interval estimation. It was noted that Bayesian HPD intervals perform better that classical interval estimation. The paper also includes a procedure for estimation of predictive distribution. From the discussion mentioned above, we may conclude that the Bayes procedures discussed in the paper can be recommended for their use.

Acknowledgements

The authors thank the referees and editor for their valuable comments and suggestions for improving the paper. The third author (Manoj Kumar) thanks UGC-BHU, Varanasi, India, for providing financial assistance. The authors also thank DST-CIMS, B. H. U., Varanasi, India, for providing computational facilities.

References

Aitchison, J. and Dunsmore, I. R. (1975). *Statistical Predictive Analysis.* Cambridge University press, Cambridge.

Al-Hussaini, E. K. (1999). Predicting observables from a general class of distributions. *Journal of Planning and Statistical Inference* **79**, 79-91.

Al-Jarallah, R. A. and Al-Hussaini, E. K. (2007). Bayes inference under a finite mixture of two-compound Gompertz components model. *Journal of Statistical Computation and Simulation* **77**, 915-927.

Barreto-Souza, W. and Cribari-Neto, F. (2009). A generalization of the exponential-Poisson distribution. *Statistics and Probability Letters* **79**, 2493-2500.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis.* Springer, New York.

Calabria, R. and Pulcini, G. (1994). An engineering approach to Bayes estimation for the Weibull distribution. *Microelectronics Reliability* **34**, 789-802.

Cancho, V. G., Louzada-Neto, F. and Barriga, G. D. C. (2011). The Poisson-exponential lifetime distribution. *Computational Statistics and Data Analysis* **55**, 677-686.

Chen, M. H., Shao, Q. M. and Ibrahim, J. G. (2000). *Monte Carlo Methods in Bayesian Computation.* Springer, New York.

Dellaportas, P. and Wright, D. E. (1991). Numerical prediction for two-parameter Weibull distribution. *Statistician* **40**, 365-372.

Geisser, S. (1971). The inferential use of predictive distributions. In *Foundations of Statistical Inference* (Edited by V. P. Godambe and D. A. Sprott), 456-469. Holt, Rinehart and Winston, Toronto.

Jain, M. K., Iyenger, S. R. K. and Jain, R. K. (1984). *Numerical Methods for Scientific and Engineering Computation*, fifth edition. New Age International (P) Ltd., Publishers, New Delhi.

Kuş, C. (2007). A new lifetime distribution. *Computational Statistics and Data Analysis* **51**, 4497-4509.

Lawless, J. F. (1982). *Statistical Models and Methods for Life Time Data.* Wiley, New York.

Louzada-Neto, F. (1999). Modelling lifetime data: a graphical approach. *Applied Stochastic Models in Business and Industry* **15**, 123-129.

Louzada-Neto, F., Cancho, V. G. and Barrigac, G. D. C. (2011). The Poisson-exponential distribution: a Bayesian approach. *Journal of Applied Statistics* **38**, 1239-1248.

Mukherjee, B., Gupta, A. and Upadhyay, S. K. (2010). A Bayesian study for the comparison of generalized gamma model with its components. *Sankhya B* **72**, 154-174.

Pradhan, B. and Kundu, D. (2011). Bayes estimation and prediction of the two-parameter gamma distribution. *Journal of Statistical Computation and Simulation* **81**, 1187-1198.

Ren, C., Sun, D. and Dey, D. K. (2006). Bayesian and frequentist estimation and prediction for exponential distributions. *Journal of Statistical Planning and Inference* **136**, 2873-2897.

Ristić, M. M. and Nadarajah, S. (2010). A new lifetime distribution. Research Report No. 21, Probability and Statistics Group School of Mathematics, University of Manchester, Manchester.

Singh, S. K., Singh, U. and Kumar, D. (2011). Bayesian estimation of the exponentiated gamma parameter and reliability function under asymmetric loss function. *REVSTAT Statistical Journal* **9**, 247-260.

Smith, R. L. (1997). Statistics for exceptional athletics records: letter to the editor. *Applied Statistics* **46**, 123-127.

Smith, R. L. (1999). Bayesian and frequentist approaches to parametric predictive inference (with discussion). In *Bayesian Statistics* (Edited by J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith), Volume 65, 89-612. Oxford University Press, Oxford.

Upadhyay, S. K. and Gupta, A. (2009). A Bayes analysis of modified Weibull distribution via Markov chain Monte Carlo simulation. *Journal of Statistical Computation and Simulation* **80**, 241-254.

Varian, H. R. (1975). A Bayesian approach to real estate assessment. In *Studies in Bayesian Econometrics and Statistics in Honour of Lenonard J. Savage* (Edited by S. E. Fienberg and A. Zellner). Amsterdam, North Holland.

Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *Journal of American Statistical Association* **81**, 446-451.

Received January 18, 2013; accepted July 11, 2013.

Sanjay Kumar Singh
Department of Statistics and DST-CIMS
Banaras Hindu University
Varanasi-221005, Uttar Pradesh, India
singhsk64@gmail.com

Umesh Singh
Department of Statistics and DST-CIMS
Banaras Hindu University
Varanasi-221005, Uttar Pradesh, India
umeshsingh52@gmail.com

Manoj Kumar
Department of Statistics and DST-CIMS
Banaras Hindu University
Varanasi-221005, Uttar Pradesh, India
manustats@gmail.com