Two-Parameter Rayleigh Distribution: Different Methods of Estimation

Sanku Dey[*], Tanujit Dey[†] and Debasis Kundu[‡]

[*]Department of Statistics, St. Anthony's College, Shillong, Meghalaya, India. Email: sanku_dey2k2003@yahoo.co.in

[†]Department of Mathematics, Jones Hall, The College of William & Mary, Williamsburg, Virginia, USA. Email: tdey@wm.edu

[‡]Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur, India. Email: kundu@iitk.ac.in

Abstract

In this paper we have considered different methods of estimation of the unknown parameters of a two-parameter Rayleigh distribution both from the frequentists and Bayesian view points. First we briefly describe different frequentists approaches, namely maximum likelihood estimators, moments estimators, L-moment estimators, percentile based estimators and least squares estimators, and compare them using extensive numerical simulations. We have also considered Bayesian inferences of the unknown parameters. It is observed that the Bayes estimates and the associated credible intervals cannot be obtained in explicit forms, and we have suggested to use importance sampling technique to compute the Bayes estimates and the associated credible intervals. We analyze one data set for illustrative purposes.

Key Words and Phrases: Maximum likelihood estimators; method of moment estimators; L-moment estimators; least squares estimators; weighted least squares estimators; percentile based estimators; Bayes estimators, asymptotic distribution; simulation consistent estimators.

1 Introduction

Lord Rayleigh (1880) introduced the Rayleigh distribution in connection with a problem in the field of acoustics. Since then, extensive work has taken place related to this distribution in different areas of science and technology. It has some nice relations with some of the well known distributions like Weibull, chi-square or extreme value distributions. An important characteristic of the Rayleigh distribution is that its hazard function is an increasing function of time. It means that when the failure times are distributed according to the Rayleigh law, an intense aging of the equipment/ item takes place. Estimations, predictions and inferential issues for one parameter Rayleigh distribution have been extensively studied by several authors. Interested readers may have a look at the book by Johnson, Kotz and Balakrishnan (1994) for an excellent exposure of the Rayleigh distribution, and see also Abd-Elfattah, Hassan and Ziedean (2006), Dey and Das (2007), Dey (2009) for some recent references.

In this paper we consider two-parameter Rayleigh distribution; one scale and one location parameter, and it has the following probability density function (PDF)

$$f(x;\lambda,\mu) = 2\lambda(x-\mu)e^{-\lambda(x-\mu)^2}; \qquad x > \mu, \quad \lambda > 0. \tag{1}$$

The corresponding cumulative distribution function (CDF) for $x > \mu$, is as follows;

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

Here λ and μ are the scale and location parameters respectively. Interestingly, although extensive work has been done on one-parameter Rayleigh distribution, not much attention has been paid to two-parameter Rayleigh distribution, although Johnson, Kotz and Balakrishnan (1994) briefly mentioned about the two-parameter Rayleigh distribution in their book. Recently, Khan, Provost and Singh (2010) considered two-parameter Rayleigh distribution and discussed some inferential issues.

The main aim of this paper is to consider different estimation procedures of the two-parameter Rayleigh distribution both from the Bayesian and frequentist points of view. We first consider the most natural frequentist's estimators namely the maximum likelihood estimators (MLEs). The MLEs of the two-parameter Rayleigh distribution cannot be obtained in explicit forms. They can be obtained by solving a one dimensional optimization process. We have provided a very simple iterative technique which can be used to compute the MLEs of the unknown parameters. The two-parameter Rayleigh distribution does not satisfy the standard regularity conditions of Cramer. In fact, some elements of the expected Fisher information matrix are not finite. Therefore, the standard asymptotic normality results of the maximum likelihood estimators do not hold here. We provide the asymptotic distribution of the MLEs, based on the results of Smith (1995).

Although, the MLEs do not have explicit forms, the method of moment estimators (MMEs) can be obtained in explicit forms. We provide the MMEs and study their asymptotic properties. We also consider other estimators, namely least squares estimators, weighted least squares estimators, percentile estimators and L-moment estimators. We further consider the Bayes estimators of the unknown parameters under the assumptions of independent gamma and uniform priors on the scale and shape parameters respectively. It is observed that the Bayes estimators under the squared error loss function cannot be obtained in closed forms. We have suggested use of importance sampling procedure to compute simulation consistent Bayes estimate and the associated credible interval. We compare the performances of the different methods using extensive computer simulations. Finally, we analyze a data set for illustrative purposes.

Rest of the paper is organized as follows. In Section 2, we provide the MLEs of the unknown parameters and discuss their asymptotic properties. The MMEs and their asymptotic properties are presented in Section 3. Percentile estimators and least squares estimators are provided in Section 4 and Section 5 respectively. L-moment estimators are provided in Section 6. Bayes estimators are presented in Section 7. In Section 8 we present the Monte Carlo simulation results. The analysis of a real data set is provided in Section 9, and finally conclusions appear in Section 10.

2 Maximum Likelihood Estimators

Let $x_1, \cdots, x_n$ be a random sample of size n from (1), then the log-likelihood function $l(\mu, \lambda)$ without the additive constant for $\mu < x_{(1)} = \min\{x_1, \cdots, x_n\}$ can be written as

$$l(\mu, \lambda) = C + n \ln \lambda + \sum_{i=1}^{n} \ln(x_i - \mu) - \lambda \sum_{i=1}^{n} (x_i - \mu)^2. \tag{3}$$

The two normal equations become

$$\frac{\partial}{\partial \lambda} l(\mu, \lambda) = \frac{n}{\lambda} - \sum_{i=1}^{n} (x_i - \mu)^2 = 0 \tag{4}$$

$$\frac{\partial}{\partial \mu} l(\mu, \lambda) = -\sum_{i=1}^{n} (x_i - \mu)^{-1} + 2\lambda \sum_{i=1}^{n} (x_i - \mu) = 0. \tag{5}$$

From (4), we obtain the MLE of λ as a function of μ, say $\widehat{\lambda}(\mu)$, as

$$\widehat{\lambda}(\mu) = \frac{n}{\sum_{i=1}^{n} (x_i - \mu)^2}. \tag{6}$$

Substituting $\widehat{\lambda}(\mu)$ in (3), we obtain the profile log-likelihood function of μ without the additive constant as

$$g(\mu) = -n \ln(\sum_{i=1}^{n} (x_i - \mu)^2) + \sum_{i=1}^{n} \ln(x_i - \mu) = \sum_{i=1}^{n} \ln \left[\frac{(x_i - \mu)}{\sum_{i=1}^{n} (x_i - \mu)^2} \right]. \tag{7}$$

Therefore, the MLE of μ, say $\widehat{\mu}_{MLE}$, can be obtained by maximizing (7) with respect to μ. It can be shown that the maximum of (7) can be obtained as a fixed point solution of the following equation

$$h(\mu) = \mu, \tag{8}$$

where

$$h(\mu) = 2\sum_{i=1}^{n}(x_i - \mu)^2 \times \sum_{i=1}^{n}(x_i - \mu) \times \sum_{i=1}^{n}(x_i - \mu)^{-1}. \tag{9}$$

Once $\widehat{\mu}_{MLE}$ is obtained, the MLE of λ, $\widehat{\lambda}_{MLE} = \widehat{\lambda}(\widehat{\mu}_{MLE})$ can be easily obtained. Observe that very simple iterative technique namely $h(\mu^{(j)}) = \mu^{(j+1)}$, where $\mu^{(j)}$ is the $j-th$ iterate, can be used to solve (8).

Now we want to study the variances and distributional properties of $\widehat{\mu}_{MLE}$ and $\widehat{\lambda}_{MLE}$. Since they are not in explicit forms, it is expected that the exact distributions of the MLEs will not be possible to obtain. We therefore, mainly rely on the asymptotic properties of the MLEs. The two-parameter Rayleigh distribution does not satisfy the standard regularity conditions of Cramer. Note that

$$E\left(\frac{\partial^2}{\partial\mu^2} l(\mu,\lambda)\right) = 2n\lambda + E\left(\sum_{i=1}^{n}(X_i - \mu)^2\right) = \infty. \tag{10}$$

From Theorem 3 of Smith (1985) it follows that $(\widehat{\mu}_{MLE} - \mu)$ is asymptotically normally distributed with mean 0 and $V(\widehat{\mu}_{MLE} - \mu) = O\left(\frac{1}{n \ln n}\right)$, and $(\widehat{\lambda}_{MLE} - \lambda)$ is asymptotically normally distributed with mean 0 and $V(\widehat{\lambda}_{MLE} - \lambda) = -\left[E\left(\frac{\partial^2 l}{\partial \lambda^2}\right)\right]^{-1} = \frac{\lambda^2}{n}$. Moreover, $\widehat{\mu}$ and $\widehat{\lambda}$ are asymptotically independent. Although, the exact asymptotic variance of $\widehat{\mu}$ cannot be obtained in explicit form, from Corollary of Theorem 3 of Smith (1985), it follows that $V(\widehat{\mu} - \mu)$ can be well approximated by the inverse of the observed information, and in this case it is

$$V(\widehat{\mu}_{MLE} - \mu) \approx \frac{1}{2\mu + \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^{-2}}. \tag{11}$$

Therefore, $100(1-\alpha)\%$ approximate confidence intervals of λ and μ can be obtained as

$$\widehat{\lambda}_{MLE} \pm z_{\alpha/2}\frac{\widehat{\lambda}_{MLE}}{\sqrt{n}} \quad \text{and} \quad \widehat{\mu}_{MLE} \pm z_{\alpha/2}\left(\frac{1}{2\widehat{\mu}_{MLE} + \frac{1}{n}\sum_{i=1}^{n}(x_i - \widehat{\mu}_{MLE})^{-2}}\right)^{1/2}, \tag{12}$$

respectively, where $z_{\alpha/2}$ is the $\alpha/2$-th percentile point of the standard normal distribution.

3 Method of Moment Estimators

The MMEs of the two-parameter Rayleigh distribution can be obtained as

$$\widehat{\lambda}_{MME} = \frac{1}{s^2}\left[1 - \Gamma^2(3/2)\right] \qquad \text{and} \qquad \widehat{\mu}_{MME} = \bar{x} - \widehat{\lambda}_{MME}^{-1/2}\Gamma(3/2). \tag{13}$$

Here $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ are the sample mean and sample variance respectively. The exact distributions of $\widehat{\lambda}_{MME}$ and $\widehat{\mu}_{MME}$ are not possible to obtain. The asymptotic distribution of $\widehat{\lambda}_{MME}$ and $\widehat{\mu}_{MME}$ can be obtained. For that we need the following notations. Suppose the random variable X has a Rayleigh distribution with parameters λ and μ. Let us define

$$a_k = EX^k \quad \text{and} \quad b_k = E(X-\mu)^k = \frac{1}{\lambda^{k/2}}\Gamma\left(\frac{k}{2}+1\right); k = 1, 2, \cdots. \tag{14}$$

Then

$$a_1 = b_1 + \mu, \quad a_2 = b_2 + 2\mu a_1 - \mu^2, \quad a_3 = b_3 + 3\mu a_2 - 3\mu^2 a_1 + \mu^3 \tag{15}$$

$$a_4 = b_4 + 4a_3\mu - 6a_2\mu^2 + 4a_1\mu^3 - \mu^4. \tag{16}$$

Now we will provide the asymptotic properties of the MMEs. Since the first two moments exist, it is immediate that the MMEs are consistent estimators to the corresponding parameters. Using the δ-method we can easily obtain the asymptotic distribution of the MMEs.

$$\left(\sqrt{n}(\widehat{\mu}_{MME} - \mu), \sqrt{n}(\widehat{\lambda}_{MME} - \lambda)\right) \xrightarrow{d} N_2[\mathbf{0}, \mathbf{\Sigma}], \tag{17}$$

where '$\xrightarrow{d}$' means convergence in distribution, and $\mathbf{\Sigma}$ is a 2×2 matrix and it has the following form;

$$\mathbf{\Sigma} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}^{-1},$$

and if we denote

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

then for $c_1 = \dfrac{\Gamma(3/2)}{\sqrt{1-\Gamma^2(3/2)}}$ and $c_2 = 1 - \Gamma^2(3/2)$

$$\begin{aligned} a_{11} &= a_2 - a_1^2, \quad a_{12} = a_{21} = a_3 - a_1 a_2, \quad a_{22} = a_4 - a_2^2 \\ c_{11} &= 1 + c_1(a_2 - a_1^2)^{-1/2} a_1, \quad c_{12} = c_{21} = -\frac{c_1}{2}(a_2 - a_1^2)^{-1/2}, \quad c_{22} = -c_2(a_2 - a_1^2)^{-2}. \end{aligned}$$

It is clear that although both $\widehat{\mu}_{MLE}$ and $\widehat{\mu}_{MME}$ are consistent estimators of μ, $\widehat{\mu}_{MLE}$ converges to μ faster than $\widehat{\mu}_{MME}$.

4 L-Moments Estimators

In this section we provide the L-moments estimators, which can be obtained as the linear combinations of order statistics. The L-moments estimators were originally proposed by Hosking (1990), and it is observed that the L-moments estimators are more robust than the usual moment estimators. The L-moment estimators are also obtained the same way as the ordinary moment estimators, *i.e.* by equating the sample L-moments with the population L-moments.

In this case the L-moments estimators can be obtained by equating the first two sample L-moments with the corresponding population L-moments. The first two sample L-moments are

$$l_1 = \frac{1}{n}\sum_{i=1}^{n} x_{(i)}, \quad \text{and} \quad l_2 = \frac{2}{n(n-1)}\sum_{i=1}^{n}(i-1)x_{(i)} - l_1, \tag{18}$$

and the first two population L-moments are

$$\lambda_1 = E(X_{1,1}) = E(X) = \frac{1}{\lambda^{1/2}}\Gamma(3/2) + \mu \tag{19}$$

$$\lambda_2 = \frac{1}{2}(E(X_{2:2}) - E(X_{2:1})) = \frac{\Gamma(3/2)}{\lambda^{1/2}} \times \frac{\sqrt{2}-1}{\sqrt{2}}. \tag{20}$$

Here $X_{m,n}$ denotes the m-th order statistic of a sample of size n. Therefore, the L-moment estimators of μ and λ are

$$\widehat{\mu}_{LME} = l_1 - \frac{\sqrt{2}}{\sqrt{2}-1} l_2 \quad \text{and} \quad \widehat{\lambda}_{LME} = \frac{\Gamma^2(3/2)}{l_2^2} \times \frac{3-2\sqrt{2}}{2}. \tag{21}$$

5 Estimators Based on Percentiles

Kao (1958, 1959) had suggested the estimators based on percentiles, when the distribution function is in closed form. Kao (1958) had proposed and then successfully implemented, see Kao (1959), the percentile based estimators for Weibull distribution. Later on, it has been used quite successfully for other distributions, where the distribution functions are in compact form. Gupta and Kundu (2001) and Kundu and Raqab (2005) used the percentile based estimators and compared with the other estimators for generalized exponential distribution and generalized Rayleigh distribution respectively.

The main advantage of the percentile based estimators is that in many situations, they can be obtained in explicit forms. The percentile based estimators are mainly obtained by minimizing the Euclidean distance between the sample percentile and population percentile points. Since the cumulative distribution function of the two-parameter Rayleigh distribution can be written in the form;

$$F(x;\lambda,\mu) = 1 - e^{-\lambda(x-\mu)^2}, \quad \Rightarrow \quad x = \mu + \left[-\frac{1}{\lambda}\ln(1-F(x;\lambda,\mu))\right]^{1/2}, \tag{22}$$

the percentile based estimators of μ and λ can be obtained by minimizing

$$\sum_{i=1}^{n}\left[x_{(i)} - \mu - \frac{1}{\sqrt{\lambda}}\left(-\ln(1-p_i)\right)^{1/2}\right]^2 \tag{23}$$

with respect to μ and λ. Here $x_{(1)} < \cdots < x_{(n)}$ is the sample order statistics, and p_i denotes some estimate of $F(x_{(i)};\mu,\lambda)$. Although, several estimates of p's are available in the literature, see for example Mann, Schafer and Singpurwalla (1974), $p_i = \frac{i}{n+1}$ seems to an acceptable choice. We have also used $p_i = \frac{i}{n+1}$ in this paper. The minimization of (23) can be performed explicitly, using the linear regression technique and the percentile based estimators (PCE) of μ and λ are

$$\widehat{\mu}_{PCE} = \frac{A\bar{x} - BC}{A - B^2}, \quad \text{and} \quad \widehat{\lambda}_{PCE} = \frac{(A - B^2)^2}{(C - B\bar{x})^2}, \tag{24}$$

where

$$A = \frac{1}{n}\sum_{i=1}^{n} -ln(1 - p_i), \quad B = \frac{1}{n}\sum_{i=1}^{n} [-\ln(1 - p_i)]^{1/2}, \quad C = \frac{1}{n}\sum_{i=1}^{n} x_{(i)} [-\ln(1 - p_i)]^{1/2},$$

6 Least Squares Estimators

The least squares estimators and weighted least squares estimators (LSEs) were proposed by Swain, Venkataraman and Wilson (1988) to estimate the parameters of a Beta distribution. The LSEs of the unknown parameters of a two-parameter Rayleigh distribution can be obtained by minimizing

$$\sum_{j=1}^{n} \left(F(x_{(j)}) - \frac{j}{n+1} \right)^2$$

with respect to unknown parameters μ and λ. $F(\cdot)$ denotes the CDF of the two-parameter Rayleigh distribution as mentioned before. Therefore, in this case, the LSEs of μ and λ, say, $\widehat{\mu}_{LSE}$ and $\widehat{\lambda}_{LSE}$, respectively, can be obtained by minimizing

$$\sum_{j=1}^{n} \left(\left(1 - \exp(-\lambda(x_{(j)} - \mu)^2)\right) - \frac{j}{n+1} \right)^2$$

with respect to μ and λ.

The weighted least squares estimators (WLSEs) of the unknown parameters can be obtained by minimizing

$$\sum_{j=1}^{n} w_j \left(F(x_{(j)}) - \frac{j}{n+1} \right)^2$$

with respect to μ and λ, where w_j denotes the weight function at the j-th point. The weights are taken to be

$$w_j = \frac{1}{V(F(X_{(j)})} = \frac{(n+1)^2(n+2)}{j(n-j+1)}.$$

Therefore, in this case, the WLSEs of μ and λ, say, $\hat{\mu}_{WLSE}$ and $\hat{\lambda}_{WLSE}$, respectively, can be obtained by minimizing

$$\sum_{j=1}^{n} \frac{(n+1)^2(n+2)}{j(n-j+1)} \left(\left(1 - \exp(-\lambda(x_j - \mu)^2)\right) - \frac{j}{n+1} \right)^2$$

with respect to μ and λ.

7 Bayes Estimators

So far we have discussed the estimation of the unknown parameters using different frequentists method. In this section we consider the Bayesian inference of the unknown parameters of the two-parameter Rayleigh distribution when both the parameters are unknown. We mainly discuss the Bayes estimates and the associated credible intervals. Although, we have developed the estimates based on the squared error loss functions, other loss functions can be easily incorporated.

It is well known that when the both the parameters are unknown, the conjugate priors do not exist. Moreover, since all the elements of the expected Fisher information matrix are not finite, Jeffreys' priors also do not exist. We consider the following fairly general priors on λ and μ.

Note that for known μ, λ has a conjugate gamma prior, and therefore, we consider the gamma prior on λ even in this case also as follows:

$$\pi_1(\lambda; a, b) \propto \lambda^{a-1} e^{-b\lambda}; \quad \lambda > 0,\ a > 0,\ b > 0. \tag{25}$$

We have considered the following non-proper uniform prior on μ as follows;

$$\pi_2(\mu) \propto d\mu, \quad -\infty < \mu < \infty, \tag{26}$$

and they are assumed to be independently distributed.

Now based on the observed sample, the joint posterior density function of λ and μ becomes;

$$\pi(\mu,\lambda|Data) = \frac{\lambda^{n+a-1}e^{-\lambda(b+\sum_{i=1}^{n}(x_i-\mu)^2)}\prod_{i=1}^{n}(x_i-\mu)}{\int_0^{\infty}\int_{-\infty}^{x_{(1)}}\lambda^{n+a-1}e^{-\lambda(b+\sum_{i=1}^{n}(x_i-\mu)^2)}\prod_{i=1}^{n}(x_i-\mu)d\mu d\lambda}; \quad 0<\lambda<\infty, -\infty<\mu<x_{(1)}. \tag{27}$$

Therefore, if we want to compute the Bayes estimate of any function of λ and μ, say $\theta(\lambda,\mu)$, it can be obtained as the posterior mean as follows;

$$\widehat{\theta}_{Bayes} = \frac{\int_0^{\infty}\int_{-\infty}^{x_{(1)}}\theta(\lambda,\mu)\lambda^{n+a-1}e^{-\lambda(b+\sum_{i=1}^{n}(x_i-\mu)^2)}\prod_{i=1}^{n}(x_i-\mu)d\mu d\lambda}{\int_0^{\infty}\int_{-\infty}^{x_{(1)}}\lambda^{n+a-1}e^{-\lambda(b+\sum_{i=1}^{n}(x_i-\mu)^2)}\prod_{i=1}^{n}(x_i-\mu)d\mu d\lambda}. \tag{28}$$

Clearly, for general $\theta(\lambda,\mu)$ it cannot be obtained in explicit form. Numerical procedure, or some approximation like Lindley's approximations may be used to compute (28). But unfortunately, by this method although the Bayes estimate can be obtained, the associated credible interval cannot be constructed.

Due to this reason, we propose to use the importance sampling method to compute the Bayes estimate and also to compute the associated credible interval. It is known that by proper importance sampling method, simulation consistent Bayes estimate and the associated credible interval can be constructed (Kundu and Pradhan (2009)). To implement the importance sampling procedure we re-write (28) as follows;

$$\pi(\lambda,\mu|Data) = \frac{f_1(\lambda|\mu,Data)f_2(\mu|Data)h(\mu)}{\int_0^{\infty}\int_{-\infty}^{x_{(1)}}f_1(\lambda|\mu,Data)f_2(\mu|Data)h(\mu)d\mu d\lambda}, \tag{29}$$

where

$$f_1(\lambda|\mu,Data) = \frac{(b+\sum_{i=1}^{n}(x_i-\mu)^2)^{n+a}}{\Gamma(n+a)}\lambda^{n+a-1}e^{-\lambda(b+\sum_{i=1}^{n}(x_i-\mu)^2)}; \quad \lambda>0, \tag{30}$$

$$f_2(\mu|Data) = \frac{k\ \ 2(\sum_{i=1}^{n}x_i-\mu)}{(\mu^2-2\mu\sum_{i=1}^{n}x_i+b+\sum_{i=1}^{n}x_i^2)^{n+a}}; \quad -\infty<\mu<x_{(1)} \tag{31}$$

$$k = (n+a-1)\left(x_{(1)}^2-2x_{(1)}\sum_{i=1}^{n}x_i+b+\sum_{i=1}^{n}x_i^2\right)^{n+a-1}$$

and

$$h(\mu) = \begin{cases} \prod_{i=1}^{n}(x_i - \mu) \Big/ 2(\sum_{i=1}^{n} x_i - \mu) & \text{for} \quad \mu < x_{(1)} \\ 0 & \text{for} \quad \mu \geq x_{(1)}. \end{cases} \tag{32}$$

Note that, $f_1(\lambda|\mu, Data)$ is a gamma density function with the shape and scale parameters as $n + a$ and $b + \sum_{i=1}^{n}(x_i - \mu)^2$ respectively. $f_2(\mu|Data)$ is a proper density function and it has the distribution function for $\mu < x_{(1)}$ as

$$F_2(\mu|Data) = \frac{\left(x_{(1)}^2 - 2x_{(1)} \sum_{i=1}^{n} x_i + b + \sum_{i=1}^{n} x_i^2\right)^{n+a-1}}{(\mu^2 - 2\mu \sum_{i=1}^{n} x_i + b + \sum_{i=1}^{n} x_i^2)^{n+a-1}}, \tag{33}$$

which is easily invertible.

Now we propose to use the following procedure to compute the Bayes estimate of $g(\lambda, \mu)$ and the associated credible interval.

Step 1: Generate (μ_1, λ_1) as

$$\mu_1 \sim f_2(\mu|Data) \quad \text{and} \quad \lambda_1|\mu_1 \sim gamma(n + a, b + \sum_{i=1}^{n}(x_i - \mu_1)^2)$$

Step 2: Repeat Step 1 $(N - 1)$ times to obtain $(\mu_2, \lambda_2), \cdots, (\mu_N, \lambda_N)$

Step 3: Then a simulation consistent estimator of (28) can be obtained as

$$\frac{\sum_{i=1}^{N} \theta(\lambda_i, \mu_i)h(\mu_i)}{\sum_{i=1}^{N} h(\mu_i)} \tag{34}$$

Now we would like to construct the highest posterior density credible interval (HPD) of θ using the generated importance sampling procedure. Suppose θ_p is such that $P(\theta \leq \theta_p|Data) = p$, for $0 < p < 1$. Consider the following function

$$g(\lambda, \mu) = \begin{cases} 1 & \text{if} \quad \theta \leq \theta_p \\ 0 & \text{if} \quad \theta > \theta_p. \end{cases} \tag{35}$$

Clearly, $E(g(\lambda, \mu)|Data) = p$. Therefore, a simulation consistent Bayes estimate of θ_p under squared error loss function can be obtained from the generated sample $\{(\mu_1, \lambda_1), \cdots, (\mu_N, \lambda_N)\}$

as follows. Let

$$w_i = \frac{h(\mu_i)}{\sum_{i=1}^{N} h(\mu_i)},$$

and $\theta_1 = \theta(\lambda_1, \mu_1), \cdots, \theta_N = \theta(\lambda_N, \mu_N)$. Rearrange, $(\theta_1, w_1), \cdots, (\theta_N, w_N)$ as follows $\{(\theta_{(1)}, w_{[1]}), \cdots, (\theta_{(N)}, w_{[N]})\}$, where $\theta_{(1)} < \cdots < \theta_{(N)}$. Note that, $w_{[i]}$'s are not ordered, they are associated with $\theta_{(i)}$'s. Then a simulation consistent Bayes estimate of θ_p can be obtained as $\widehat{\theta}_p = \theta_{(N_p)}$, where

$$\sum_{i=1}^{N_p} w_{[i]} \le p < \sum_{i=1}^{N_p+1} w_{[i]}.$$

Now using the above procedure, a 100(1-α)% credible interval can be obtained as

$$(\widehat{\theta}_\delta, \widehat{\theta}_{\delta+1-\alpha}), \quad \text{for} \quad \delta = w_{[1]}, w_{[1]} + w_{[2]}, \cdots, \sum_{i=1}^{N_{1-\alpha}} w_{[i]}. \tag{36}$$

Therefore, a 100(1-α)% HPD credible interval of θ becomes, $((\widehat{\theta}_{\delta^*}, \widehat{\theta}_{\delta^*+1-\alpha})$, where δ^* is such that for all δ,

$$(\widehat{\theta}_{\delta^*+1-\alpha} - \widehat{\theta}_{\delta^*}) \le (\widehat{\theta}_{\delta+1-\alpha} - \widehat{\theta}_\delta)$$

8 Numerical Simulations

In this section we present some experimental results which have been performed to see the effectiveness of the different methods for different sample sizes. We mainly compare different methods in terms of the biases and mean squared errors. When we compare the different frequentists methods with the Bayes estimate we have taken non-informative prior namely $a = b = 0$. In each case we have taken without loss of generality $\lambda = 1$ and $\mu = 0$. We have considered different sample sizes n = 10, 15 (small), 25, 30 (moderate), 50 (large) and 100 (very large). We present the average estimates of λ and μ based on 10,000 replications based on different methods and the associated mean squared errors are presented in parentheses below. The results are presented in Tables 1 and 2.

Some of the points are quite cleat from Tables 1 and 2. In all the cases it is observed that as sample size increases, the average biases and the mean squared errors decrease. It verifies the consistency properties of all the estimators.

Now comparing the performances of the different estimators, it is observed that among all the estimators presented here, the L-moment estimators have the smallest biases for both μ and λ. The L-moment estimators perform better than the moment estimators in all the cases considered in terms of biases and mean squared errors. In case of μ the L-moment estimators behave better than percentile estimators and least squares estimators in terms of biases and mean squared errors. In case of λ the mean squared errors of the L-moment estimators are almost same with the percentile and least squares estimators.

The biases of the Bayes estimators are slightly larger than the L-moment estimators although they have smaller mean squared errors. The performances of the MLEs are also quite satisfactory, although it has slightly higher biases and mean squared errors compared to the Bayes estimators. Comparing all these we propose to use the MLEs or Bayes estimators for all practical purposes in estimating the parameters of the two-parameter Rayleigh distribution.

9 Data Analysis

In this section we present analysis for illustrative purposes using strength data set originally reported by Badar and Priest (1982). It describes the strength measured in GPA for single carbon fibers and impregnated 1000-carbon fiber tows. The data are presented in Table 3.

Preliminary data analysis indicates that the data are positively skewed. A device called scaled TTT (total time on test) transform and its empirical version are relevant to study the

Table 1: Average estimates of μ and their associated MSE's.

n	$\widehat{\mu}_{MLE}$	$\widehat{\mu}_{MME}$	$\widehat{\mu}_{LME}$	$\widehat{\mu}_{PCE}$	$\widehat{\mu}_{LSE}$	$\widehat{\mu}_{BAYES}$
10	0.1073	0.0384	0.0113	-0.0620	-0.0328	0.0264
	(0.0382)	(0.0433)	(0.0422)	(0.0542)	(0.0579)	(0.0245)
15	0.0713	0.0211	0.0034	-0.0534	-0.0204	0.0248
	(0.0211)	(0.0267)	(0.0256)	(0.0329)	(0.0324)	(0.0111)
25	0.0507	0.0153	0.0054	-0.0359	-0.0028	0.0210
	(0.0117)	(0.0167)	(0.0159)	(0.0194)	(0.0182)	(0.0051)
30	0.0420	0.0110	0.0003	-0.0357	-0.0113	0.0205
	(0.0088)	(0.0127)	(0.0117)	(0.0147)	(0.0125)	(0.0038)
50	0.0226	0.0019	-0.0034	-0.0292	-0.0056	0.0152
	(0.0042)	(0.0082)	(0.0075)	(0.0094)	(0.0077)	(0.0022)
100	0.0156	0.0003	-0.0015	-0.0182	-0.0027	0.0137
	(0.0021)	(0.0041)	(0.0037)	(0.0045)	(0.0036)	(0.0018)

shape of the hazard function of a distribution function. For a family with survival function $S(y)$ and distribution function $F(y)$ with positive supports, the scaled TTT transform is defined for $0 < u < 1$, as $\phi_F(u) = H_F^{-1}(u)/H_F^{-1}(1)$, where $H_F^{-1}(u) = \int_0^{F^{-1}(u)} S(y)dy$. The empirical version of the scaled TTT transform for $j = 1, \ldots, n$, is given by

$$\phi_n(j/n) = H_n^{-1}(j/n)/H_n^{-1}(1) = \left(\sum_{i=1}^{j} x_{(i)} + (n-j)x_{(j)}\right) \Bigg/ \left(\sum_{i=1}^{n} x_{(i)}\right),$$

see Aarset (1987). Aarset (1987) showed that the scaled TTT transform is concave (convex) if the hazard function is increasing (decreasing). To check the shape of the empirical hazard function, we have plotted the scaled TTT transform in Figure 1. Since it is a concave function, it indicates that the hazard function of the distribution will be an increasing function. Therefore, Rayleigh distribution can be used to analyze this data set.

First we would like to compute the maximum likelihood estimators of the unknown parameters. The profile log-likelihood $g(\mu)$ of μ as provided in (7) is provided in Figure 2.

Table 2: Average estimates of λ and their associated MSE's.

n	$\widehat{\lambda}_{MLE}$	$\widehat{\lambda}_{MME}$	$\widehat{\lambda}_{LME}$	$\widehat{\lambda}_{PCE}$	$\widehat{\lambda}_{LSE}$	$\widehat{\lambda}_{BAYES}$
10	1.1268	1.0117	0.9466	0.7770	0.8909	1.2365
	(0.1432)	(0.1615)	(0.1442)	(0.1526)	(0.3147)	(0.1343)
15	1.1383	1.0636	1.0148	0.8702	0.9710	1.2101
	(0.1257)	(0.1504)	(0.1286)	(0.1184)	(0.2028)	(0.1197)
25	1.1213	1.0697	1.0422	0.9303	1.0250	1.1783
	(0.0852)	(0.1063)	(0.0940)	(0.0833)	(0.1190)	(0.0557)
30	1.0954	1.0428	1.0154	0.9197	0.9864	1.1579
	(0.0614)	(0.0704)	(0.0619)	(0.0606)	(0.0702)	(0.0447)
50	1.0743	1.0424	1.0280	0.9576	1.0228	1.1434
	(0.0438)	(0.0535)	(0.0480)	(0.0452)	(0.0557)	(0.0284)
100	1.0442	1.0203	1.0150	0.9700	1.0161	1.1103
	(0.0206)	(0.0261)	(0.0243)	(0.0241)	(0.0259)	(0.0153)

Clearly, it is an unimodal function, and the MLE of μ is obtained as 0.560, which maximizes $g(\mu)$. The MLE of λ becomes 0.648. The 95% confidence intervals of λ and μ based on MLEs as suggested in Section 2, can be obtained as (0.495,0.799) and (0.531,0.619) respectively. The Kolmogorov-Smirnov distance between the fitted and empirical cumulative distribution functions is 0.128 and the associated p values is 0.207. Therefore, based on the MLEs we cannot reject the null hypothesis that the data are coming from a two-parameter Rayleigh

Table 3: Strength Data

0.562	0.564	0.729	0.802	0.950	1.053	1.111	1.115	1.194	1.208
1.216	1.247	1.256	1.271	1.277	1.305	1.313	1.348	1.390	1.429
1.474	1.490	1.503	1.520	1.522	1.524	1.551	1.551	1.609	1.632
1.632	1.676	1.684	1.685	1.728	1.740	1.761	1.764	1.785	1.804
1.816	1.824	1.836	1.879	1.883	1.892	1.898	1.934	1.947	1.976
2.020	2.023	2.050	2.059	2.068	2.071	2.098	2.130	2.204	2.262
2.317	2.334	2.340	2.346	2.378	2.483	2.683	2.835	2.835	

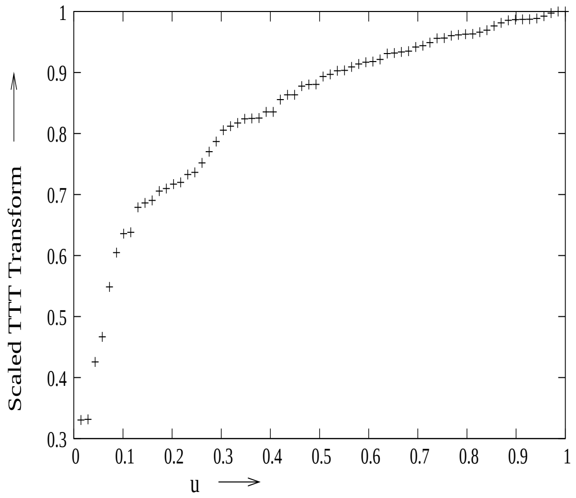

Figure 1: Scaled TTT transform of the strength data.

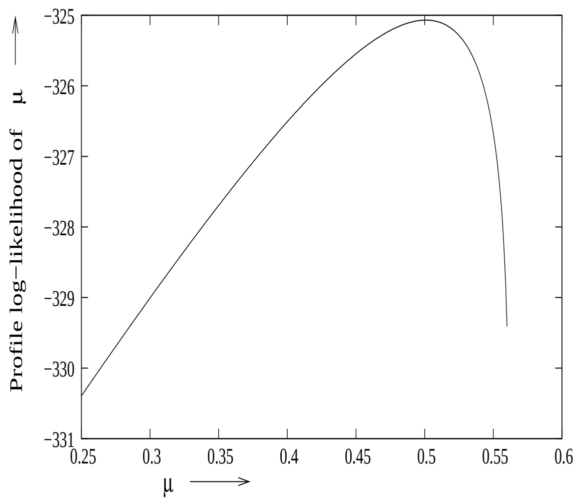

Figure 2: Profile log-likelihood function of μ of the strength data.

distribution.

To compute the Bayes estimates we have assumed non-informative priors namely $a = b = c = d = 0$. Based on 1,000 importance sampling we obtain the Bayes estimates of μ and λ as 0.543 and 0.629 respectively. The 95% symmetric credible intervals of λ and μ are (0.445, 0.798) and (0.530, 0.562) respectively. The Kolmogorov-Smirnov distance between the fitted and empirical cumulative distribution functions is 0.130 and the associated p values is 0.191. Therefore, Bayes estimates also provide satisfactory goodness of fit of the two-parameter Rayleigh distribution to the strength data.

10 Conclusions

In this paper we have considered different estimation procedures for estimating the unknown location and scale parameters of a two-parameter Rayleigh distribution. We have mainly considered the maximum likelihood estimators, the method of moment estimators, the L-moment estimators, percentile estimators, least squares estimators and the Bayes estimators. It is not possible to compare different methods theoretically, and we have used some simulations to compare different estimators. We have compared different estimators mainly with respect to biases and mean squared errors. It is observed that the Bayes estimators with non-informative priors work very well in terms of biases and mean squared errors, although it is quite involved computationally. The performances of the maximum likelihood estimators are also quite satisfactory. We recommend use of the MLEs or Bayes estimators for all practical purposes.

Acknowledgments

The authors would like to thank the editors and the reviewers who helped to substantially improve the paper. The authors are thankful to Professor R.G. Surles for providing the strength data.

References

[1] Abd-Elfattah, A. M., Hassan, A. S. & Ziedan, D. M. (2006). Efficiency of maximum likelihood estimators under different censored sampling schemes for Rayleigh distribution. *Interstat*, March issue, 1.

[2] Aarset, M.V. (1987). How to identify a bathtub shaped hazard rate?. *IEEE Transactions on Reliability*, 36, 106 - 108.

[3] Badar, M.G. & Priest, A.M. (1982). Statistical aspects of fiber and bundle strength in hybrids composite. *Progress in Science and Engineering Composites (T. Hayakashi, K. Kawata & S. Umekawa eds.)*, ICCM-IV, Tokyo, 1129 - 1136.

[4] 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, 247 - 264.

[5] Dey, S. & Das, M.K. (2007). A Note on Prediction Interval for a Rayleigh Distribution: Bayesian Approach. *American Journal of Mathematical and Management Science*, 1&2, 43 - 48.

[6] Gupta, R.D. & Kundu, D. (2001). Generalized exponential distributions: different methods of estimation. *Journal of Statistical Computation and Simulation*, 69, 315 - 338.

[7] Hosking, J.R.M. (1990). L-moment: analysis and estimation of distributions using linear combinations of order statistics. *Journal of Royal Statistical Society, Ser. B*, 52, 105 - 124.

[8] Johnson, N.L., Kotz, S. & Balakrishnan, N. (1994). *Continuous Univariate Distribution* (Volume 1). John Wiley & Sons, New York.

[9] Kao, J.H.K. (1958). Computer methods for estimating Weibull parameters in reliability studies. *Trans. IRE Reliability Quality Control*, 13, 15 - 22.

[10] Kao, J.H.K. (1959). A graphical estimation of mixed Weibull parameters in life testing electron tube. *Technometrics*, 1, 389 - 407.

[11] Khan, H.M.R., Provost, S.B. & Singh, A. (2010). Predictive inference from a two-parameter Rayleigh life model given a doubly censored sample. *Communications in Statistics - Theory and Methods*, 39, 1237 - 1246.

[12] Kundu, D. & Raqab, M.Z. (2005). Generalized Rayleigh distribution: different methods of estimation. *Computational Statistics and Data Analysis*, 49, 187 - 200.

[13] Rayleigh, J.W.S. (1880). On the resultant of a large number of vibrations of the some pitch and of arbitrary phase. *Philosophical Magazine, 5-th Series*, 10, 73 - 78.

[14] Smith, R.L. (1995). Maximum likelihood estimation in a class of non-regular cases. *Biometrika*, 72, 67 - 90.

[15] Swain, J.J., Venkataraman, S. & Wilson, J.R. (1988). Least squares estimation of distribution functions in Johnson's translation system. *Journal of Statistical Computation and Simulation*, 29, 271 - 297.