Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.12785/jsap/030208

Product Spacings as an Alternative to Likelihood for Bayesian Inferences

Umesh Singh, Sanjay Kumar Singh and Rajwant Kumar Singh∗

Department of Statistics and DST-CIMS. Banaras Hindu University, Varanasi-221005, India

Received: 22 Apr. 2014, Revised: 1 May 2014, Accepted: 2 May 2014
Published online: 1 Jul. 2014

Abstract: The paper aims to study the consequences of using product of spacings (PS) as an alternative to the traditional likelihood under Bayesian set up. For this purpose we have considered the problem of point estimation of the parameter of exponential distribution. We have also obtained the asymptotic and HPD confidence intervals of the parameter. The proposed estimates have been compared with those based on usual likelihood on the basis of simulated samples from exponential distribution.

Keywords: Exponential distribution, Product spacings (PS), Bayes estimates, MLE and Interval Estimation.

1 Introduction

Various classical estimation techniques such as the method of moments, method of least square, method of chi square and method of maximum likelihood estimation (MLE) etc. are discussed in statistical literature. Each of these are having their own advantages and limitations but the most popular method of estimation among these is method of MLE, which can be justified on the ground of its various useful properties like consistency, sufficiency, invariance, asymptotic efficiency and above all its easy computation. The method of MLE works well if each of the contributions to the likelihood function is bounded above. It is true with all discrete distributions, but for continuous distribution, it may not be. Various authors have noted the limitations of MLE in different contexts. Its greatest limitation is that it can not work for 'heavy tailed' continuous distribution with unknown location and scale parameters (Pitman, 1979 , p . 70). It also creates problem in mixture of continuous distributions and in such cases MLE method can break down. It is well known that often, MLE does not give satisfactory estimate for certain three parameter distributions, such as gamma, Weibull, and log normal distributions. In all these cases, the major difficulty is that there are paths in the parameter space with location parameter tending to smallest observation along which the likelihood becomes infinite. Unfortunately in such situations estimates of other parameters becomes inconsistent; see Harter and Moore [4]. Further they reiterated the view of Huzurbazar [21] that no stationary point (and hence no local maximum) can provide a consistent estimator, when the concerned distribution is J-shape, as in the case of Weibull and gamma distribution when the shape parameter is less than unity. Thus, whether a global or a local maximum is considered, MLE is bound to fail in some situations.

In order to overcome these shortcomings and having better applicability in such types of situations which possesses properties similar to MLE, Cheng and Amin [1] introduced the Maximum Product of Spacings (MPS) method as an alternative to MLE for the estimation of parameters of continuous univariate distributions. They proposed to replace the likelihood function by product of spacings and claimed that it retains most of the properties of the method of maximum likelihood. Ranneby [2] independently developed the same method as an approximation to the Kullback-Leibler measure of information. The approach of Cheng and Amin is more intuitively attractive and can, to some extent, be regarded as a pragmatic solution to the problems linked with likelihood (Titterington, [11]), but that of Ranneby is more powerful theoretically and allows the derivation of the properties of MPS estimators. It may be noted that MPS method is especially suited to the cases where one of the parameter has an unknown shifted origin, as it is the case in three parameter lognormal, gamma and weibull distribution or to the distribution having J-shape.

∗ Corresponding author e-mail: rajwant37@gmail.com

The above discussed problem related to likelihood function (LF) is not the matter of concern in classical paradigm only, but it may create problem in the derivation of a posterior density function under Bayesian set up. It is well known that the LF is the probability of observing the given sample. The principle of Maximum likelihood states that the MLE of parameter is that value of parameter which maximizes the likelihood function i.e it is the value that makes the observed data the "most probable". In other words, MLE procedure adjusts the shape of the product of density by tuning the parameter so that its value is maximized for the given sample values. In other words we try to maximize the product of density based on each observation i.e the joint density, where as the principle of MPS is based on the product of successive spacings of observations. It is the product of probabilities of a new observation falling between each of the two neighbouring sample points . It may be noted here that both of these methods rely on joint probabilities of sample observation i.e they work on almost similar principle. Therefore, it seems logical to use product spacing (PS) as an alternative to traditional LF in Bayesian paradigm. It is worthwhile to mention here that under classical set up MPS method provides the estimators which possess most of the large sample optimum properties like sufficiency, consistency and asymptotic efficiency being possessed by MLE (for details see Cheng and Amin [1]).

In certain cases, it is possible to obtain the distributional behaviour of MPS estimator for all sample sizes n; e.g. for the uniform distribution with unknown endpoints, the MPS estimator is precisely the minimum variance unbiased (MVU) estimator and its distribution is known exactly. For a general distribution, however, the small sample behaviour of MPS estimators, like ML estimators, is usually difficult to obtain. However, the asymptotic properties of consistency and asymptotic efficiency are readily obtainable.

The consistency of MPS estimators have been discussed in detail by Cheng and Amin [16] and it is concluded that MPS estimators, when exit are at least asymptotically as efficient as MLE . For distribution where the end points are unknown and the density is J-shaped, the MLE is bound to fail, but MPS gives asymptotically efficient estimators. MPS estimators may not necessarily be function of sufficient statistics in general. However, for the case when the support of density functions are known, MPS estimator will show the same asymptotic properties as ML estimators including asymptotic sufficiency. Through examples Cheng and Amin [1] and Nan Zhang [15] have illustrated unbiasedness, consistency and efficiency properties of MPS. The invariance property of it is same as that of MLE, this is shown by Coolen and Newby [13].

The objective of the present study is to propose the use of PS as an alternative to usual LF in Bayesian paradigm and study the performance of the estimator thus obtained. For this purpose, we have considered the problem of point estimation of the parameter of exponential distribution using PS.

The organisation of rest of the paper is as follows:
Section 2, discusses PS briefly. Estimation procedures are discussed in section 3. It includes the development of point estimators and asymptotic confidence intervals based on PS under classical set up. Further, under Bayesian set-up, PS is proposed as an alternative to traditional likelihood function and Bayes estimators have been obtained. A comparison of the estimators based on simulation study is provided in section 4. Finally, the concluding remark is given in section 5.

2 Product of Spacings

Consider that a random sample $x_1, x_2, \cdots, x_n$ of size n is available from a univariate distribution F($x|\theta$) with corresponding probability density function f($x|\theta$) and it is required to estimate θ. The density is assumed to be strictly positive in an interval (a, b) and zero elsewhere, (a= -∞ and b=∞ may also be taken). Now F($x|\theta$) and f($x|\theta$) are equal to zero for $x < a$:, but F($x|\theta$)=1. and f($x|\theta$)=0 for $x > b$. Let $x_{i:n}$ denote the i^{th} order statistics. The spacings $D_i's$ are defined as follows:

$D_1 = F(x_{1:n}, \theta)$, $D_{n+1} = 1 - F(x_{n:n}, \theta)$, $D_i = F(x_{i:n}, \theta) - F(x_{i-1:n}, \theta), i = 2, 3, \cdots, n$ as the spacings of the sample. Clearly the spacings sum to unity i.e $\sum D_i = 1$. The PS is defined as the product of D_i's i.e. S = $\prod_{i=1}^{n+1} D_i$. The average spacing, denoted by G, can be measured by the geometric mean of the spacings i.e. $G = (S)^{\frac{1}{n+1}}$. Naturally G will be maximum if all D_i's are equal i.e. $F(x_{i:n})$ are equally spaced in the interval [0,1]. If sample in hand is most probable sample (as assumed in justifying the use of likelihood function), it is expected that spacings induced by the sample will be more or less equally spaced. MPS method chooses that value of the parameter θ as its estimate which makes the observed spacings as uniform as possible. Thus one can choose a value of θ which provides the maximum for S or G. Cheng and Amin [1] proposed maximizing G as a method of parameter estimation. It is expected to be as efficient as ML estimation.

If there are ties in the data an anticipated difficulty may arise in drawing inferences based on PS. In such situations, the D_i's corresponding the tied observations would be zero resulting into the corresponding G and S to be zero. One can argue at this stage that at least theoretically there is no chance of ties in the data obtained from a continuous distribution but practically we often encounter with the data set which comprises of repetition of values in the data. But the problem of ties poses no serious problem because it can be easily tackled as discussed below:

Suppose that among the n observations $x_1, x_2, \cdots, x_n$ there are m distinct values expressed in ascending order of their magnitude as $y_{j:m}, j = 1, 2, \cdots, m$. Denote 'a' by $y_{0:m}$ and 'b' by $y_{m+1:m}$ i.e. $F(y_{0:m}) = 0$ and $F(y_{m+1:m}) = 1$. Let l_j denote the number of observations in $(y_{j-1:m}, y_{j:m}]$ i.e. exactly l_j out of n observations are equal to $y_{j:m}$, naturally, $\sum_{j=1}^{m} l_j = n$. To get rid of the problem of ties, one can suggest simply to drop the repeated observations use sample of distinct observations , i.e. use $y_{j:m}$'s only. However, it will result into reduction of the sample size from n to m, leading to the loss of information. In order to retain the maximum information, we can use the method suggested by Shao and Hahn [16]. They argue that since the observations are i.i.d., each of the l_j tied observations has the same probability of occurrence. Thus, their contribution to the product spacings should be equal i.e. corresponding to each of the l_j observations, the contribution should be $\frac{F(y_{j:m}) - F(y_{j-1:m})}{l_j}$ so that the sum of the contributions due to these observations remain $\left[F(y_{j:m}) - F(y_{j-1:m})\right]$. Hence, the sum of the contributions due to all the n observations remains equal to 1. In light of the above, the product spacing, in the presence of ties, can now be easily modified and can be expressed in terms of $y_{j:m}$ as the following (assuming $l_{m+1} = 1$):

$$S = \prod_{j=1}^{m+1} \left[\frac{F(y_{j:m}) - F(y_{j-1:m})}{l_j} \right]^{l_j} \tag{1}$$

The other way to tackle this problem would be to consider that all the equal observations are in fact unequal but differ by the amount smaller than the least count of the measurement and hence noted as equal. Suppose two observations x and y are equal, i.e. x=y. Then we may consider that actually the observations are x and x +dx (dx tending to zero). Hence, such tied observation should contribute to the PS equal to $\lim_{dx \to 0} (F(x+dx) - F(x))$ which can be approximated by f(x) where f(.) denotes the density function corresponding to F. Thus the modified PS can be given as follows:

$$S = \prod_{j=1}^{m+1} \left[F(y_{j:m}) - F(y_{j-1:m})\right] \left[f(y_{j:m})\right]^{l_j - 1} \tag{2}$$

It may be seen that if $l_j = 1$ for all $j's$, the above expressions reduce to the original expression.

3 Point estimation of the Parameter

In this section, we shall try to develop classical as well as Bayes estimators of the parameter of the exponential distribution. The estimators thus obtained will be compared with each other.

3.1 Estimation under classical Paradigm

Under classical paradigm, a number of estimation procedure are available. But we shall be considering here only two of such methods, namely MLE and MPS.

3.1.1 Maximum likelihood estimator

The likelihood function for a sample of size n, say $x_1, x_2, \cdots, x_n$, drawn from exponential distribution having pdf $f(x,\theta)$=$\theta e^{-\theta x}$ is given by

$$L(x,\theta) = \theta^n e^{-\theta \sum_{i=1}^{n} x_i}, x \geq 0 \, , \theta > 0 \tag{3}$$

After differentiating the normal equation with respect to parameter θ and then equating it to zero, we get the well known estimate of θ as

$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n} x_i} \tag{4}$$

3.1.2 Maximum Product of spacings estimator

The CDF of the exponential distribution is $F(x)=1-e^{-\theta x}$ and thus the, spacings can be defined as follows:

$$D_1 = F(x_{1:n}) = \left[1-e^{-\theta x_{1:n}}\right]$$

$$D_{(n+1)} = 1-F(x_{n:n}) = \left[1-e^{-\theta x_{n:n}}\right]$$

$$D_i = F(x_{i:n}) - F(x_{i-1:n}) = \left[e^{-\theta x_{i-1:n}} - e^{-\theta x_{i:n}}\right]$$

Such that $\sum D_i = 1$, MPS estimator chooses θ which maximizes the product of the spacings.

$$S(\theta|x) = \left(\prod_{i=1}^{n+1} D_i\right) = (1-e^{-\theta x_{1:n}})(1-e^{-\theta x_{2:n}})\cdots \tag{5}$$

Taking the logarithm of S we get,

$$G = \sum_{i=1}^{n+1} \ln D_i \tag{6}$$

$$G = \left[\ln(1-e^{-\theta x_{1:n}}) + \sum_{i=2}^{n} \ln(e^{-\theta x_{i-1:n}} - e^{-\theta x_{i:n}}) - \theta x_{n:n}\right] \tag{7}$$

After differentiating the above equation with respect to parameter θ and then equating it to zero, we get the normal equation as follow:

$$\frac{x_{1:n}e^{-\theta x_{1:n}}}{1-e^{-\theta x_{1:n}}} + \sum_{i=2}^{n} \frac{x_{i:n}(e^{-\theta x_{i:n}}) - x_{i-1:n}(e^{-\theta x_{i-1:n}})}{(e^{-\theta x_{i-1:n}} - e^{-\theta x_{i:n}})} - x_{n:n} = 0 \tag{8}$$

The above normal equations cannot be solved analytically. Therefore, we can use any iterative procedure. We propose to use Newton-Rapson method.

3.2 Asymptotic confidence intervals

In this section, we propose the asymptotic confidence intervals using PS, as it was mentioned by Cheng and Amin [1], Anatolyevin and Kosenok [14], Singh et.al. [22] and Ghosh and Jammalamadaka [5] that the MPS method is asymptotically equivalent to MLE. Keeping this in mind, we may propose the asymptotic confidence intervals using PS. The exact distribution of the PS cannot be obtained explicitly. Therefore, the asymptotic properties of PS can be used to construct the confidence intervals for the parameter θ. Anatolyevin and Kosenok [14] show mathematically that $\hat{\theta}_{MPS} = \hat{\theta}_{ML} + o(n^{-\frac{1}{2}})$ i.e. it implies that both are asymptotically equivalent and hence the asymptotic or bootstrap inference around θ based on MPS estimator may be carried out by utilizing the ML asymptotics.
Using the concept of large sample theory we may write the asymptotic confidence interval. We obtain the information matrix I($\hat{\theta}$). We may write the asymptotic confidence interval using MLE as,

$$C.I_{ML} = \left[\hat{\theta}_{ML} \pm 1.96\sqrt{V(\hat{\theta_{ML}})}\right] \tag{9}$$

and utilizing the concept of Anatolyevin and Kosenok [14], we may write the asymptotic confidence interval using PS as

$$C.I_{PS} = \left[\hat{\theta}_{PS} \pm 1.96\sqrt{V(\hat{\theta_{PS}})}\right] \tag{10}$$

3.3 Bayesian Estimation

In this section, we have developed the Bayesian estimation procedure for the parameter θ and HPD interval using PS as an alternative to the usual likelihood. We have taken the idea from a note due to Coolen and Newby [13]. Keeping his idea in mind we have derived the expressions. The product spacings i.e $S(\theta|x) = \prod_{i=1}^{n+1} D_i$ is used in place of the traditional likelihood. In Bayesian analysis, the parameter of interest is assume to be a random variable having some prior distribution. The prior distribution is selected on the basis of type of information available to us. We write $p(\theta)$ for prior density and $S(\underset{\sim}{x}|\theta)$ as PS. According to Bayes theorem, we may write the posterior density using PS as

$$\pi(\theta|\underset{\sim}{x}) = \frac{S(\underset{\sim}{x}|\theta)p(\theta)}{\int_\theta S(\underset{\sim}{x}|\theta)p(\theta)\mathrm{d}\theta} \tag{11}$$

3.3.1 Bayesian Estimation of Parameter θ using PS

In this section, we have provided prior and posterior distribution for considered model for parameter θ. Here, we have considered both informative as well as non informative priors. We have chosen Gamma prior as informative prior and it can be justified on the basis of its flexibility. When we have little or no information about the parameter, a non-informative prior should be used. Jeffery's prior is one of the general class of non-informative priors. Several authors have given justification for using Jaffery's prior for an exponential family. For this reason we are motivated to take Jaffery's non-informative prior for the parameter.

Bayes estimator of θ using an informative prior

Here, we take an informative prior distribution for the parameter θ as the Gamma prior having pdf

$$p_1(\theta) = \frac{\beta^\alpha \theta^{\alpha-1} e^{-\beta\theta}}{\Gamma(\alpha)}\ ;\alpha\ ,\beta\ \geq 0 \tag{12}$$

Then the posterior can be written as,

$$\pi_1(\theta|\underline{\mathrm{x}}) \propto S(\underset{\sim}{x}|\theta)p_1(\theta)$$

$$\pi_1(\theta|\underline{\mathrm{x}}) \propto \left(\prod_{i=1}^{n+1}\left[e^{-\theta x_{i-1:n}} - e^{-\theta x_{i:n}}\right]\right)\theta^{\alpha-1}e^{-\beta\theta} \tag{13}$$

Bayes estimates of θ based on the squared error loss function (SELF) is the posterior mean and can be derived as

$$E(\theta|\underline{\mathrm{x}}) = \int_{\theta=0}^{\infty} \theta\ \pi_1(\theta|\underline{\mathrm{x}})\ d\theta$$

Substituting the value of $\pi_1(\theta|\underline{\mathrm{x}})$ from equation (11), we get

$$E(\theta|\underline{\mathrm{x}}) = \int_{\theta=0}^{\infty} \left(\prod_{i=1}^{n+1}\left[e^{-\theta x_{i-1:n}} - e^{-\theta x_{i:n}}\right]\right)\theta^{\alpha}e^{-\beta\theta}\ d\theta \tag{14}$$

Bayes estimator of θ using a non-informative prior

The pdf of the Jaffery's non-informative prior distribution for the parameter θ is given as

$$p_{12}(\theta) \propto \frac{1}{\theta};\ \theta > 0 \tag{15}$$

Thus, the posterior in this case can be written as,

$$\pi_{12}(\theta|\underline{x}) \propto \left(\prod_{i=1}^{n+1}\left[e^{-\theta x_{i-1:n}} - e^{-\theta x_{i:n}}\right]\right)\frac{1}{\theta} \tag{16}$$

Hence, the Bayes estimates of θ under squared error loss function (SELF) can be derived as

$$E(\theta|\underline{x}) = \int_{\theta=0}^{\infty}\left(\prod_{i=1}^{n+1}\left[e^{-\theta x_{i-1:n}} - e^{-\theta x_{i:n}}\right]\right)d\theta \tag{17}$$

3.3.2 Bayesian Estimation of Parameter θ using usual likelihood

Bayes estimator of θ using an informative prior

Here, we take an informative prior distribution for the parameter θ as the Gamma prior whose pdf is given in equation (12). Combining the likelihood function with the considered prior density of parameter θ, we get the posterior density of θ as,

$$\pi_{21}(\theta|\underline{x}) \propto \theta^{n+\alpha-1}e^{-\theta(\sum x_i+\beta)} \tag{18}$$

Hence, the Bayes estimates of θ under SELF can easily obtained as

$$\hat{\theta}_{BS1} = \frac{\Gamma(n+\alpha+1)}{(\sum x_i+\beta)^{n+\alpha+1}} \tag{19}$$

Bayes estimator of θ using a non-informative prior

In this subsection, we take a non-informative prior distribution for the parameter θ as the Jaffery's prior whose pdf is given in equation (15). Combining the likelihood function with the considered prior density of parameter θ, we get the posterior density of θ as,

$$\pi_{22}(\theta|\underline{x}) \propto \theta^{n-1}e^{-\theta(\sum x_i)} \tag{20}$$

and the Bayes estimates of θ under SELF as

$$\hat{\theta}_{BS2} = \left[\frac{\Gamma(n+1)}{(\sum x_i)^{n+1}}\right] \tag{21}$$

It may be noted here that the solution of the Bayes estimators using PS are not analytically possible, but in this era, it is not a matter of concern due to advancement of numerical methods to solve any numerical equation or integral. Therefore, we use Monte Carlo Markov Chain Method to solve the integral, which is described as follows.

3.3.3 Gibbs Sampling Method

In this subsection, we discuss about the Gibbs sampling procedure to generate sample from posterior under the considered prior for the parameter θ, for more details about MCMC method see Gelfand and Smith (1990) and Singh et al. [19]. Thus utilizing the concept of Gibbs sampling procedure, we generate samples from the posterior density function. For implementing the Gibbs algorithm, the full conditional posterior densities of parameter θ for informative and non informative prior are.

$$\pi_{11}(\theta|\underline{x}) \propto \left(\prod_{i=1}^{n+1}\left[e^{-\theta x_{i-1:n}} - e^{-\theta x_{i:n}}\right]\right)\theta^{\alpha-1}e^{-\beta\theta} \tag{22}$$

and

$$\pi_{12}(\theta|\underline{x}) \propto \left(\prod_{i=1}^{n+1}\left[e^{-\theta x_{i-1:n}} - e^{-\theta x_{i:n}}\right]\right)\frac{1}{\theta} \tag{23}$$

respectively.

The simulation algorithm consists of the following steps.

Step 1 :Start with j = 1 and the initial values of $\{\theta^0\}$
Step 2 :Generate θ from (13) and (16) respectively.
Step 3 :Obtain the posterior sample $\theta_1, \theta_2, \cdots, \theta_M$ by repeating step 2, M times.
Step 4 :The Bayes estimates of the parameter θ under SELF of the parameters can be obtained as the mean of the generated samples from the posterior densities i.e.

$$\hat{\theta} = [E_\pi(\theta|\underline{x})] \cong \left(\frac{1}{M}\sum_{k=1}^{M}\theta_k\right) \tag{24}$$

Step 5 :After extracting the posterior samples, we can easily construct the HPD credible intervals for θ. For this purpose, order θ_i's as $\theta_1 < \theta_2 < \cdots < \theta_M$. Then construct all the $100(1-\beta)$ % credible intervals of θ as

$$\left(\theta_{[1]}, \theta_{[M(1-\beta)+1]}\right), \cdots, \left(\theta_{[M\beta]}, \theta_{[M]}\right)$$

Here, $[x]$ denotes the largest integer less than or equal to x. Then, the HPD credible interval is that interval which has the shortest length.

4 Comparison of estimators

In this section, we compare the various estimators obtained in section 3. This section consists of the simulation results to compare the performance of the classical and Bayesian estimation procedures. The comparison between the PS, MLEs and Bayes estimators using PS and Bayes estimators using usual likelihood of the model parameter has been performed. The comparison is based on the risks (average loss over sample space) under SELF. We have also compared the average lengths of the asymptotic confidence intervals and HPD credible intervals. Here, we investigate the performance of the proposed estimators through a simulation study. For this purpose, we generate the sample of sizes $n = 20$ (small), 30 (medium), and 50 (large) from exponential distribution for fixed values of $\theta = 2$. Asymptotic/HPD intervals and corresponding coverage probabilities (CP) were also calculated for different value of θ ($\theta = 0.5, 1, 2, 3$). For more detail about CP readers are requested to see Krishna and Kumar [18].

The choice of the hyper parameters is the main issue in the Bayesian analysis. Berger [17] argues that when information is not in compact form, it is better to perform the Bayesian analysis under the assumption of non-informative prior. For the choice of hyper parameters of informative prior, we have taken prior means equals to the true values of the parameter with varying prior variances. The prior variance indicates our confidence in the prior guess. A large prior variance shows less confidence in prior guess i.e. the prior distribution is relatively flat. On other hand, small prior variance indicates greater confidence in prior guess. In this study, we have taken prior variance equals to 0.5 (small) and 100 (large).

For obtaining the Bayes estimates, we generate samples from the posterior of θ using the algorithm discussed in Section 3. First thousand MCMC iterations (Burn-in period) have been discarded from the generated sequence. We have also checked the convergence of the sequences of θ for its stationary distributions through different starting values. It was observed that all the Markov chains reached to the stationary condition very quickly.

For the unknown model parameter, we have computed MLEs, PS and Bayes estimates under informative and non-informative priors along with their asymptotic confidence/HPD intervals. We repeat the process 1000 times, and the average estimates, risk of the estimators, and average confidence/HPD intervals are recorded.

On the basis of the simulated results which are summarised in Tables 1-3, the following conclusions can be made as follows:

(i)The risk of all the considered estimators and Bayes estimators decrease as sample size n increases, which is quite obvious. This confirms that all estimators are consistent.

(ii)The Bayes estimator based on PS perform well (in the sense of having smaller risk) in comparison to Bayes estimator based on usual likelihood (UL) and other considered estimators. It is also observed that classical estimator based on PS perform better than Bayes estimator under UL and MLE .

(iii)The risk of the all considered estimators increase as θ increases but trend of the associated risk remains same. It is also noticed that risk of the proposed Bayes estimators are not very much affected by the variation of hyper

parameters. Table 2 shows slight changes in risk of the estimators with variation of prior variance (small, moderate, large). Thus in all the considered cases, the risk of the estimators can be ordered in the following way;
Risk (Bayes PS)$<$ Risk (classical PS) $<$ Risk (Bayes UL)$<$ Risk (MLE).

(iv)The HPD credible intervals have shorter average length than the asymptotic confidence intervals. Further more, for fixed value of θ, it is observed that the average length of confidence interval decreases as sample size n increases and the HPD credible intervals based on PS provide smaller length as compared to Bayes estimator based on UL and other considered estimators. It is interesting to note here that coverage probability i.e. $P(\hat{\theta_L} < \theta < \hat{\theta_U})$ for interval estimators based on PS are more than prefixed confidence interval (Here it is obtained as 100% in all the considered cases). However, the coverage probability for other intervals are either equal to prefixed value (95%) or slightly less than that, see Table 1.

(v)It is observed from table 2 that, the average length of asymptotic/HPD intervals increases as the value of θ increases, but the HPD credible intervals based on PS maintain their superiority in term of the smaller average length among all considered estimators. Further, it is noticed that in some cases Bayes estimators using PS attains 100% coverage probability for some value of θ under different priors.
Similar trend has been observed in case of non informative prior also.

Table 1: Average estimates (in the first row of each cell) under SELF using non-informative prior and corresponding risks, Coverage probabilities and corresponding confidence intervals of the estimators of θ for fixed values of $\theta = 2$ with varying sample size i.e n.

n		estimate	mse	CI/HPD	cv percentage
20	mle	2.12306	0.24868	(1.1884,3.0576)	94.2053
	mps	1.87548	0.19869	(1.0535,2.6974)	92.0530
	Bayes ml	2.11296	0.22533	(1.2414,3.0681)	94.5364
	Bps	1.87466	0.02479	(1.2975,2.4047)	100.0000
30	mle	2.07197	0.15421	(1.3251,2.8187)	94.6000
	mps	1.89446	0.13614	(1.2165,2.5723)	91.3000
	Bayes ml	2.06254	0.14167	(1.3600,2.8232)	95.6000
	Bps	1.89562	0.01724	(1.4245,2.3343)	100.0000
50	mle	2.03955	0.09092	(1.4673,2.6117)	94.2053
	mps	1.92222	0.08526	(1.3894,2.4550)	92.0530
	Bayes ml	2.02355	0.09075	(1.4915,2.6116)	94.5364
	Bps	1.93850	0.00789	(1.5659,2.2842)	100.0000

5 Concluding Remarks

In this paper, we have proposed PS as an alternative to the traditional likelihood in Bayesian set up. We have found that Bayesian procedure under PS provides the better estimates and smaller HPD interval of the unknown parameter of exponential model. On the basis of above simulated results and findings, we recommend to use the PS as an alternative to traditional likelihood in Bayesian set up . The methodology developed in this paper will be very useful to the researchers, engineers, and statisticians for the further advancement in this area. It motivated the researchers to use PS as an alternative to UL to get more efficient estimators. For future, same methodology has been considered by us for some other lifetime models and work is under progress for Generalized inverted exponential distribution (GIED) and Flexible Weibull distribution.

Table 2: Coverage probabilities and corresponding confidence intervals/HPD intervals of the estimators of θ for fixed values of sample size $n = 20$ with different values of θ i.e $\theta = 0.5, 1, 2, 3$ with different prior scheme.

theta	prior scheme	mle	mps	Bmle	Bps
0.5	1	92.9(0.3086,0.7380)	88.3(0.2566,0.6674)	94.7(0.3059,0.7559)	91.4(0.3037,0.5546)
	2	96.4(0.2632,0.7906)	88.5(0.2609,0.6702)	95.4(0.3092,0.7561)	100(0.3564,0.6649)
	3	95.0(0.2890,0.7632)	88.6(0.2597,0.6694)	96.4(0.3075,0.7596)	100(0.3266,0.6065)
	4	95.7(0.2528,0.7958)	87.4(0.2591,0.6674)	94.0(0.3063,0.7573)	100(0.3488,0.6571)
1	1	95.8(0.5097,1.5807)	84.6(0.5188,1.3284)	94.7(0.6109,1.5106)	100(0.6858,1.2876)
	2	81.9(0.7410,1.3579)	86.6(0.5207,1.3334)	95.7(0.6261,1.4802)	85.0(0.5149,1.1595)
	3	91.3(0.6181,1.5202)	87.9(0.5307,1.3590)	94.8(0.6257,1.5421)	99.6(0.6382,1.1705)
	4	92.4(0.6001,1.5195)	87.5(0.5256,1.3458)	94.0(0.6191,1.5305)	100(0.6423,1.1877)
2	1	94.2(1.1884,3.0576)	92.0 (1.0535,2.6974)	94.5(1.2414,3.0681)	100 (1.2975,2.4047)
	2	90.6(1.2798,2.9449)	87.7(1.0481,2.6836)	97.4(1.3254,2.8264)	100(1.5171,2.3704)
	3	90.3(1.3316,2.8558)	88.5(1.0355,2.6636)	95.9(1.2298,3.0054)	75.7(1.1702,2.0832)
	4	91.0(1.2958,2.8866)	87.1(1.0377,2.6570)	96.1(1.2203,3.0170)	95.2(1.1639,2.1047)
3	1	89.5(1.9618,4.3600)	87.2(1.5687,4.0167)	95.0(1.8444,4.5646)	78.8(1.7682,3.2064)
	2	95.3(1.6492,4.6123)	86.9(1.5539,3.9788)	98.7(2.0994,4.0040)	100(2.8159,4.1089)
	3	95.1(1.6234,4.6323)	87.5(1.5521,3.9741)	96.1(1.8509,4.45839)	100(2.1087,3.8440)
	4	92.0(1.8405,4.5046)	87.5(1.5744,4.0312)	94.5(1.8555,4.5773)	95.7(1.8557,3.3923)

Table 3: Average estimates (in the first row of each cell) under SELF using informative prior and corresponding Risks (in brackets) of θ for fixed values of sample size $n = 20$ for different values of θ i.e $\theta = 0.5, 1, 2, 3$ with different prior scheme and prior variance.

	prior scheme	mle	mps	Bmle	Bps
0.5	1(00)	0.5233(0.0146)	0.4620(0.0125)	0.5232(0.0144)	0.4286(0.0055)
	2(var=.5)	0.5269(0.0157)	0.4656(0.0129)	0.5255(0.0146)	0.5238(0.0012)
	3(var=8)	0.5261(0.01395)	0.4645(0.0116)	0.5260(0.0138)	0.4737(0.0032)
	4(var=100)	0.5243(0.0158)	0.4632(0.0132)	0.5243(0.0158)	0.5202(0.0059)
1	1(00)	1.0452(0.0669)	0.9236(0.0569)	1.0452(0.0669)	1.0178(0.0174)
	2(var=.5)	1.0494(0.0620)	0.9271(0.0523)	1.0399(0.0486)	0.6557(0.0105)
	3(var=8)	1.0692(0.0746)	0.9449(0.0577)	1.0679(0.0732)	0.9080(0.0106)
	4(var=100)	1.0598(0.0697)	0.9357(0.0559)	1.0598(0.0696)	0.9249(0.0128)
2	1(00)	2.1230(0.2486)	1.8754(0.1986)	2.1129(0.2253)	1.8746(0.2431)
	2(var=.5)	2.1124(0.2609)	1.8658(0.2125)	2.0558(0.2145)	1.8269(0.0167)
	3(var=8)	2.0937(0.2288)	1.8491(0.1945)	2.0887(0.2145)	1.5952(0.0238)
	4(var=100)	2.0912(0.2443)	1.8473(0.2093)	2.0908(0.2429)	1.6183(0.1512)
3	1(00)	3.1609(0.6045)	2.7927(0.4967)	3.1609(0.6045)	2.4677(0.2972)
	2(var=.5)	3.1308(0.5278)	2.7664(0.4566)	3.0292(0.4378)	3.2104(0.0243)
	3(var=8)	3.1279(0.5439)	2.7631(0.4678)	3.1121(0.4695)	2.9965(0.0448)
	4(var=100)	3.1725(0.6273)	2.8028(0.5094)	3.1708(0.6193)	2.6252(0.1581)

References

[1] Cheng, R. C. H. and Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. Journal of the Royal Statistical Society B **45**, 394-403.

[2] Ranneby, B., (1984). The Maximum Spacings Method. An Estimation Method Related to the Maximum Likelihood Method. Scand. J. Stat., **11**, 93-112.

[3] Shah, A. and Gokhale, D. V., (1993). On Maximum Product of Spacings Estimation for Burr XII Distributions. Commun. Stat. Simulat. Computat., **22**, 615-641.

[4] Harter, H. L. and Moore, A. H. (1965). Maximum likelihood estimation of the parameters of Gamma and Weibull populations from complete and from censored samples.

[5] Ghosh, S.R. Jammalamadaka (2001) A general estimation method using spacings. Journal of Statistical Planning and Inference **93**.

[6] Shao, Y. (2001). Consistency of the maximum product of spacings method and estimation of a unimodal distribution. Statistica Sinica **11**, 1125-1140.

[7] Wong, T. S. T. and Li, W. K. (2006). A note on the estimation of extreme value distributions using maximum product of spacings. IMS Lecture Notes Monograph Series **52**, 272-283.

[8] Mezbahur Rahman and Larry M. Pearson (2002): Estimation in two-parameter Exponential distributions. Journal of Stat. Computat. Simulat., **70**, 371-386.
[9] Mezbahur Rahman and Larry M. Pearson (2003): A note on estimating parameters in two-parameter Pareto distributions, International Journal of Mathematical Education in Science and Technology, **34**, 298-306.
[10] Mezbahur Rahman, Larry M. Pearson and Uros R Martinovic (2007): Method of product of spacings in the two parameter gamma distribution, Journal of Statistical Research Bangladesh **41**, 51-58.
[11] Titterington, D.M. (1985) Comment on Estimating parameters in continuous univariate distributions". Journal of the Royal Statistical Society Series B **47**, 115 .
[12] Cheng, R.C.H.; Amin, N.A.K. 1979: Maximum product-of-spacings estimation with applications to the lognormal distribution, University of Wales IST, Math Report79-1.
[13] FC and M.J. Newby (1991). A note on the use of the product of spacings in Bayesian inference. Kwantitatieve Methoden **37**, 19-32.
[14] Anatolyev, Stanislav and Grigory Kosenok (2005) ?An alternative to maximum likelihood based on spacings, Econometric Theory, **21**, 472?-476.
[15] Nan Zhang, Yongcheng Qi and Kang James (2010) "Comparision of study of estimation methods of non-regular distribution"
[16] Shao, Y. and Nahn, M.G. (1999) "Maximum product spacing method: A unified formulation with illustration of strong consistency", Illiinois journal of mathematics **43**, 190.
[17] H. Krishna and Kapil. Sharma (2013), Reliability estimation in Generalized inverted exponential distribution with progressively type-II censored data, J. Statist. Comput. Simul. **83**.
[18] S. K. Singh, U. Singh, and V. K. Sharma (2013), "Bayesian estimation and prediction for Flexible Weibull model under Type-II Censoring Scheme", Journal of Probability and Statistics. **2013**, Article ID 146140,
[19] J. O. Berger, Statistical Decision Theory and Bayesian Analysis, Springer Series in Statistics, Springer, New York, NY, USA, 2nd edition, (1985).
[20] S. K. Singh, U. Singh, and D. Kumar (2012), ?Bayes estimators of the reliability function and parameter of inverted exponential distribution using informative and non-informative priors,? Journal of Statistical Computation and Simulation, **1**, article 12.
[21] Huzurbazar, V. S. (1948) The likelihood equation, consistency and the maxima of the likelihood function. Ann. Eugenics, **14**, 185-200.
[22] U.Singh, S.K.Singh and R.K.Singh (2014),"Comparative study of traditional estimation method and maximum product spacing method in Generalised inverted exponential Distribution", J. Stat. Appl. Pro. **3**, 1-17.