Minimax Estimation of the Parameter of Maxwell Distribution Under Different Loss Functions

Lanping Li

Department of Basic Subjects, Hunan University of Finance and Economics, Changsha, China

Email address: lilanping1981@163.com

To cite this article:

Received: May 21, 2016; Accepted: June 6, 2016; Published: June 23, 2016

Abstract: The aim of this article is to study the Bayes estimation and minimax estimation of the parameter of Maxwell distribution. Bayes estimators are obtained with non-informative quasi-prior distribution under different loss functions, namely, weighted squared error loss, squared log error loss and entropy loss functions. Then the minimax estimators of the parameter are obtained by using Lehmann’s theorem. Finally, performances of these estimators are compared in terms of risks.

Keywords: Bayes Estimator, Minimax Estimator, Squared Log Error Loss, Entropy Loss, Maxwell Distribution

1. Introduction

The Maxwell distribution was first introduced by Maxwell in 1860, and since then, the study and application of Maxwell distribution have been received great attention. In 1989, Tyagi and Bhattacharya [1] firstly used Maxwell distribution to model products’ life, and they obtained the minimum variance unbiased estimator and Bayes estimator of the parameter and reliability. Chaturvedi and Rani [2] studied Bayesian reliability estimation of the generalized Maxwell failure distribution. Podder and Roy [3] studied the estimation of the parameter of this distribution under modified linear exponential loss (MLINEX) function. Bekker and Roux [4] discussed the maximum likelihood estimator (MLE), Bayes estimators of the truncated first moment and hazard function of the Maxwell distribution. Dey and Maiti [5] derived Bayes estimators of Maxwell distribution by considering non-informative and conjugate prior distributions under three loss functions, namely, quadratic loss function, squared-log error loss function and MLINEX function. The references [6-8] studied the reliability estimation of Maxwell distribution based on Type-II censored sample, progressively Type-II censored sample and random censored sample, respectively. Reference [9] obtained some Bayes estimators under quadratic loss function using non-informative prior, represented by Jefferys prior and Informative priors as Gumbel Type II and Conjugate (Inverted Gamma and Inverted Levy) priors. More details about Maxwell distribution can be found in references [10-12].

This paper is devoted to the minimax estimation problem of the unknown scale parameter θ in the Maxwell distribution with the following probability density function (pdf) (Dey and Maiti [5]):

\[ f(x; \theta) = \frac{\sqrt{2}}{\sqrt{\pi}} \theta^2 x^2 e^{-\theta x}, \quad x > 0, \theta > 0 \]  

(1)

The minimax estimation is an important part in the statistical estimation area, which was introduced by Abraham Wald [13]. The minimax estimation theory has drawn great attention by many researchers. Roy et al. [14] studied the minimax estimation of the scale parameter of the Weibull distribution for quadratic and MLINEX loss functions. Podder et al. [15] derived the minimax estimator of the parameter of the Pareto distribution under Quaratic and MLINEX loss functions. Dey et al. [16] discussed the minimax estimator of the parameter of Rayleigh distribution. Shadrokh and Pazira [17] studied the minimax estimator of the parameter of the minimax distribution under several loss functions.

The purpose of this paper is to study maximum likelihood estimation (MLE) and Bayes estimation of the parameter of Maxwell distribution. Further, by using Lehmann’s theorem we derive minimax estimators under three loss functions, namely, weighted squared error loss, squared log error loss and entropy loss functions.
2. Preliminary Knowledge

2.1. Maximum Likelihood Estimation

Let \( X = (X_1, X_2, \ldots, X_n) \) be a sequence of independent and identically distributed (i. i. d.) random variables of Maxwell distribution with pdf \((1)\), and \( x=(x_1, x_2, \ldots, x_n) \) is the observation of \( X \). The likelihood function of \( \theta \) for the given sample observation is

\[
l(\theta; x) = \left( \frac{2}{\pi} \right)^{\frac{n}{2}} \cdot \theta^{\frac{3n}{2}} \cdot \exp\left( \ln | \right) \cdot e^{-\frac{\theta}{2} \sum_{i=1}^{n} x_i^2} \tag{2}
\]

That is

\[
l(\theta; x) \propto \theta^{\frac{3n}{2}} e^{-\frac{1}{2} \sum_{i=1}^{n} x_i^2 \theta} \propto \theta^{\frac{3n}{2}} e^{-\frac{\theta}{2} t} \tag{3}
\]

where \( t = \sum_{i=1}^{n} x_i^2 \) is the observation of \( T = \sum_{i=1}^{n} X_i^2 \).

By solving log likelihood equation, the maximum likelihood estimator of \( \theta \) is easily derived as follows:

\[
\hat{\theta}_M = \frac{3n}{t} \tag{4}
\]

And by Eq. (1), we can easily show that \( T \) is distributed as a Gamma distribution \( \Gamma\left(\frac{3}{2}, \frac{\theta}{2}\right) \), which has the following probability density function ([5], p. 283):

\[
f_t(t; \theta) = \left( \frac{\theta}{2} \right)^{\frac{3}{2}-1} t^{\frac{3}{2}-1} e^{-\frac{\theta}{2} t}, \quad t > 0, \theta > 0 \tag{5}
\]

Here, \( \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \) is the Gamma function.

2.2. Loss Function

Loss function plays an important role in Bayesian analysis and the most common loss are symmetric loss function, especially squared error loss function are considered most. Under squared loss function, it is to be thought the overestimation and underestimation have the same estimated risks. However, in many practical problems, overestimation and underestimation will have different consequences. To overcome this difficulty, Zellner [18] proposed an asymmetric loss function known as the LINEX loss function. Brown proposed a new asymmetric loss function for scale parameter estimation in 1968, which is called squared log error loss (Kiapoura and Nematollahib [19]):

\[
L(\theta, \delta) = (\ln \delta - \ln \theta)^2 \tag{6}
\]

Which is balanced and \( L(\theta, \delta) \to \infty \) as \( \delta \to 0 \) or \( \infty \). The risk function has minimum at the following Bayes estimator:

\[
\hat{\delta}_M = \exp(E(\ln \theta | X)) \tag{7}
\]

In many practical situations, it appears to be more realistic to express the loss in terms of the ratio \( \delta / \theta \). In this case, Dey et al. [20] pointed out that a useful asymmetric loss function is entropy loss function:

\[
L(\hat{\theta}, \theta) = \frac{\delta}{\theta} - \ln \frac{\delta}{\theta} - 1 \tag{8}
\]

The Bayes estimator under the entropy loss is denoted by \( \hat{\theta}_{BE} \), given by

\[
\hat{\theta}_{BE} = [E(\theta^{-1} | X)]^{-1} \tag{9}
\]

3. Bayesian Estimation

In this section, we estimate \( \theta \) by considering weighted square error loss, squared log error loss and entropy loss functions.

We further assume that some prior knowledge about the parameter \( \theta \) is available to the investigation from past experience with the Maxwell model. The prior knowledge can often be summarized in terms of the so-called prior densities on parameter space of \( \theta \). In the following discussion, we assume the following Jeffrey’s non-informative quasi-prior density defined as,

\[
\pi(\theta) \propto \frac{1}{\theta^d}, \quad \theta > 0 \tag{10}
\]

Hence, \( d = 0 \) leads to a diffuse prior and \( d = 1 \) to a non-informative prior.

Combining the likelihood function (3) and the prior density (10), we obtain the posterior density of \( \theta \) is

\[
h(\theta | x) = \frac{\left(\frac{\theta}{2}\right)^{\frac{3}{2}-d+1} t^{-\frac{3}{2}-d+1} e^{-\frac{\theta}{2} t}}{\Gamma\left(\frac{3}{2}, n-d+1\right)} \tag{11}
\]

This is a Gamma distribution \( \Gamma\left(\frac{3}{2}, n-d+1, \frac{L}{2}\right) \).

Theorem 1. Let \( X = (X_1, X_2, \ldots, X_n) \) be a sample of Maxwell distribution with pdf \((1)\), and \( x=(x_1, x_2, \ldots, x_n) \) is the observation of \( X \). \( t = \sum_{i=1}^{n} x_i^2 \) is the observation of \( T = \sum_{i=1}^{n} X_i^2 \).

Then

(i) Under the weighted square error loss function

\[
L_1(\theta, \delta) = (\delta - \theta)^2 \tag{12}
\]
the Bayes estimator is given by

\[
\hat{\delta}_{BS} = \frac{E[\theta^{-1} | X]}{E[\theta^{-2} | X]} = \frac{3n-2d-2}{T}
\]

(iii) The Bayes estimator under the squared log error loss function is come out to be

\[
\hat{\delta}_{SL} = \exp\left[\ln\theta \mid X\right] = \frac{e^{\Psi_{\frac{n-d-1}{2}}}}{T / 2}
\]

Where

\[
\Psi(n) = \frac{d}{dn} \ln \Gamma(n) = \int_0^{\infty} \ln y \cdot y^{n-1} e^{-y} dy / \Gamma(n)
\]

is a Digamma function.

The Bayes estimator under the entropy loss function is obtained as

\[
\hat{\delta}_{BE} = [E(\theta^{-1} | X)]^{-1} = \frac{n-d}{n / d}
\]

Proof.

(i) By formula (11), we know that

\[
\theta \mid X \sim \Gamma\left(\frac{3}{2}, n-d + 1, \frac{T}{2}\right)
\]

Then

\[
E[\theta^{-1} | X] = \frac{\frac{T}{2}}{\frac{3}{2} n-d}
\]

\[
E[\theta^{-2} | X] = \frac{\left(\frac{T}{2}\right)^2}{\left(\frac{3}{2} n-d\right)\left(\frac{3}{2} n-d-1\right)}
\]

Thus, the Bayes estimator under the weighted square error loss function is given by

\[
\hat{\delta}_{BS} = \frac{E[\theta^{-1} | X]}{E[\theta^{-2} | X]} = \frac{\frac{3}{2} n-d-1}{\frac{T}{2}} = \frac{3n-2d-2}{T}
\]

For the case

(ii) By using (11),

\[
E(\ln \theta \mid X) = \frac{\Psi_{\frac{n-d-1}{2}}}{\Gamma\left(\frac{3}{2} n-d+1\right)} \int_0^{\infty} \ln \theta \cdot \theta^{\frac{3}{2} n-d-1+1} e^{-\theta T/2} d\theta
\]

\[
= \frac{d}{dn} \ln \Gamma\left(\frac{3}{2} n-d+1\right) - \ln(T / 2) = \Psi\left(\frac{3}{2} n-d+1\right) - \ln(T / 2)
\]

Then the Bayes estimator under the squared log error loss function is come out to be

\[
\hat{\delta}_{SL} = \exp[\ln\theta \mid X] = \frac{e^{\Psi_{\frac{n-d-1}{2}}}}{T / 2}
\]

(iii) By Eqs. (10) and (17), the Bayes estimator under the entropy loss function is given by

\[
\hat{\delta}_{BE} = [E(\theta^{-1} | X)]^{-1} = \frac{3}{2} n-d = \frac{3n-2d}{T}
\]

4. Minimax Estimation

The most important elements in the minimax approach are the specification of the prior distribution and the loss functions by using a Bayesian method. The derivation of minimax estimators depends primarily on a theorem due to Lehmann which can be stated as follows:

Lemma 1 (Lehmann’s Theorem) If \( \tau = \{\tau, \theta \in \Theta\} \) be a family of distribution functions and \( D \) a class of estimators of \( \theta \). Suppose that \( \delta^* \in D \) is a Bayes estimator against a prior distribution \( \delta^* (\theta) \) on \( \Theta \), and the risk function \( R(\delta^*, \theta) \) equals constant on \( \Theta \); then \( \delta^* \) is a minimax estimator of \( \theta \).

Theorem 2 Let \( X = (X_1, X_2, \ldots, X_n) \) be a sample of Maxwell distribution with probability density function (1), then

(i) \( \hat{\delta}_{BS} = \frac{3n-2d-2}{T} \) is the minimax estimator of the parameter \( \theta \) for the weighted square error loss function.

(ii) \( \hat{\delta}_{SL} = \frac{e^{\Psi_{\frac{n-d-1}{2}}}}{T / 2} \) is the minimax estimator of the parameter \( \theta \) for the squared log error loss function.

(iii) \( \hat{\delta}_{BE} = \frac{3n-2d}{T} \) is the minimax estimator of the parameter \( \theta \) for the entropy loss function.

Proof. To prove the theorem we shall use Lehmann’s theorem, which has been stated before. Then we have to find the Bayes estimator \( \delta^* \) of \( \theta \). Thus if we can show that the risk of \( \delta^* \) is constant, then the theorem 2 will be proved.

(i) The risk function of the estimator \( \hat{\delta}_{BS} \) is

\[
R(\theta) = \int \left[ \frac{3n-2d-2}{T} \right] \left[ 3n-2d-2 \right] \left( T^{-1} \right) + \theta^2
\]

From the conclusion \( T \sim \Gamma\left(\frac{3}{2} n, \frac{\theta}{2}\right) \), we have
\[ E(T^{-1}) = \frac{\theta}{3n-2}, \]
\[ E(T^{-1}) = \frac{\theta^2}{(3n-2)(3n-4)} \]

Then
\[ R(\theta) = \frac{1}{\theta^2} \left[ (3n-2d-2)^2 \frac{\theta^2}{(3n-2)(3n-4)} - 2\theta(3n-2d-2) \theta + \theta^2 \right] \]
\[ = \frac{6n-4-8d+4d^2}{(3n-2)(3n-4)} \]

Then \( R(\theta) \) is a constant. So, according to the Lehmann’s theorem it follows that, \( \hat{\delta}_{BS} \) is the minimax estimator for the parameter \( \theta \) of the Maxwell distribution under the weighted square error loss function.

Now we are going to prove the case (ii) The risk function of the estimator \( \hat{\delta}_{SL} \) is
\[ R(\theta) = E \left[ L(\theta, \hat{\delta}_{SL}) \right] = E \left[ (\ln \hat{\delta}_{SL} - \ln \theta)^2 \right] \]
\[ = E[\ln \hat{\delta}_{SL}]^2 - 2 \ln \theta \cdot E[\ln \hat{\delta}_{SL}] + (\ln \theta)^2 \]

From the conclusion \( T \sim \Gamma \left( \frac{3}{2}, \frac{n}{\theta} \right) \), we have
\[ E(\ln T) = \Psi \left( \frac{3}{2}, n \right) - \ln \theta + \ln 2 \]

Thus
\[ E[\ln \hat{\delta}_{SL}] = E\left[ \Psi \left( \frac{3}{2}, n \right) - \ln T + \ln 2 \right] \]
\[ = \Psi \left( \frac{3}{2}, n \right) - \Psi \left( \frac{3}{2}, \frac{n}{\theta} \right) + \ln \theta \]
\[ = \Psi \left( \frac{3}{2}, n \right) - \Psi \left( \frac{3}{2}, n \right) - \ln(T / 2)^2 \]
\[ = \Psi \left( \frac{3}{2}, n \right) - \Psi \left( \frac{3}{2}, n \right) - \ln(T / 2)^2 \]
\[ = \Psi \left( \frac{3}{2}, n \right) - \Psi \left( \frac{3}{2}, n \right) + \ln(T / 2)^2 \]

Using the fact
\[ \Psi(n) = \int_0^y (\ln y)^{\alpha-1} e^{-y} \frac{dy}{\Gamma(n)} \int_0^y (\ln y)^{\alpha-1} e^{-y} \Psi(n)dy \]
\[ = E(\ln Y)^2 - \Psi(n) \]

where \( Y \sim \Gamma(n, 1) \). and we can show that
\[ U = \frac{T \theta}{2} \sim \Gamma \left( \frac{3}{2}, n, 1 \right) \]

Thus
\[ \Psi \left( \frac{3}{2}, n \right) + \Psi \left( \frac{3}{2}, \frac{n}{\theta} \right) = E[\ln U]^2 = E[\ln(T / 2) + \ln \theta]^2 \]
\[ = E[\ln(T / 2)^2] + 2 \ln \theta \cdot E[\ln(T / 2)] + (\ln \theta)^2 \]
\[ = E[\ln(T / 2)^2] + 2 \ln \theta \cdot (\Psi \left( \frac{3}{2}, n \right) - \ln \theta + (\ln \theta)^2 \]

Then we get the fact
\[ E[\ln(T / 2)^2] = \Psi \left( \frac{3}{2}, n \right) + \Psi \left( \frac{3}{2}, \frac{n}{\theta} \right) - 2 \ln \theta \cdot \Psi \left( \frac{3}{2}, n \right) + (\ln \theta)^2 \]

Therefore
\[ E[\ln \hat{\delta}_{SL}] = \Psi \left( \frac{3}{2}, n \right) + \Psi \left( \frac{3}{2}, \frac{n}{\theta} \right) - 2 \ln \theta \cdot \Psi \left( \frac{3}{2}, n \right) + (\ln \theta)^2 \]

Then \( R(\theta) \) is a constant. So, according to the Lehmann’s theorem it follows that, \( \hat{\delta}_{BS} \) is the minimax estimator for the parameter \( \theta \) of the Maxwell distribution under the weighted log error loss function.

Finally we are going to prove the case (iii) The risk function of the estimator \( \hat{\delta}_{BE} \) is
\[ R(\theta) = E \left[ L(\theta, \hat{\delta}_{BE}) \right] = E \left[ (\ln \hat{\delta}_{BE} - \ln \theta)^2 \right] \]
\[ = \Psi \left( \frac{3}{2}, n \right) - \Psi \left( \frac{3}{2}, \frac{n}{\theta} \right) + \ln \theta \]
\[ = \Psi \left( \frac{3}{2}, n \right) - \Psi \left( \frac{3}{2}, n \right) - \ln(T / 2)^2 \]
\[ = \Psi \left( \frac{3}{2}, n \right) - \Psi \left( \frac{3}{2}, n \right) + \ln(T / 2)^2 \]

Using the fact
\[ \Psi(n) = \int_0^y (\ln y)^{\alpha-1} e^{-y} \frac{dy}{\Gamma(n)} \int_0^y (\ln y)^{\alpha-1} e^{-y} \Psi(n)dy \]
\[ = E(\ln Y)^2 - \Psi(n) \]

where \( Y \sim \Gamma(n, 1) \). and we can show that
\[ U = \frac{T \theta}{2} \sim \Gamma \left( \frac{3}{2}, n, 1 \right) \]

Therefore
\[ E[\ln \hat{\delta}_{BE}] = \Psi \left( \frac{3}{2}, n \right) - \Psi \left( \frac{3}{2}, \frac{n}{\theta} \right) + \ln \theta \]

Then \( R(\theta) \) is a constant. So, according to the Lehmann’s theorem it follows that, \( \hat{\delta}_{BE} \) is the minimax estimator for the parameter \( \theta \) of the Maxwell distribution under the entropy loss function.
5. Risk Function

The risk functions of the estimators $\hat{\delta}_{BS}$, $\hat{\delta}_{BL}$ and $\hat{\delta}_{BE}$ relative to squared error loss $L(\theta, \hat{\theta}) = (\delta - \theta)^2$ are denoted by $R(\hat{\delta}_{BS})$, $R(\hat{\delta}_{BL})$ and $R(\hat{\delta}_{BE})$, respectively, and can be easily shown as

$$R(\hat{\delta}_{BS}) = E[L(\hat{\delta}_{BS}, \theta)] = E[(3n - 2d - 2\theta)^2 + \theta^2] f(t) dt$$

$$R(\hat{\delta}_{BL}) = E[L(\hat{\delta}_{BL}, \theta)] = E[\frac{3n - 2d - 2\theta}{t^2} + \theta^2]$$

$$R(\hat{\delta}_{BE}) = E[L(\hat{\delta}_{BE}, \theta)] = E[(\frac{3n - 2d}{t^2} - \theta)^2]$$

The Figs. 1-4 have plotted the ratio of the risk functions to $\theta^2$, i.e.

$$\frac{R(\hat{\delta}_{BS})}{\theta^2} = L_1, \quad \frac{R(\hat{\delta}_{BL})}{\theta^2} = L_2, \quad \text{and} \quad \frac{R(\hat{\delta}_{BE})}{\theta^2} = L_3$$

From Figs. 1-4, it is clear that no of the estimators uniformly dominates any other. We therefore recommend choose the estimators according to the value of $d$ when the quasi-prior density is used as the prior distribution.
6. Conclusion

In this paper, we have derived Bayes estimators of the parameter of Maxwell distribution under weighted squared error loss, squared log error loss and entropy loss functions. Simulation results show that the risk functions of the estimators \( \hat{\delta}_{BS}, \hat{\delta}_{BL} \) and \( \hat{\delta}_{BE} \) relative to squared error loss decrease as sample size increases. When the sample size n is large, such as n>50, the risks are almost the same.

Acknowledgement

This study is partially supported by Natural Science Foundation of Hunan Province (No. 2015JJ3030) and Foundation of Hunan Educational Committee (No. 15C0228). The author also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References