Pakistan Journal of Statistics and Operation Research

Bayesian Estimation in Rayleigh Distribution under a Distance Type Loss Function

Babulal Seal[1], Proloy Banerjee[2]*, Shreya Bhunia[3], Sanjoy Kumar Ghosh[4]

*Corresponding author

1. Department of Mathematics and Statistics, Aliah University, Newtown, Kolkata, India, babulal_seal@yahoo.com
2. Department of Mathematics and Statistics, Aliah University, Newtown, Kolkata, India, proloy.stat@gmail.com
3. Department of Mathematics and Statistics, Aliah University, Newtown, Kolkata, India, shreyabhunia.stat@gmail.com
4. Department of Statistics, Vidyasagar Metropolitan College, Kolkata, India, sanjoykghosh2@gmail.com

Abstract

Estimation of unknown parameters using different loss functions encompasses a major area in the decision theory. Specifically, distance loss functions are preferable as it measures the discrepancies between two probability density functions from the same family indexed by different parameters. In this article, Hellinger distance loss function is considered for scale parameter λ of two-parameter Rayleigh distribution. After simplifications, form of loss is obtained and that is meaningful if parameter is not large and Bayes estimate of λ is calculated under that loss function. So, the Bayes estimate may be termed as 'Pseudo Bayes estimate' with respect to the actual Hellinger distance loss function as it is obtained using approximations to actual loss. To compare the performance of the estimator under these loss functions, we also consider weighted squared error loss function (WSELF) which is usually used for the estimation of the scale parameter. An extensive simulation process is carried out to study the behaviour of the Bayes estimators under the three different loss functions, i.e. simplified, actual and WSE loss functions. From the numerical results it is found that the estimators perform well under the Hellinger distance loss function in comparison with the traditionally used WSELF. Also, we demonstrate the methodology by analyzing two real-life datasets.

Key Words: Two-parameter Rayleigh Distribution; Hellinger Divergence Measure; Risk Function.

Mathematical Subject Classification: 62C10, 62C05, 91G70.

1. Introduction

Estimation of model parameters is one of the fundamental steps in statistical modeling. In life testing experiments the situations when censoring occurs, i.e. the values of measurements or observations are partially known, the Rayleigh distribution is useful in handling such type of data. Since its inception in statistics literature by Rayleigh (1880) relating with the problems in the field of acoustics several extensive works have been done on the estimation of one-parameter Rayleigh distribution, e.g. Dey and Das (2007), Howlader and Hossain (1995), Mkolesia et al. (2016) etc. Inferential issues and problems relating to estimation from data which follow Rayleigh distribution are discussed by several authors. Some of them are Akhter and Hirai (2009), Dey (2009), Ahmed et al. (2013), Ahmed et al. (2015), Tahmasebi et al. (2017), Banerjee and Bhunia (2022) and many others.

Although we find extensive works on one-parameter Rayleigh distribution, Dey et al. (2016) mentioned less attention was given to two-parameter Rayleigh distribution. A brief description of this distribution is given in the book by Johnson et al. (1994). For two-parameter Rayleigh distribution, different estimation techniques were proposed by Dey et al. (2014). Based on progressive type-II censored samples using NR method and EM algorithm, Fundi et al. (2017) obtained MLE for both the parameters of this distribution. In case of upper record values, Seo et al. (2016) provided the exact confidence intervals for unknown parameters and exact predictive intervals. On the basis of a doubly censored sample from a two-parameter Rayleigh lifetime model, inference about future responses is done by Khan et al. (2010).

Here in our work, we have considered two-parameter Rayleigh distribution and our aim is to estimate the scale parameter by using Hellinger distance measure. By using this distance measure, loss is directly obtained from the discrepancies between two densities instead of the difference between parameter and its estimate. A new goodness-of-fit test for Rayleigh distribution based on Hellinger distance and its related properties are discussed by Jahanshahi et al. (2016). The probability density function and the corresponding cumulative density function are in the following.

$$f(x;\lambda,\mu) = 2\lambda(x-\mu)e^{-\lambda(x-\mu)^2}, \;\; x>\mu, \tag{1}$$
$$F(x;\lambda,\mu) = 1-e^{-\lambda(x-\mu)^2},$$

where, $\lambda > 0$ and $0 < \mu < \infty$ are respectively the scale parameter and location parameter.

It has also been observed that the two-parameter Rayleigh model is the special case of three parameter Weibull distribution. Presence of the location parameter in the two-parameter model gives much more flexibility to analyze the real-life data than the one-parameter Rayleigh distribution.

In this paper our focus is to derive a loss function using Hellinger distance measure and use this loss to find the Bayes estimate for the scale parameter λ. As it is difficult to obtain the risk directly by using the derived loss, here we propose a intermediate way to find the Bayes estimate by considering some simplification in actual loss function. Also this is explained in a following section for general λ. For comparison, we derive the Bayes estimate and the integrated risk of λ under weighted squared error loss function. The performance of the Bayes estimate under these different loss functions are studied numerically through the integrated risk calculation.

The article is organized as follows. Section 2 provides the posterior density for scale parameter by using the conjugate prior. In Section 3, loss function for the scale parameter is derived by considering the two-parameter Rayleigh distribution and Hellinger distance. Section 4 is dedicated to find the risks and Bayes estimates under derived Helinger distance loss and weighted squared error loss function respectively. Further, all the methods are studied numerically in Section 5. In Section 6, two datasets have been used to demonstrate the performance of the estimators in real situation. Finally, a conclusion of the study is provided in Section 7.

2. Posterior Density under Conjugate Prior Distribution

The selection of prior distribution is essential for the parameter estimation in Bayesian setup. After selection of the prior, the posterior distribution summarizes the probabilistic information available for the parameter of interest by combining two components, i.e. prior probability and the likelihood function for the observed data drawn from the statistical model.

Suppose, $x_1, x_2, \cdots, x_n$ is a random sample from two-parameter Rayleigh distribution (1). The likelihood function for the given sample is,

$$L(\lambda|\boldsymbol{x}) = (2\lambda)^n \prod_{i=1}^{n} (x_i-\mu)\, e^{-\lambda\sum_{i=1}^{n}(x_i-\mu)^2}. \tag{2}$$

It is postulated that, if μ is known, then λ has a conjugate gamma prior. But if both the parameters are unknown, then the joint conjugate prior does not exist (Dey et al., 2016). So, it is reasonable to assume conjugate gamma prior for the scale parameter λ as it is very flexible and can also be converted to non-informative prior. It is expected that the conjugate prior minimizes the effect of the data on the posterior distribution depending on its hyperparameters. The prior density for λ is given by

$$\Pi(\lambda) = \frac{p^{\alpha}}{\Gamma\alpha} e^{-p\lambda}\lambda^{\alpha-1}, \;\; \alpha > 0, \text{ and } p > 0, \tag{3}$$

where α and p are the hyperparameters and both being known.

Now, using likelihood function (2) and the prior density (3), the posterior distribution of λ for the given sample $x_1, x_2, \cdots, x_n$ is

$$\begin{aligned}\Pi(\lambda|\boldsymbol{x}) &= \frac{f(\boldsymbol{x}|\lambda)\Pi(\lambda)}{\int_0^{\infty} f(\boldsymbol{x}|\lambda)\Pi(\lambda)d\lambda} \\ &= \frac{\lambda^n e^{-\lambda\sum_{i=1}^{n}(x_i-\mu)^2} e^{-p\lambda}\lambda^{\alpha-1}}{\int_0^{\infty} \lambda^n e^{-\lambda\sum_{i=1}^{n}(x_i-\mu)^2} e^{-p\lambda}\lambda^{\alpha-1}\, d\lambda} \\ &= \frac{\left\{\sum_{i=1}^{n}(x_i-\mu)^2 + p\right\}^{n+\alpha}}{\Gamma(n+\alpha)} e^{-\lambda\left\{\sum_{i=1}^{n}(x_i-\mu)^2+p\right\}} \lambda^{(n+\alpha)-1}. \end{aligned} \tag{4}$$

$$\text{Therefore,} \quad \lambda|\boldsymbol{x} \sim \text{Gamma}(n+\alpha, \sum_{i=1}^{n}(x_i-\mu)^2 + p).$$

So, the posterior distribution is also a gamma distribution with parameters $(n+\alpha)$ and $\sum_{i=1}^{n}(x_i-\mu)^2 + p$.

3. Derivation of Hellinger Distance Loss for Scale Parameter

In Bayesian inference, choice of an appropriate loss function is an integral part and there are various types of loss functions in the literature which have been used for many years. A loss function must be specified in order to select the 'best' estimator from a decision-theoretic context and it is used to indicate a penalty associated with each of the alternative estimates. In statistics, the Hellinger distance loss function is used to measure the discrepancies between two probability distributions. Let, $f(x|a)$ and $f(x|\theta)$ be two probability densities over same measurable space Ω. A new class of priors based on Hellinger information is introduced by Shemyakin (2014). In non-regular situations, i.e. when Fisher information does not exist, Hellinger distance can also be utilized to investigate the information aspects of the parametric set as suggested by Birgé (1985) and Le Cam (2012). According to Zamanzade and Mahdizadeh (2017), Hellinger distance between $f(x|a)$ from $f(x|\theta)$ is denoted by $D_H(f(x|a), f(x|\theta))$, where

$$D_H(f(x|a), f(x|\theta)) = \frac{1}{2}\int f_{\theta}(x)\left(\sqrt{\frac{f(x|a)}{f(x|\theta)}} - 1\right)^2 dx \;\; = \frac{1}{2}E_{\theta}\left(\sqrt{\frac{f(x|a)}{f(x|\theta)}} - 1\right)^2.$$

Now, we consider two continuous density functions from the two-parameter Rayleigh distribution, denoted as $f(x|\lambda_1)$ and $f(x|\lambda_2)$ and derive the Hellinger distance between these as follows:

$$\begin{aligned} D_H(f(.|\lambda_1), f(.|\lambda_2)) &= \frac{1}{2}E_{\lambda_1}\left[\sqrt{\frac{f(x|\lambda_2)}{f(x|\lambda_1)}} - 1\right]^2 \\ &= \frac{1}{2}E_{\lambda_1}\left[\sqrt{\frac{2\lambda_2(x-\mu)e^{-\lambda_2(x-\mu)^2}}{2\lambda_1(x-\mu)e^{-\lambda_1(x-\mu)^2}}} - 1\right]^2 \\ &= \frac{1}{2}E_{\lambda_1}\left[\frac{\lambda_2}{\lambda_1}e^{(\lambda_1-\lambda_2)(x-\mu)^2} - 2\left(\sqrt{\frac{\lambda_2}{\lambda_1}}\right)e^{\left(\frac{\lambda_1-\lambda_2}{2}\right)(x-\mu)^2} + 1\right] \\ &= \frac{1}{2}\left[\frac{\lambda_2}{\lambda_1}\left\{E_{\lambda_1}e^{(\lambda_1-\lambda_2)(x-\mu)^2}\right\} - 2\left(\sqrt{\frac{\lambda_2}{\lambda_1}}\right)E_{\lambda_1}\left\{e^{\left(\frac{\lambda_1-\lambda_2}{2}\right)(x-\mu)^2}\right\} + 1\right] \\ &= \frac{1}{2}\left[\left(\frac{\lambda_2}{\lambda_1}\times\frac{\lambda_1}{\lambda_2}\right) - 2\sqrt{\frac{\lambda_2}{\lambda_1}}\times\frac{2\lambda_1}{\lambda_1+\lambda_2} + 1\right] \end{aligned}$$

$$= \frac{1}{2}\left[2 - 4\frac{\sqrt{\lambda_1}\sqrt{\lambda_2}}{\lambda_1+\lambda_2}\right]$$
$$= 1 - \frac{2\sqrt{\lambda_1}\sqrt{\lambda_2}}{\lambda_1+\lambda_2}.$$

Thus, we define loss function as,

$$L\left(\delta(x),\lambda\right) = 1 - \frac{2\sqrt{\delta(x)}\sqrt{\lambda}}{\delta(x)+\lambda}. \tag{5}$$

The loss defined in (5), is the loss function for the scale parameter under Hellinger distance measure between two density functions of the two-parameter Rayleigh distribution.

Now, as the loss function contains the parameter of interest λ and $\delta(x)$ in the denominator, it will be complicated while integrating the loss with respect to λ to attain the risk. Therefore, a modification has been done by ignoring λ from the denominator and it is renamed as simplified loss function. This is reasonable as $1 - \frac{\sqrt{\delta(x)}\sqrt{\lambda}}{\delta(x)+\lambda}$ is more than the actual loss $L_a(\delta(x),\lambda)$. Moreover, $\sqrt{\frac{\delta(x)}{\lambda}}$ should be close to one when λ from denominator is ignored and thus logically $L_s(\delta(x),\lambda)$ will not be too bad. So, from Equation (5), we are defining two loss functions such as actual loss function (ALF) in (6) and simplified loss function (SLF) in (7) which is used to find the risk for the scale parameter.

$$L_{\mathsf{a}}\left(\delta(x),\lambda\right) = 1 - \frac{2\sqrt{\delta(x)}\sqrt{\lambda}}{\delta(x)+\lambda}, \tag{6}$$

$$L_{\mathsf{s}}\left(\delta(x),\lambda\right) = 1 - \frac{\sqrt{\delta(x)}\sqrt{\lambda}}{\delta(x)}. \tag{7}$$

Though the simplified loss is slightly different from actual loss, it will be preferred into consideration which is discussed in numerical section. Here our purpose is to find the risk in terms of λ under the actual loss function. But due to the complexity in calculation, we use the following steps to achieve the target which is the basic idea behind our methodology.

1. Obtain the posterior risk under SLF, $L_{\mathsf{s}}\left(\delta(x),\lambda\right)$.
2. Compute the simplified Bayes estimate $\delta_s(x)$ under SLF.
3. Substitute the obtained $\delta_s(x)$ in ALF, i.e. $L_{\mathsf{a}}\left(\delta(x),\lambda\right)$.
4. Find the risk under ALF by using Step 3 and then compare previous step with Step in 1.

4. Posterior Risk and Bayes Estimate under Distance Type Loss Function and Weighted Squared Error Loss Function

4.1. Under Hellinger Type Distance Loss Function

Bayes action is a decision rule that minimizes the posterior expected value of a loss function. The quality of an estimator is measured by the expected loss, known as risk of that estimator. In this section, our focus is on the derivation of the Bayes estimate and the posterior risk of the scale parameter under the simplified loss function. So, the posterior risk of the scale parameter λ under SLF is derived in the following.

$$E_{\lambda|\tilde{x}}\left(L_{\mathsf{s}}\left(\delta(x),\lambda\right)\right) = \int_0^\infty L_{\mathsf{s}}(\delta(x),\lambda)\,\pi(\lambda|\tilde{x})\,d\lambda$$
$$= \int_0^\infty \left\{1 - \frac{\sqrt{\delta(x)}\sqrt{\lambda}}{\delta(x)}\right\} \frac{\left\{\sum_{i=1}^n (x_i-\mu)^2 + p\right\}^{n+\alpha}}{\Gamma(n+\alpha)} e^{-\lambda\left\{\sum_{i=1}^n (x_i-\mu)^2+p\right\}} \lambda^{n+\alpha-1}\,d\lambda$$
$$= 1 - \frac{\sqrt{\delta(x)}}{\delta(x)} \frac{\left(\sum_{i=1}^n (x_i-\mu)^2 + p\right)^{n+\alpha}}{\Gamma(n+\alpha)} \int_0^\infty e^{-\left\{\sum_{i=1}^n (x_i-\mu)^2+p\right\}\lambda} \lambda^{(n+\alpha+\frac{1}{2})-1}\,d\lambda$$

$$
\begin{aligned}
=&1-\frac{\sqrt{\delta(x)}}{\delta(x)}\ \frac{\left(\sum_{i=1}^{n}(x_i-\mu)^2+p\right)^{n+\alpha}}{\Gamma(n+\alpha)}\ \frac{\Gamma(n+\alpha+\frac{1}{2})}{(\sum_{i=1}^{n}(x_i-\mu)^2+p)^{n+\alpha+\frac{1}{2}}}\\
=&1-\frac{\sqrt{\delta(x)}}{\delta(x)}\ \frac{\Gamma(n+\alpha+\frac{1}{2})}{\Gamma(n+\alpha)}\ \frac{1}{(\sum_{i=1}^{n}(x_i-\mu)^2+p)^{\frac{1}{2}}}\\
=&1-\frac{\Gamma(n+\alpha+\frac{1}{2})}{\delta(x)\,\Gamma(n+\alpha)\sqrt{\sum_{i=1}^{n}(x_i-\mu)^2+p}}.
\end{aligned}
\tag{8}
$$

Now, we find the simplified Bayes estimate $\delta_{\mathsf{s}}(x)$, for which the posterior risk will be minimum. The Equation (8) attains least value of risk, when

$$
\delta_{\mathsf{s}}(x)=\left\{\frac{2\,\Gamma(n+\alpha+\frac{1}{2})}{\Gamma(n+\alpha)\sqrt{\sum_{i=1}^{n}(x_i-\mu)^2+p}}\right\}^2. \tag{9}
$$

Therefore, the Bayes estimate $\delta_{\mathsf{s}}(x)$ is obtained under the SLF which is simplification of the actual loss function. That is why this simplified Bayes estimate may be termed as 'Pseudo Bayes estimate' (PBE) with respect to the actual Hellinger distance loss function. According to the steps mentioned in the previous section, we substitute that $\delta_{\mathsf{s}}(x)$ into the actual loss function (6) and then we compute the risk of the scale parameter λ under ALF by carrying out a simulation study in the numerical section.

4.2. Under Weighted Squared Error Loss Function

In case of parameter estimation, several loss functions are used in the Bayesian inference. The most frequently used loss function is the squared error loss function (SELF) in the field of decision theory as it is symmetric and penalizes both over and under estimation of the same magnitude (Kamińska and Porosiński, 2009). But when our parameter of interest is the scale parameter, then a modification has been done in the SELF and it is defined as,

$$
L(\lambda,\delta(x))=\left(\frac{\delta(x)}{\lambda}-1\right)^2, \tag{10}
$$

where, $\delta(x)$ is the estimate of λ. Generalization on the squared error loss is termed as the weighted squared error loss function (WSELF) (Ferguson, 1967). So, the WSELF for the scale parameter λ is,

$$
L(\lambda,\delta(x))=w(\lambda)(\lambda-\delta(x))^2,\ \ w(\lambda)>0.
$$

Here, in particular we consider the following loss function

$$
L(\lambda,\delta(x))=(\delta(x)-\lambda)^2\ \frac{1}{\lambda^2};\ \text{ with } w(\lambda)=\frac{1}{\lambda^2}.
$$

As WSELF is usually used for the estimation of scale parameter, we study the performance of the estimator under this loss function also. The Bayes estimate under weighted squared error loss function becomes,

$$
\begin{aligned}
\delta_w(x)&=\frac{E\left\{\lambda\, w(\lambda)|X=x\right\}}{E\left\{w(\lambda)|X=x\right\}}\\
&=\frac{\int_0^\infty e^{-\left\{\sum_{i=1}^{n}(x_i-\mu)^2+p\right\}\lambda}\lambda^{(n+\alpha-1)-1}\,d\lambda}{\int_0^\infty e^{-\left\{\sum_{i=1}^{n}(x_i-\mu)^2+p\right\}\lambda}\lambda^{(n+\alpha-2)-1}\,d\lambda}\\
&=\frac{\Gamma(n+\alpha-1)}{\Gamma(n+\alpha-2)}\times\frac{\left\{\sum_{i=1}^{n}(x_i-\mu)^2+p\right\}^{n+\alpha-2}}{\left\{\sum_{i=1}^{n}(x_i-\mu)^2+p\right\}^{n+\alpha-1}}\\
&=\frac{(n+\alpha-2)\Gamma(n+\alpha-2)}{\Gamma(n+\alpha-2)}\left\{\sum_{i=1}^{n}(x_i-\mu)^2+p\right\}^{-1}
\end{aligned}
$$

$$= \frac{n+\alpha-2}{\{\sum_{i=1}^{n}(x_i-\mu)^2+p\}}. \tag{11}$$

Now, the integrated risk of the Bayes estimator is as follows,

$$\begin{aligned} E_\lambda[R(\lambda,\delta(x))] &= E_\lambda E_{x|\lambda}\left[L(\lambda,\delta(x))\right] \\ &= \int_0^\infty \int_\mu^\infty L(\lambda,\delta(x)) f(x|\lambda)\pi(\lambda) d\lambda dx \\ &= \frac{2p^\alpha}{\Gamma\alpha}\int_0^\infty \int_\mu^\infty \left(\frac{\delta(x)}{\lambda}-1\right)^2 \lambda^\alpha \left(x-\mu\right) e^{-\left\{(x-\mu)^2+p\right\}\lambda} d\lambda dx. \end{aligned}$$

Due to the analytical complexity, risk under WSELF is studied numerically in next section.

5. Simulation Studies for the Performance of the Estimators

In this section, we have implemented our methodology with the help of an extensive simulation process to study the behavior of the Bayes estimators under the different loss functions. The complete structure of the numerical work is discussed below.

We assume the location parameter μ is known and we fix it at w.l.g $\mu = 1$. It has been observed that there is no effect of changing the location parameter on the Bayes estimate as well as the risk. This happens due to the value of the random sample x_i which is generated from two-parameter Rayleigh distribution in such a way that the difference $x_i - \mu$ is same for any value of μ. Hence the different values of μ result the same Bayes estimate of λ as it contains $(x_i - \mu)$ in (9) and (11).

Also, we need to keep in mind that to derive the simplified loss function, we have omitted λ from the denominator of the actual loss function. Therefore that λ should be small so that it does not make any large changes between both the loss functions. Now, to generate small values of λ we choose some combinations of hyperparameters (α, p) such as (0.50, 5.00), (0.25, 1.00), (0.75, 3.00), (1.00, 5.00), (3.50, 0.75), (5.00, 3.00), (0.10, 0.25). For every combinations of (α, p), λ of sizes *100* are simulated to study the integrated risk of $\hat{\lambda}$ under the different loss functions.

Subsequently, for fixed μ and previously generated λ, we simulate random samples of sizes $n = 10, 25, 50, 75, 100, 150$ from two-parameter Rayleigh distribution by using Inverse-transformation method. We carried out the Monte Carlo simulation process for $K = 1000$ times and based on that simulated random samples, we calculate the simplified Bayes estimate mentioned in (9) and substitute it to the actual loss (6) and simplified loss (7) respectively. So, we obtain both the loss values considering the same Bayes estimate and take an average of that to calculate the integrated risk at $\hat{\lambda}$ under both the loss functions separately. Also, we consider the weighted squared error loss function to calculate the integrated risk of $\hat{\lambda}$ by substituting the Bayes estimate (11) into the WSELF (10) in similar manner. The structure of the three loss functions are shown by the histograms in Figure 2. The following algorithm is used to obtain the integrated risk of $\hat{\lambda}$.

Step 1: Choose some combinations of hyperparameters (α, p).

Step 2: Fix (α_1, p_1) and generate λ_1, λ_2,..., λ_m from $gamma(\alpha_1, p_1)$.

Step 3: Set $\mu = 1$ and generate $\mathbf{X}$ of size n from $Rayleigh(\mu, \lambda_1)$.

Step 4: Repeat **Step 3** for *K* times to generate random samples $\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_K}$.

Step 5: Calculate the Bayes estimates $\delta_s(x)$ and $\delta_w(x)$ for K times from the Equation (9) and (11) respectively.

Step 6: Substitute $\delta_s(x)$ in Equation (6) and (7) to calculate actual loss values $L_a\left(\delta_s(x), \lambda_1\right)$ K times and also in simplified loss values $L_s\left(\delta_s(x), \lambda_1\right)$.

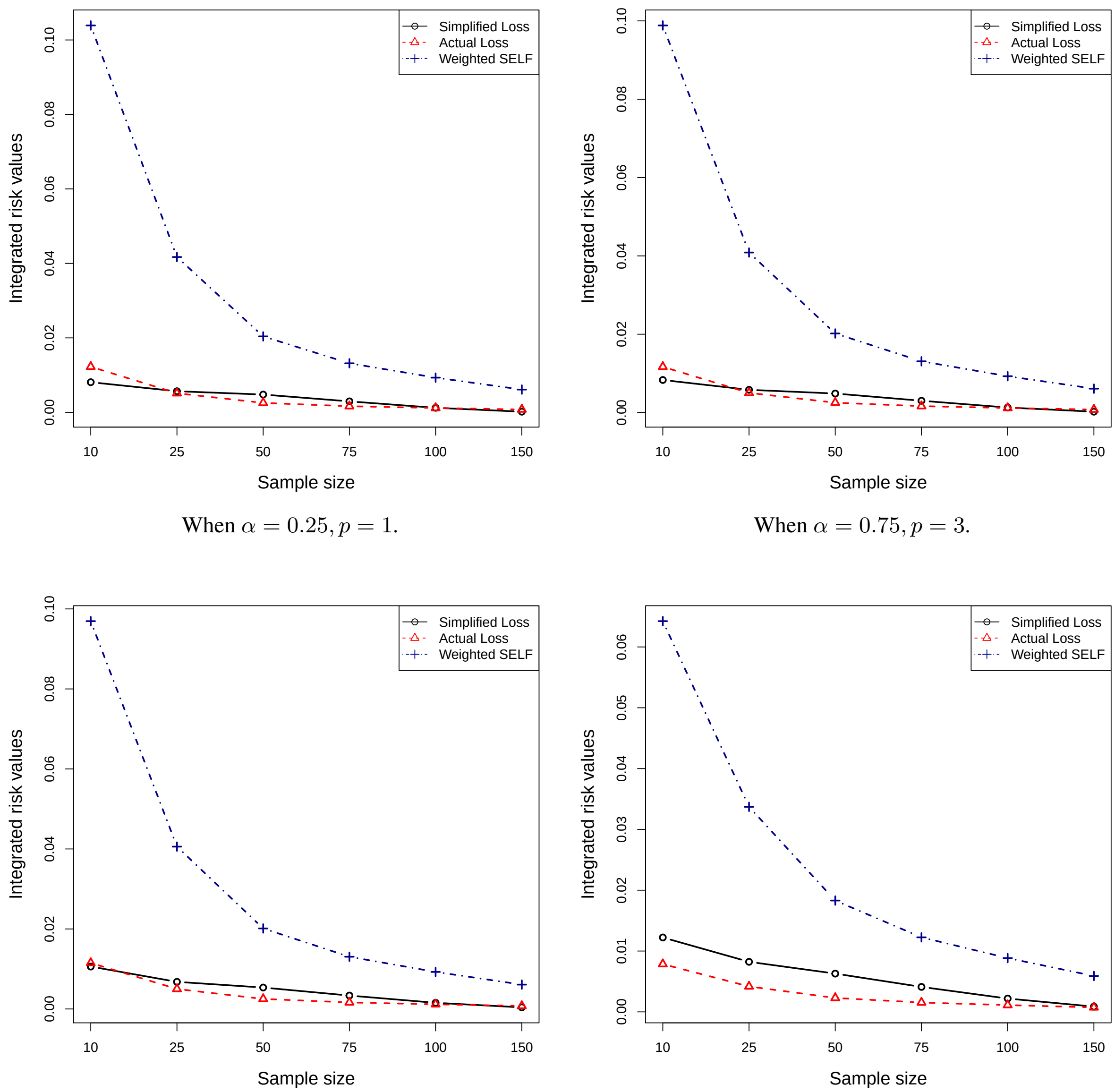
Step 7: Substitute $\delta_w(x)$ in (10) to calculate *K* times weighted squared error loss values $L\left(\delta_w(x), \lambda_1\right)$.

Step 8: Take an average over the K replicated loss values $E_{X|\lambda_1} L\left(\delta(x), \lambda_1\right)$ under the three different loss functions to obtain the risk values at λ_1.

Step 9: Repeat **Step 3** - **Step 8** to compute the risks for remaining $\lambda_i' s;\ i = 2, 3, \cdots, m$.

Step 10: Finally, take an average over all the m replicated risk values of λ_i i.e. $E_\lambda E_{X|\lambda} L\left(\delta(x), \lambda\right)$ to obtain the integrated risk of $\hat{\lambda}$ under three loss functions.

Step 11: Repeat **Step 2** - **Step 9** to calculate the integrated risk of $\hat{\lambda}$ for remaining combinations of (α, p).



When $\alpha = 0.25, p = 1.$ When $\alpha = 0.75, p = 3.$

When $\alpha = 1.0, p = 5.0.$ When $\alpha = 5.0, p = 3.0.$

Figure 1: Risk values for different combinations of hyperparameters (α, p).

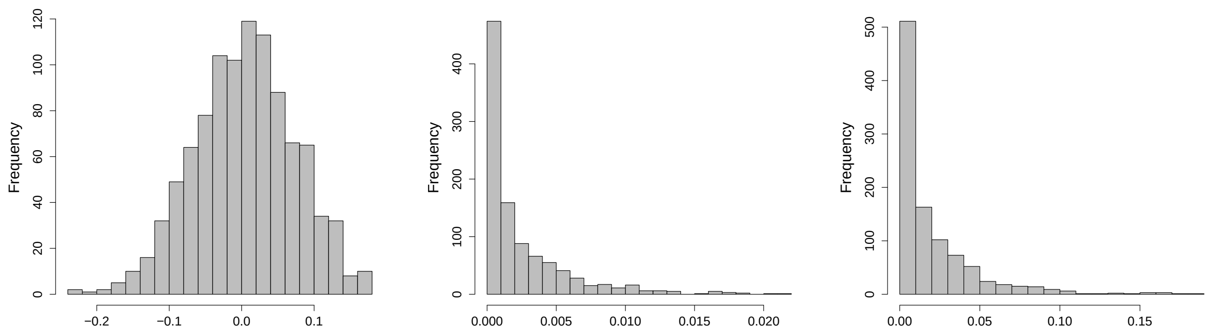
Numerical outcomes of the above algorithm are presented in Table 1. It provides the average Bayes estimate and the corresponding integrated risk values of $\hat{\lambda}$ under the three different loss functions. The following observations can be listed from the Table 1.

Table 1: Average Bayes estimates (BE) and Integrated risk (IR) of the scale parameter λ.

Parameter choices	Sample sizes	BE under SLF	IR under SLF	IR under ALF	Absolute difference	BE under WSELF	IR under WSELF
(0.50, 5.00)	10	0.8949944	0.0087930	0.0119920	0.0031990	0.7419701	0.1015763
	25	0.9034320	0.0059734	0.0050845	0.0008889	0.8407767	0.0413315
	50	0.9048837	0.0049359	0.0025270	0.0024089	0.8733595	0.0203037
	75	0.9024652	0.0030574	0.0016413	0.0014161	0.8814728	0.0131329
	100	0.8996372	0.0013113	0.0011687	0.0001426	0.8839301	0.0093009
	150	0.8983384	0.0002393	0.0007615	0.0005222	0.8878740	0.0061098
(0.25, 1.00)	10	0.2170043	0.0080962	0.0122479	0.0041517	0.1789727	0.1038946
	25	0.2189387	0.0056705	0.0051300	0.0005405	0.2036028	0.0417114
	50	0.2192221	0.0047788	0.0025379	0.0022410	0.2115466	0.0203939
	75	0.2186095	0.0029465	0.0016459	0.0013006	0.2135074	0.0131714
	100	0.2179107	0.0012237	0.0011711	0.0000526	0.2140966	0.0093207
	150	0.2175846	0.0001792	0.0007626	0.0005834	0.2150459	0.0061189
(0.75, 3.00)	10	0.2340238	0.0083166	0.0116777	0.0033611	0.1949647	0.0988354
	25	0.2360836	0.0058068	0.0050265	0.0007802	0.2198713	0.0408618
	50	0.2364325	0.0048584	0.0025123	0.0023460	0.2282365	0.0201873
	75	0.2357928	0.0030123	0.0016351	0.0013772	0.2303261	0.0130837
	100	0.2350501	0.0012820	0.0011656	0.0001164	0.2309565	0.0092767
	150	0.2347047	0.0002216	0.0007602	0.0005386	0.2319753	0.0060992
(1.00, 5.00)	10	0.1805371	0.0106018	0.0115066	0.0009048	0.1511065	0.0969443
	25	0.1815707	0.0067804	0.0049977	0.0017828	0.1692230	0.0405855
	50	0.1816869	0.0053568	0.0025062	0.0028506	0.1754197	0.0201285
	75	0.1811506	0.0033518	0.0016325	0.0017192	0.1769647	0.0130581
	100	0.1805589	0.0015412	0.0011641	0.0003770	0.1774221	0.0092617
	150	0.1802714	0.0003964	0.0007595	0.0003631	0.1781785	0.0060916
(3.50, 5.00)	10	0.6633656	0.0143058	0.0089814	0.0053244	0.5756489	0.0739874
	25	0.6593862	0.0089594	0.0044709	0.0044885	0.6185151	0.0361098
	50	0.6564393	0.0066112	0.0023692	0.0042420	0.6348592	0.0189857
	75	0.6532789	0.0042764	0.0015728	0.0027036	0.6386656	0.0125604
	100	0.6505059	0.0022886	0.0011323	0.0011563	0.6394785	0.0089998
	150	0.6487543	0.0009199	0.0007453	0.0001746	0.6413451	0.0059720
(5.00, 3.00)	10	1.6151610	0.0122326	0.0078522	0.0043804	1.4233272	0.0642430
	25	1.6096121	0.0082284	0.0041792	0.0040493	1.5148751	0.0337116
	50	1.6040354	0.0062919	0.0022848	0.0040071	1.5527492	0.0183038
	75	1.5968800	0.0041054	0.0015350	0.0025704	1.5618314	0.0122570
	100	1.5903742	0.0021890	0.0011123	0.0010767	1.5638001	0.0088434
	150	1.5862591	0.0008668	0.0007365	0.0001302	1.5683183	0.0059017
(0.10, 0.25)	10	0.2933254	0.0079970	0.0124091	0.0044121	0.2411342	0.1053085
	25	0.2919597	0.0056215	0.0051587	0.0004628	0.2713854	0.0419446
	50	0.2909065	0.0047519	0.0025449	0.0022070	0.2806906	0.0204502
	75	0.2896090	0.0029249	0.0016489	0.0012761	0.2828364	0.0131951
	100	0.2884410	0.0012049	0.0011726	0.0000323	0.2833849	0.0093325
	150	0.2877719	0.0001657	0.0007632	0.0005976	0.2844108	0.0061242

- As our target is to obtain the risk of $\hat{\lambda}$ under the actual loss function, we compare it with a well known loss function used for the scale parameter i.e. weighted squared error loss function. It may be hypothesized that though the actual loss is a distance loss function derived using two probability density functions, the actual loss function may attain smaller risk value at $\hat{\lambda}$ than the weighted squared error loss function. From the numerical results presented in Table 1, it is observed that the integrated risk is smaller under actual loss function than the WSELF. Also the obtained results are graphically verified from Figure 1.

- In most of the cases, the Bayes estimator performs well under the actual loss function for the chosen parameter combinations. As the sample size increases, the integrated risks of $\hat{\lambda}$ under SLF and ALF is approximately close.
- Though we have modified the actual loss to define the simplified loss function by omitting the parameter of interest λ from denominator, but from the absolute difference column in Table 1, we observe that the difference between the integrated risk values under both the loss functions is very small. Therefore, our approach to attain the risk at $\hat{\lambda}$ under actual loss function by using simplified Bayes estimate is satisfactory.



Simplified loss values | Actual loss values | Weighted squared error loss values

Figure 2: Histogram of the loss functions for randomly chosen generated λ.

6. Real Data Analysis

In this section, we demonstrate the results obtained in the above sections by analyzing two real-life data examples. Before applying these two datasets for illustration, a natural concern is how well these datasets fit with the considered model, i.e. two-parameter Rayleigh distribution (Banerjee and Seal, 2022). In order to evaluate the fits of the model with the available data three goodness-of-fit statistics are used. The measures including Anderson-Darling (AD), Cramèr-Von Mises (CVM) and Kolmogorov-Smirnov (K-S) statistics are used to compute goodness-of-fits. These three test statistics values are frequently used to observe how well a given CDF fits with the empirical distribution for a given sample. The detail descriptions of AD and CVM test statistics are found in Chen and Balakrishnan (1995).

Table 2: Parameter estimates with their standard error (in parentheses) and goodness-of-fit measures for real datasets.

Data	MLE (S.E.)	K-S (P- value)	CVM (P- value)	AD (P- value)
Dataset I	$\hat{\mu}$ = 1.84055 (0.05511) $\hat{\lambda}$ = 0.53625 (0.07784)	0.06813 (0.93177)	0.04247 (0.92153)	0.27956 (0.95251)
Dataset II	$\hat{\mu}$ = 0.02284 (0.02637) $\hat{\lambda}$ = 4.80708 (0.84164)	0.09224 (0.78852)	0.08440 (0.66887)	0.44968 (0.79766)

The smaller values of these statistics are the indications of better fit for the statistical model. The MLEs and their corresponding standard error (S.E.) of the model parameters are listed in Table 2. The numerical values along with the associated *P-values* at 5% level of significance of the mentioned test statistics are also listed in the same Table. Also, to verify the fitting of the considered datasets in two-parameter Rayleigh distribution, some diagnostic plots like histogram with estimated PDF, estimated CDF, *P-P* plots and the empirical CDF F_n versus fitted CDF $F(X)$ plots are separately provided for each dataset in Figures 3 and 4. All the required computational results are obtained by using the R Core Team (2020) software.

6.1. Application I

The data used for the first example was originated by Bader and Priest (1982) and it includes 63 strength measures in GPA (Giga-Pascals), for single carbon fibers and impregnated 1000 carbon-fiber tows tested under tension at gauge lengths of 10 mm. This strength dataset in Table 3 was also analyzed previously by Kundu and Gupta (2006) and Bi and Gui (2017). The results provided in Table 2 for the dataset I, indicate that the data fitted well and we use this for further analysis.

Table 3: Strength measured in GPA for carbon fibers tested under tension at gauge lengths of 10 mm.

1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445	2.454	2.474
2.518	2.522	2.525	2.532	2.575	2.614	2.616	2.618	2.624	2.659	2.675	2.738
2.740	2.856	2.917	2.928	2.937	2.937	2.977	2.996	3.030	3.125	3.139	3.145
3.220	3.223	3.235	3.243	3.264	3.272	3.294	3.332	3.346	3.377	3.408	3.435
3.493	3.501	3.537	3.554	3.562	3.628	3.852	3.871	3.886	3.971	4.024	4.027
4.225	4.395	5.020									

The sample mean and the sample variance for strength data are 3.059 and 0.386 respectively. We calculate the risk values under three different loss functions, i.e. simplified loss function (SLF), actual loss function (ALF) and weighted squared error loss function (WSELF) for different choices of hyperparameter values. Also, to find the average simplified Bayes estimate and average BE under WSELF, we adopt the similar method discussed in the previous sections. We see from Table 4, that the risk values using actual Hellinger distance loss for the scale parameter λ is smaller compared with the traditional weighted squared error loss function. Further, the absolute difference column in Table 4, shows small values and hence the modification done in ALF to obtain the Bayes estimate also has been justified for the real dataset I.

Table 4: Average Bayes estimates (BE) and Risk values of the scale parameter λ.

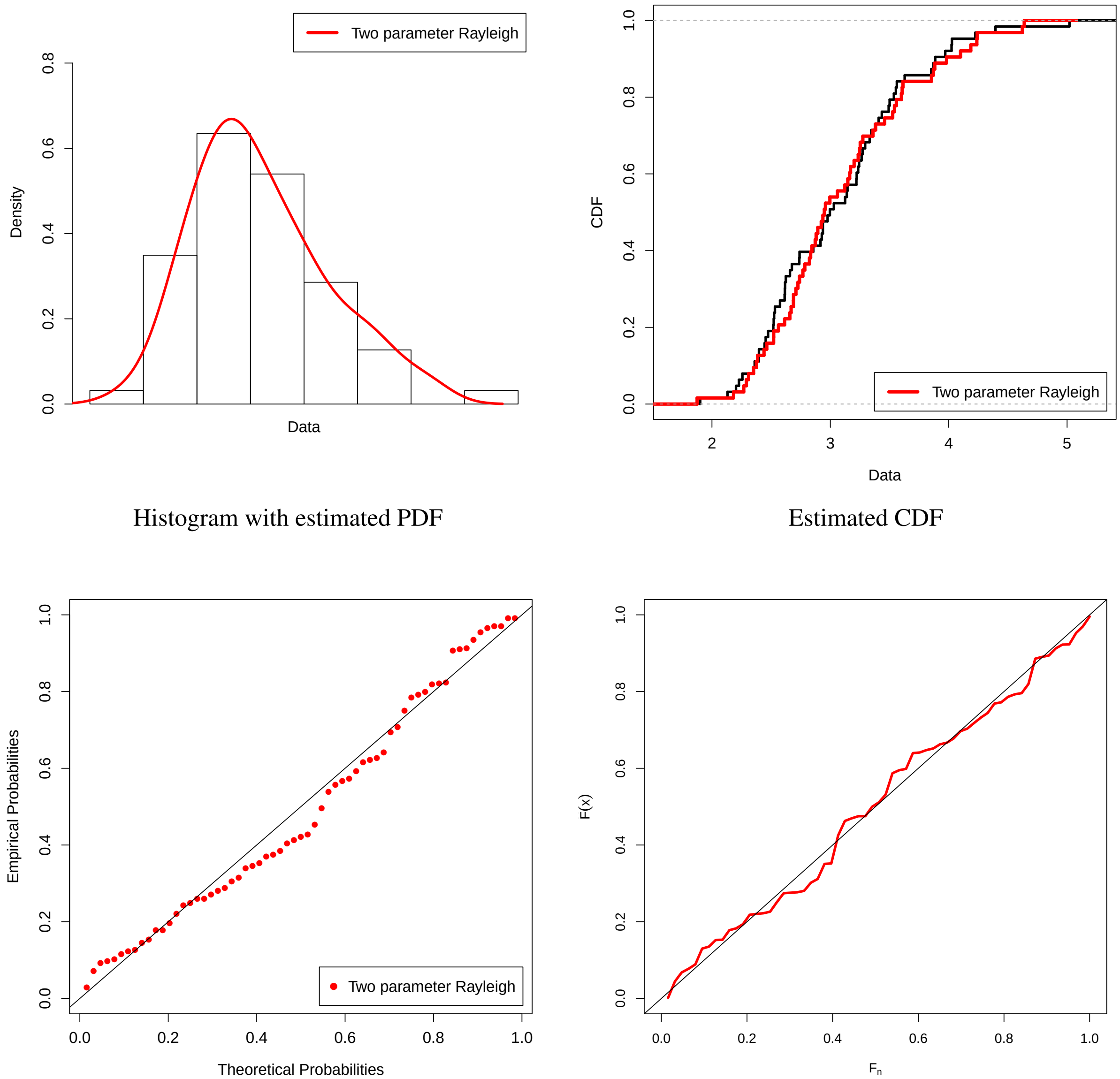
Parameter choices	Simplified BE	Risk values under SLF	Risk values under ALF	Absolute difference	BE under WSELF	Risk values under WSELF
$(0.75, 0.25)$	0.551337	0.007913	0.002014	0.005899	0.536138	0.015964
$(1.00, 1.00)$	0.549870	0.006667	0.001977	0.004690	0.534771	0.015679
$(1.00, 0.50)$	0.552290	0.008794	0.002014	0.006780	0.537125	0.015954
$(0.25, 1.0)$	0.543401	0.000772	0.001943	0.001172	0.528302	0.015514
$(2.0, 0.25)$	0.562190	0.017535	0.002156	0.015379	0.546991	0.017018
$(0.50, 1.50)$	0.543178	0.000617	0.001926	0.001310	0.528145	0.015379

6.2. Application II

The second data used in this example for illustrative purpose is about the lifetime (Hours) of T8 fluorescent lamps for 50 devices. The dataset is collected from the UK National Physical Laboratory and for additional information about the data, one can visit: `http://www.npl.co.uk/`. Recently, Ahmed (2020) also analyzed this dataset. It is noted from Table 2 that the dataset fitted well with considered model and the null hypothesis that the data follows the two-parameter Rayleigh distribution cannot be rejected.

Table 5: Lifetime (Hours) of T8 fluorescent lamps for 50 devices.

0.445	0.493	0.285	0.564	0.760	0.381	0.690	0.579	0.636	0.238
0.149	0.244	0.126	0.796	0.405	0.553	0.780	0.431	0.184	0.375
0.198	0.890	0.192	0.463	0.486	0.521	0.366	0.486	0.116	0.511
0.612	0.117	0.384	0.326	0.057	0.412	0.586	0.517	0.570	0.588
0.497	0.246	0.234	0.228	0.552	0.893	0.403	0.458	0.134	0.338



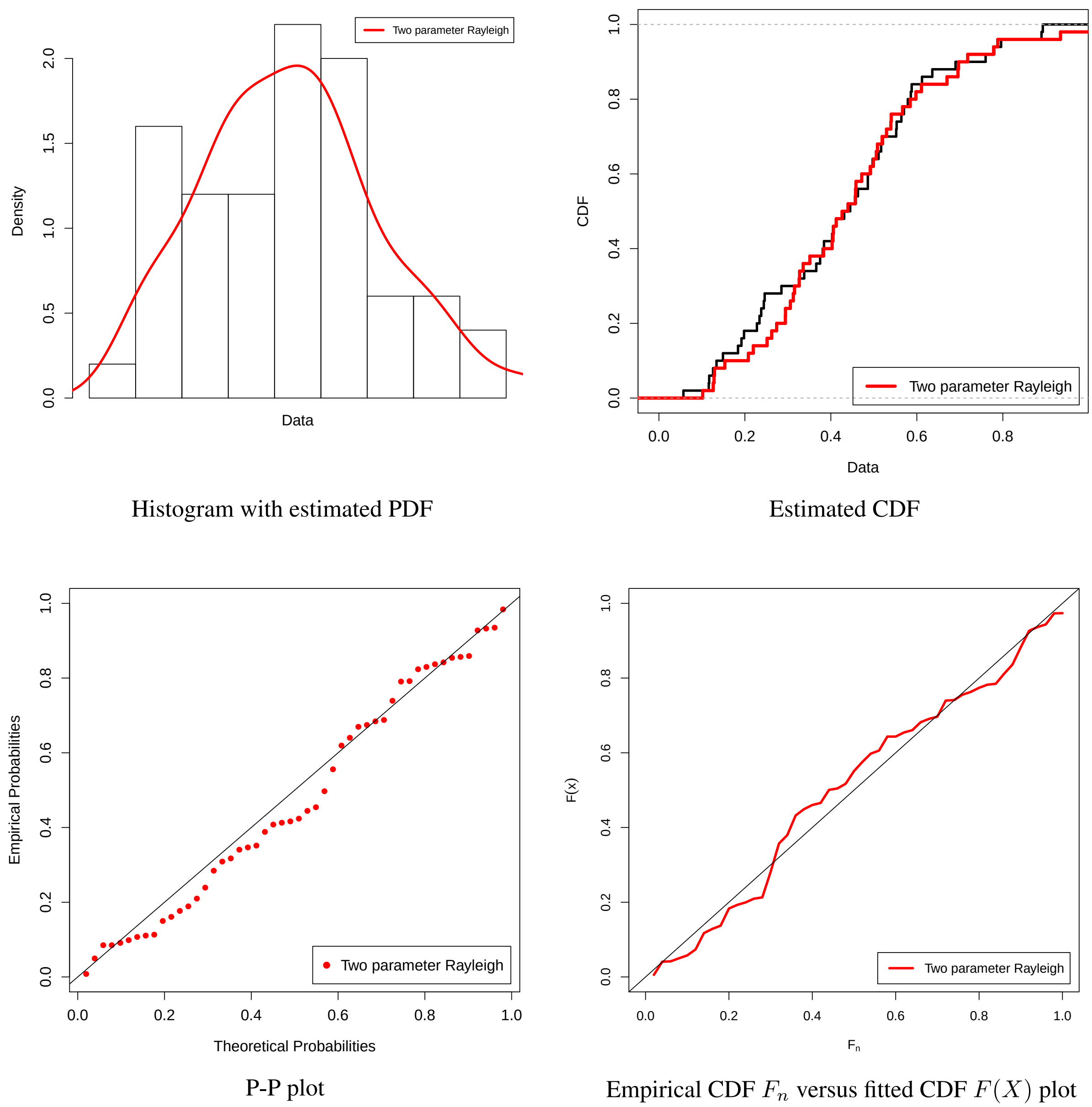
Histogram with estimated PDF — Estimated CDF

P-P plot — Empirical CDF F_n versus fitted CDF $F(X)$ plot

Figure 3: Histogram with estimated PDF, Estimated CDF, P-P plot and empirical CDF F_n versus fitted CDF $F(X)$ plots of two-parameter Rayleigh distribution for dataset I.

Table 6: Average Bayes estimates (BE) and Risk values of the scale parameter λ.

Parameter choices	Simplified BE	Risk values under SLF	Risk values under ALF	Absolute difference	BE under WSELF	Risk values under WSELF
$(0.25, 0.10)$	4.900330	0.002118	0.002485	0.000367	4.728759	0.020002
$(0.50, 0.10)$	4.924831	0.004603	0.002499	0.002104	4.753260	0.020067
$(1.00, 0.25)$	4.900347	0.002337	0.002414	0.000078	4.731312	0.019407
$(2.00, 0.25)$	4.996903	0.012023	0.002507	0.009516	4.827869	0.019968
$(2.50, 0.50)$	4.924026	0.005079	0.002318	0.002761	4.759052	0.018536
$(3.00, 0.50)$	4.971145	0.009806	0.002364	0.007441	4.806171	0.018803



Histogram with estimated PDF

Estimated CDF

P-P plot

Empirical CDF F_n versus fitted CDF $F(X)$ plot

Figure 4: Histogram with estimated PDF, Estimated CDF, P-P plot and empirical CDF F_n versus fitted CDF $F(X)$ plots of two-parameter Rayleigh distribution for dataset II.

The sample mean and sample variance are reported as 0.430 and 0.043 for the lifetime of 50, T8 fluorescent lamps data. By using this dataset, we also compare the risk values obtained using the loss derived from the Hellinger distance measure and the traditional weighted squared error loss function. Similarly, as the result mentioned in previous Subsection 6.1, here also from Table 6, we see that the risk values obtained using the actual Hellinger distance loss is smaller in comparison with the traditionally used WSELF for different combinations of hyperparameters. For calculation complexities some adjustment are done in the actual Hellinger loss function. Therefore, we also compute the risk values under this simplified loss function and further the absolute differences between actual and simplified losses are also shown in Table 6. The results show that the differences are very small and that also justified our method used in this study.

7. Conclusion

It is important to choose loss function which measures the nearness of the parameter and its estimator. Most of the works are based on the assumption that the underlying loss is squared error loss function or its variant due to the symmetry of the loss function. But it may not reflect the original phenomenon of real situation. In this study, we have taken a distance type loss function which measures the loss directly from the probability density function.

Hellinger distance loss function is considered to estimate the scale parameter of the two-parameter Rayleigh distribution. But it is intractable to obtain risk under the actual Hellinger loss function. So, we modified the loss function by adjusting the actual loss function and have calculated the Bayes estimate under simplified loss function. As our aim is to do that under ALF, so we put the simplified Bayes estimate in the ALF and that estimate is termed as Pseudo Bayes estimate with respect to the actual loss function. Finally, the performance of the estimator under both the loss functions are compared with the traditionally used weighted squared error loss function (WSELF). An extensive simulation study has been performed to investigate the performance of the estimators under three different loss functions. The performance is evaluated in terms of integrated risk values. Numerical results showed that the estimator outperforms under ALF than others as it attains minimum risk values for the chosen parameter combinations. It is also observed that the location parameter of Rayleigh distribution has no impact on the risk values of the scale parameter. Lastly, two real-life datasets have been considered to illustrate the applicability of the proposed methodology for parameter estimation.

References

1. Ahmed, A., Ahmad, S., and Reshi, J. (2013). Bayesian analysis of rayleigh distribution. *International Journal of Scientific and Research Publications*, 3(10):1–9.
2. Ahmed, M. A. (2020). On the alpha power kumaraswamy distribution: Properties, simulation and application. *Revista Colombiana de Estadística*, 43(2):285–313, doi: https://doi.org/10.15446/rce.v43n2.83598.
3. Ahmed, Z., Nofal, Z. M., Abd El Hadi, N. E., et al. (2015). Exponentiated transmuted generalized raleigh distribution: A new four parameter rayleigh distribution. *Pakistan journal of statistics and operation research*, pages 115–134, doi:https://doi.org/10.18187/pjsor.v11i1.873.
4. Akhter, A. S. and Hirai, A. S. (2009). Estimation of the scale parameter from the rayleigh distribution from type ii singly and doubly censored data. *Pakistan Journal of Statistics and Operation Research*, 5(1):31–45, doi:https://doi.org/10.18187/pjsor.v5i1.
5. Bader, M. and Priest, A. (1982). Statistical aspects of fibre and bundle strength in hybrid composites. *Progress in science and engineering of composites*, pages 1129–1136.
6. Banerjee, P. and Bhunia, S. (2022). Exponential transformed inverse rayleigh distribution: Statistical properties and different methods of estimation. *Austrian Journal of Statistics*, 51(4):60–75, doi:10.17713/ajs.v51i4.1338.
7. Banerjee, P. and Seal, B. (2022). Partial bayes estimation of two parameter gamma distribution under non-informative prior. *Statistics, Optimization & Information Computing*, 10(4):1110–1125.
8. Bi, Q. and Gui, W. (2017). Bayesian and classical estimation of stress-strength reliability for inverse weibull lifetime models. *Algorithms*, 10(2):71.
9. Birgé, L. (1985). Non-asymptotic minimax risk for hellinger balls. *Probab. Math. Statist*, 5(1):21–29.
10. Chen, G. and Balakrishnan, N. (1995). A general purpose approximate goodness-of-fit test. *Journal of Quality Technology*, 27(2):154–161.
11. Dey, S. (2009). Comparison of bayes estimators of the parameter and reliability function for rayleigh distribution under different loss functions. *Malaysian Journal of Mathematical Sciences*, 3(2):247–264.
12. Dey, S. and Das, M. (2007). A note on prediction interval for a rayleigh distribution: Bayesian approach. *American Journal of Mathematical and Management Sciences*, 27(1-2):43–48, doi:https://doi.org/10.1080/01966324.2007.10737687.
13. Dey, S., Dey, T., and Kundu, D. (2014). Two-parameter rayleigh distribution: Different methods of estimation. *American Journal of Mathematical and Management Sciences*, 33(1):55–74, doi:https://doi.org/10.1080/01966324.2013.878676.
14. Dey, T., Dey, S., and Kundu, D. (2016). On progressively type-ii censored two-parameter rayleigh distribution. *Communications in Statistics - Simulation and Computation*, 45(2):438–455, doi:https://doi.org/10.1080/03610918.2013.856921.

15. Ferguson, T. S. (1967). *Mathematical Statistics - A Decision Theoretic Approach, 1st Edition*. Academic Press.
16. Fundi, M. D., Njenga, E. G., and Keitany, K. G. (2017). Estimation of parameters of the two-parameter rayleigh distribution based on progressive type-ii censoring using maximum likelihood method via the nr and the em algorithms. *American Journal of Theoretical and Applied Statistics*, 6(1):1–9, doi: 10.11648/j.ajtas.20170601.11.
17. Howlader, H. and Hossain, A. (1995). On bayesian estimation and prediction from rayleigh based on type ii censored data. *Communications in Statistics - Theory and Methods*, 24(9):2251–2259, doi:10.1080/03610929508831614.
18. Jahanshahi, S. M. A., Rad, A. H., and Fakoor, V. (2016). Goodness-of-fit test for rayleigh distribution based on hellinger distance. *Annal Data Science*, 3:401–411, doi:https://doi.org/10.1007/s40745–016–0088–6.
19. Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distribution, Vol. 1*. John Wiley & Sons, New York, USA.
20. Kamińska, A. and Porosiński, Z. (2009). On robust bayesian estimation under some asymmetric and bounded loss function. *Statistics*, 43(3):253–265, doi:https://doi.org/10.1080/02331880802264318.
21. Khan, H. M. R., Provost, S. B., and Singh, A. (2010). Predictive inference from a two-parameter rayleigh life model given a doubly censored sample. *Communications in Statistics - Theory and Methods*, 39(7):1237–1246, doi:https://doi.org/10.1080/03610920902871453.
22. Kundu, D. and Gupta, R. D. (2006). Estimation of p [y¡ x] for weibull distributions. *IEEE Trans. Reliab.*, 55(2):270–280, doi: 10.1109/TR.2006.874918.
23. Le Cam, L. (2012). *Asymptotic methods in statistical decision theory*. Springer, Springer Science & Business Media.
24. Mkolesia, A., Kikawa, C. R., and Shatalov, M. Y. (2016). Estimation of the rayleigh distribution parameter. *Transylvanian Review*, 24(8):Special Issue.
25. R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
26. Rayleigh, L. (1880). Xii. on the resultant of a large number of vibrations of the same pitch and of arbitrary phase. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 10(60):73–78, doi:https://doi.org/10.1080/14786448008626893.
27. Seo, J. I., Jeon, J. W., and Kang, S. B. (2016). Exact interval inference for the two-parameter rayleigh distribution based on the upper record values. *Journal of Probability and Statistics*, pages 1–5.
28. Shemyakin, A. (2014). Hellinger distance and non-informative priors. *Bayesian Analysis*, 9(4):923–938, doi: 10.1214/14–BA881.
29. Tahmasebi, S., Hosseini, E. H., and Jafari, A. A. (2017). Bayesian estimation for rayleigh distribution based on ranked set sampling. *New Trends in Mathematical Sciences*, 5(4):97–106.
30. Zamanzade, E. and Mahdizadeh, M. (2017). Goodness of fit tests for rayleigh distribution based on phi-divergence. *Revista Colombiana de EstadÍstica*, 40(2):279–290, doi: https://doi.org/10.15446/rce.v40n2.60375.