Bayes and Non-Bayes Estimation Methods for the Parameter of Maxwell-Boltzmann Distribution

Iden H. Alkanani 1

Shayma G. Salman 2*

Received 8/6/2017
Accepted 9/7/2017

Abstract:
In this paper, point estimation for parameter $\theta$ of Maxwell-Boltzmann distribution has been investigated by using simulation technique, to estimate the parameter by two sections methods; the first section includes Non-Bayesian estimation methods, such as (Maximum Likelihood estimator method, and Moment estimator method), while the second section includes standard Bayesian estimation method, using two different priors (Inverse Chi-Square and Jeffrey) such as (standard Bayes estimator, and Bayes estimator based on Jeffrey's prior). Comparisons among these methods were made by employing mean square error measure. Simulation technique for different sample sizes has been used to compare between these methods.

Keywords: Maxwell-Boltzmann distribution, Maximum Likelihood method, Moment method, Standard Bayes method, Standard Bayes by Jeffrey's prior method, Mean square error, Simulation technique.

Introduction:
The statistical mechanic deals with Maxwell-Boltzmann distribution which describes the energy and velocity in gas, when the molecules motion freely between the level of collisions, and do not interact with each other, because of employment of the temperature of gas system. In the statistical mechanic, the Maxwell-Boltzmann distribution grants the velocity and energy in gas.

In the statistical mechanic, the Maxwell-Boltzmann distribution grants the velocity and energy molecules in thermal equilibrium. In (1989) the Tyagi and Bhattacherya introduced the Maxwell distribution in lifetime model [1]. In (1998) the Chaturvedi and Rani used classical and Bayesian method of the generalized Maxwell distribution in lifetime to find the reliability estimation function [2]. In (2005) Bekker and Roux made the characteristic of reliability function to Maxwell distribution and studied Bayes estimation [3]. In (2009) Krishna and Malik estimated the reliability function to Maxwell distribution by using type-two censored sample [4].

In (2011) Kasmi, and others utilized the maximum likelihood estimator in type one censored sample of mixture Maxwell distribution [5]. In (2012) Kasmi and others utilized the Bayesian estimation for two components of Maxwell distribution by using type-one censored sample [6]. In (2013) Al-Baldawi compared between some Bayesian estimators with maximum likelihood estimator for Maxwell distribution used non-informative priors [7]. In (2016) Rasheed and Khalifa estimated the parameter of Maxwell distribution by using Bayes estimator under quadratic loss function using Non-informative prior [8].

The aim of this paper is to study four methods to find and estimate the parameter of Maxwell-Boltzmann distribution: the first is maximum likelihood method, the second is the moment method, the third is standard Bayesian method, the fourth is Jeffrey's prior method, in addition to comparing between them using Mean Square Error (MSE) utilizing Monte Carlo simulation technique with various sample sizes.

Maxwell-Boltzmann Distribution [9].

Let $x$ be a random variable of Maxwell-Boltzmann distribution with one parameter, then the cumulative distribution function (cdf) of it is given:


\[ F(x; a) = \frac{2}{\sqrt{\pi}} y \left( \frac{3}{2} \frac{x^2}{a^2} \right) \]

Where \( y \left( \frac{3}{2} \frac{x^2}{a^2} \right) \) is the lower incomplete Gamma function defined by

\[ y(a, t) = \int_0^t y^{-a-1} e^{-y} dy \quad (a, t > 0). \]

and \( a \) is called scale parameter.

The probability density function (pdf) for the Maxwell-Boltzmann distribution is defined as follows:

\[
f(x; a) = \left\{ \begin{array}{ll}
\frac{2 x^2 e^{-x^2/2a^2}}{\pi a^3} & x \geq 0, a > 0 \\
0 & o.w.
\end{array} \right.
\]

let \( a^2 = \theta \) then \( f(x; a) \) became as blow

\[
f(x; \theta) = \left\{ \begin{array}{ll}
\frac{2 x^2 e^{-x^2/2\theta}}{\pi \theta^{3/2}} & x \geq 0, \theta > 0 \\
0 & o.w.
\end{array} \right.
\]

The mean and variance for this distribution define as follows:

\[
\mu_x = 2a \sqrt{\frac{2}{\pi}}, \quad \sigma_x^2 = \frac{a^2(3\pi-8)}{\pi}.
\]

The moment generating function (mgf) of this distribution is:

\[
M_x(t) = \frac{e^{a^2 t^2/2}}{\sqrt{\pi}} \left[ a^2 t^2 \Gamma \left( \frac{1}{2}, \frac{a^2 t^2}{2} \right) + 2 \Gamma \left( \frac{3}{2}, \frac{a^2 t^2}{2} \right) \right] + 2 \sqrt{\frac{2}{\pi}} at
\]

Estimation Methods:

This study has been dealing with three estimation methods as follows:

Maximum Likelihood Estimator Method (MLEM).

The likelihood function for Maxwell-Boltzmann distribution is as follows [8]:

\[
L = \prod_{i=1}^{n} \frac{\sqrt{2}}{\sqrt{\pi} \theta^{3/2}} e^{-x_i^2/2\theta}
\]

\[
L = \left( \frac{2}{\sqrt{\pi}} \right)^n \left( \frac{\sqrt{2}}{\sqrt{\pi} \theta^{3/2}} \right)^n \prod_{i=1}^{n} e^{-x_i^2/2\theta}
\]

Taking the natural logarithm for both sides, we get

\[
\ln L = \frac{n}{2} \ln 2 - \frac{n}{2} \ln \pi + 2n \ln x_i - \frac{3}{2} n \ln \theta - \frac{\sum_{i=1}^{n} x_i^2}{2\theta}
\]

The partial derivative for log-likelihood function with respect to parameter \( \theta \), and setting the results equal to zero, we obtain

\[
\frac{\partial \ln L}{\partial \theta} = -3n + \frac{\sum_{i=1}^{n} x_i^2}{2\theta^2}
\]

\[
\frac{\partial \ln L}{\partial \theta} = 0 \quad -3n + \frac{\sum_{i=1}^{n} x_i^2}{2\theta^2} = 0, \quad \sum_{i=1}^{n} x_i^2 - 3n\theta = 0
\]

\[
\hat{\theta} = \frac{\sum_{i=1}^{n} x_i^2}{3n} \quad \text{is MLE for } \theta
\]

Moment Estimator Method (MEM).

The population and sample moments for parameter of Maxwell-Boltzmann distribution are given by

\[
\hat{\mu} = 2\sqrt{\theta} \sqrt{\frac{2}{\pi}}, \quad \mu = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}
\]

Then, \( \hat{\mu} = \mu \), \( \frac{2\sqrt{\theta} \sqrt{2}}{\sqrt{\pi}} = \bar{x} \)

\[
\frac{3}{2} \sqrt{\theta} = \bar{x} \sqrt{\pi}, \quad \sqrt{\theta} = \frac{\bar{x} \sqrt{\pi}}{(2)^{3/2}}
\]

\[
\hat{\theta} = \frac{-\pi x^2}{8} \quad \text{is ME for } \theta
\]

Standard Bayes Estimator Method (SBEM) [10].

The standard Bayesian estimator method is hypothesized that unknown parameter \( \theta \) is random variable, in some real situation, there is often additional information available about \( \theta \) (that means there is a prior knowledge exists about the parameter \( \theta \).

This method is based on the notation \( g(\theta) \) that represents the prior distribution for parameter \( \theta \) which came from prior knowledge, additional information, past experience. The Bayes estimator depends on the probability density function (posterior pdf) which includes information from previous knowledge and sample information [10].

Here we used two priors distributions as follows:

Invers Chi-Square Prior.

The prior density function for \( \theta \) where \( \theta \) is scale parameter is:

\[
\theta \sim \text{Inver-chi-square}(\nu)
\]

\[
g(\theta) = \begin{bmatrix} \frac{1}{2\nu} e^{-\frac{\theta}{2\nu}} & \nu, \theta > 0 \\ 0 & \text{ow} \end{bmatrix}
\]

Where \( \nu \) is a degree of freedom and \( \nu = n - 1 \)

To find the posterior distribution as follows:

\[
f(\theta; x_1, x_2, \ldots, x_n) = \frac{1}{\prod_{i=1}^{n} f(x_1, x_2, \ldots, x_n; \theta) g(\theta)}
\]

\[
= \frac{1}{\int_0^{\infty} \prod_{i=1}^{n} f(x_1, x_2, \ldots, x_n; \theta) g(\theta) d\theta}
\]
By using squared error loss function \( L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2 \), the risk function is:
\[
R(\theta) = \int_0^\infty L(\hat{\theta}, \theta) f(\theta; x_i) d\theta
\]

by using squared error loss function \( L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2 \), the risk function is
\[
R(\theta) = \int_0^\infty L(\hat{\theta}, \theta) f(\theta; x_i) d\theta
\]

The partial derivative for \( R(\theta) \) with respect to \( \hat{\theta} \) we get
\[
\frac{\partial R(\theta)}{\partial \hat{\theta}} = 2\hat{\theta} - \frac{\sum_{i=1}^n x_i^2 + 1}{\frac{3}{2n + v - 1}} + \text{zero}
\]
the Bayes estimator \( \hat{\theta} \) is the solution of equation
\[
2\hat{\theta} - \frac{\sum_{i=1}^n x_i^2 + 1}{\frac{3}{2n + v - 1}} = 0
\]
\[
\hat{\theta} = \frac{\sum_{i=1}^n x_i^2 + 1}{3n + v - 2} \quad \text{is SBE for } \theta
\]

3.3.2 Jeffrey's Prior [11].
Consider Jeffrey's prior as \( g(\theta) \propto \sqrt{I(\theta)} \) where \( I(\theta) \) is called fisher's information and
\[
I(\theta) = -n E \left( \frac{\partial^2 \ln f(x; \theta)}{\partial \theta^2} \right) = \frac{3n}{2\theta^2}
\]
\[
g(\theta) = \frac{\frac{3n}{\sqrt{2\theta^2}}}{\sqrt{2\theta^2}}
\]
The posterior pdf of \( \theta \) has the following from.
\[
f(\theta; x_1, x_2, \ldots, x_n) = \frac{\left( \frac{\sqrt{3}}{\sqrt{2}} \right)^n \theta^{-\frac{1}{2}n} \prod_{i=1}^n x_i^2 e^{-\frac{\sum_{i=1}^n x_i^2}{2\theta}}}{\prod_{i=1}^n 2^n \frac{1}{\sqrt{2\theta}}} \int_0^{\infty} \left( \frac{\sqrt{3}}{\sqrt{2}} \right)^n \theta^{-\frac{1}{2}n} \prod_{i=1}^n x_i^2 e^{-\frac{\sum_{i=1}^n x_i^2}{2\theta}} d\theta
\]
\[
f(\theta; x_1, x_2, \ldots, x_n) = \frac{\prod_{i=1}^n \sqrt{2} \frac{\sqrt{3}}{\sqrt{2\theta}} x_i^2 e^{-\frac{\sum_{i=1}^n x_i^2}{2\theta}}}{\prod_{i=1}^n 2^n \frac{1}{\sqrt{2\theta}}} d\theta
\]
\[
f(\theta; x_1, x_2, \ldots, x_n) \quad \frac{\sum_{i=1}^n x_i^2}{2} \prod_{i=1}^n \frac{3^n}{\sqrt{2\theta}} \theta^{-\frac{1}{2}n} \prod_{i=1}^n x_i^2 e^{-\frac{\sum_{i=1}^n x_i^2}{2\theta}} d\theta
\]
\[
= \frac{\left( \frac{\sqrt{3}}{\sqrt{2}} \right)^n \theta^{-\frac{1}{2}n} \prod_{i=1}^n x_i^2 e^{-\frac{\sum_{i=1}^n x_i^2}{2\theta}}}{\prod_{i=1}^n 2^n \frac{1}{\sqrt{2\theta}}} \int_0^{\infty} \left( \frac{\sqrt{3}}{\sqrt{2}} \right)^n \theta^{-\frac{1}{2}n} \prod_{i=1}^n x_i^2 e^{-\frac{\sum_{i=1}^n x_i^2}{2\theta}} d\theta
\]
b-Obtain two standard normal variates $Y_1$ and $Y_2$ using the transformation

$$ Y_1 = \sqrt{-2 \log(X_1) \cos 2\pi(X_2)}, \quad Y_2 = \sqrt{-2 \log(X_1) \sin 2\pi(X_2)}, $$

and find $Z = \frac{Y_1 + Y_2}{\sqrt{2}}$ which is $N(0,1)$.

c- Repeating steps 1 and 2 three times generate a chi-square $X_3^2$ variate using

$$ T = \sum_{i=1}^{3} x_i^2 $$

which is gamma $G\left(\frac{3}{2}, \frac{1}{\theta}\right)$ variate.

d- Using the transformation $V = \frac{\theta}{\sqrt{\pi}}$, we get a number generated from Maxwell variate.

Second: To generate $x$ which distributed as Maxwell-Boltzmann function, we must choose many values to the parameter $\theta$ as well as we must choose many various samples sizes which are as follows:

$\theta = 0.5, 1.0, 1.5, 2.0, 2.5$

$n = 10, 50, 100, 200$

And we replicate the data of experiment (500) times, then the number of all generating experiment is (5) times.

Third: By MATLAB program, we have got the following estimated values for the scale parameter to the Maxwell-Boltzmann distribution and numerical results scheduled in Table (1)

Fourth: Computing the Mean Squares Error measure (MSE) for all situations studied in this paper, and scheduled in the Table (2).

| Table 1. The estimate values for parameter $\theta$ in several methods When $n = 10$ |
|---|---|---|---|---|
| $\theta$ | $\hat{\theta}_{MLE}$ | $\hat{\theta}_{ME}$ | $\hat{\theta}_{SBE}$ | $\hat{\theta}_{JE}$ |
| 0.5 | 0.5046 | 0.5127 | 0.4632 | 0.5407 |
| 1.0 | 1.0148 | 1.0362 | 0.8769 | 1.0873 |
| 1.5 | 1.5168 | 1.5442 | 1.2839 | 1.6252 |
| 2.0 | 1.9906 | 2.0050 | 1.6680 | 2.1327 |
| 2.5 | 2.4615 | 2.5143 | 2.0499 | 2.6374 |

| Table 2. The Mean Square Error for parameter $\theta$ in several methods When $n = 200$ |
|---|---|---|---|---|---|
| $\theta$ | $\hat{\theta}_{MLE}$ | $\hat{\theta}_{ME}$ | $\hat{\theta}_{SBE}$ | $\hat{\theta}_{JE}$ | Best |
| 0.5 | 0.4987 | 0.4999 | 0.3780 | 0.5004 | SBE |
| 1.0 | 0.9997 | 1.0001 | 0.7551 | 1.0030 | MLE |
| 1.5 | 1.5010 | 1.5027 | 1.1325 | 1.5061 | MLE |
| 2.0 | 2.0034 | 2.0042 | 1.5107 | 2.0101 | MLE |
| 2.5 | 2.5022 | 2.5035 | 1.8862 | 2.5106 | MLE |

From Table (1) we can make the following comments:

1-Noting that estimated values of parameter $\theta$ are vibrating and have (31) times smaller values compared with true values of $\theta$, and have (49) times bigger values compared with true values of $\theta$ for all samples sizes of the four studied methods.

2-Showing that the estimated values of parameter $\theta$ are converged to the true values of $\theta$ (20) times in Maximum Likelihood method, moment method, and Jeffery's prior method, and (5) times in standard Bayes method.

| Table 2. The Mean Square Error for parameter $\theta$ in several methods When $n = 10$ |
|---|---|---|---|---|---|
| $\theta$ | $\hat{\theta}_{MLE}$ | $\hat{\theta}_{ME}$ | $\hat{\theta}_{SBE}$ | $\hat{\theta}_{JE}$ | Best |
| 0.5 | 0.0179 | 0.0198 | 0.0131 | 0.0221 | SBE |
| 1.0 | 0.0673 | 0.0765 | 0.0592 | 0.0846 | SBE |
| 1.5 | 0.1789 | 0.1937 | 0.1641 | 0.2207 | SBE |
| 2.0 | 0.2708 | 0.2942 | 0.2882 | 0.3284 | MLE |
| 2.5 | 0.4135 | 0.4684 | 0.4735 | 0.4919 | MLE |

| Table 2. The Mean Square Error for parameter $\theta$ in several methods When $n = 50$ |
|---|---|---|---|---|---|
| $\theta$ | $\hat{\theta}_{MLE}$ | $\hat{\theta}_{ME}$ | $\hat{\theta}_{SBE}$ | $\hat{\theta}_{JE}$ | Best |
| 0.5 | 0.0034 | 0.0037 | 0.0141 | 0.0035 | MLE |
| 1.0 | 0.0133 | 0.0142 | 0.0596 | 0.0139 | MLE |
| 1.5 | 0.0319 | 0.0365 | 0.1371 | 0.0334 | MLE |
| 2.0 | 0.0525 | 0.0573 | 0.2555 | 0.0542 | MLE |
| 2.5 | 0.0798 | 0.0863 | 0.3873 | 0.0833 | MLE |

| Table 2. The Mean Square Error for parameter $\theta$ in several methods When $n = 100$ |
|---|---|---|---|---|---|
| $\theta$ | $\hat{\theta}_{MLE}$ | $\hat{\theta}_{ME}$ | $\hat{\theta}_{SBE}$ | $\hat{\theta}_{JE}$ | Best |
| 0.5 | 0.0014 | 0.0015 | 0.0147 | 0.0015 | MLE |
| 1.0 | 0.0060 | 0.0066 | 0.0605 | 0.0062 | MLE |
| 1.5 | 0.0167 | 0.0181 | 0.1367 | 0.0172 | MLE |
| 2.0 | 0.0256 | 0.0277 | 0.2431 | 0.0264 | MLE |
| 2.5 | 0.0418 | 0.0443 | 0.3952 | 0.0425 | MLE |

| Table 2. The Mean Square Error for parameter $\theta$ in several methods When $n = 200$ |
|---|---|---|---|---|---|
| $\theta$ | $\hat{\theta}_{MLE}$ | $\hat{\theta}_{ME}$ | $\hat{\theta}_{SBE}$ | $\hat{\theta}_{JE}$ | Best |
| 0.5 | 0.0008 | 0.0009 | 0.0154 | 0.0008 | MLE,JE |
| 1.0 | 0.0030 | 0.0031 | 0.0617 | 0.0030 | MLE,JE |
| 1.5 | 0.0079 | 0.0081 | 0.1395 | 0.0080 | MLE |
| 2.0 | 0.0131 | 0.0138 | 0.2468 | 0.0133 | MLE |
| 2.5 | 0.0209 | 0.0226 | 0.3886 | 0.0211 | MLE |

From Table 2. we can make the following comments:
The values of mean squares error for $\hat{\theta}$ are decreasing when the samples sizes are increasing for all values of $\theta$ in all methods, except standard Bayes estimators of $\theta$ that are vibrating.

2-Noting that the values of MSE are vibrating for all increasing values of $\theta$. The smallest values of MSE are ($0.0008$) when ($\theta = 0.5, n = 200$) for Maximum likelihood estimator method and Jeffry Bayes estimator.

Conclusions:
Throughout the estimator parameters for all four methods, we see that all the values of estimator parameters are close to the true values of parameters in Maxwell-Boltzmann distribution. Also we can see that the mean squared error procedure for all four methods have a smallest value, specially the maximum likelihood estimator method and far away from standard Bayesian estimator method.

References:
[1] Tyagi, R. K. and Bhattacharya, S. K. 1989. Bayes estimation of the Maxwell's velocity distribution function. Statistica, 29(4): 563-567.
[2] Chaturvedi, A. and Rani, U. 1998. Classical and Bayesian Reliability estimation of the generalized Maxwell failure distribution. Journal of Statistical Research, 32(1): 113-120.
[3] Bekker, A. and Roux, J. J. 2005. Reliability characteristics of the Maxwell distribution: A Bayes estimation study. Comm. Stat. (Theory & Math.), 34(11): 2169 - 2178.
[4] Krishna, H. and Malik, M. 2009. Reliability estimation in Maxwell distribution with Type-II censored data. Int. Journal of Quality and Reliability management, 26 (2): 184 – 195.
[5] Kasmi, A. S. M.; Aslam, M. and Ali, S. 2011. A note of Maximum likelihood estimator for the mixture of Maxwell distributions using Type-I censored scheme. The open stat. and prob. Journal, (3): 31 – 35.
[6] Kasmi, A. S. M.; Aslam, M. and Ali, S. 2012. on the Bayesian estimation for two component mixture of Maxwell distribution, assuming type I censored data. Int. J. of Applied Science and Technology, 2(1): 197-218.
[7] Al-Baldawi, T. H. K. 2013. Comparison of Maximum Likelihood and some Bayes Estimators for Maxwell Distribution based on Non-Informative Priors. Baghdad Science Journal, 10(2): 480-488.
[8] Rasheed, H. A. and Khalifa, Z. N. 2016. Bayes Estimators for The Maxwell Distribution Under Quadratic Loss Function Using Different Priors. Australian Journal of Basic and Applied Sciences, 10(6): 97-103.
[9] Papoulis, A. 1984. Random Variables, and Stochastic Processes, 2nd Edition. New York: McGraw-Hill, pp.104 and 148-149.
[10] Soong, T.T. 2004. Fundamentals of Probability and Statistics for Engineers, John Wiley and Sons, Ltd.
[11] Sanku, D. and Tanujit, D. 2011. Rayleigh Distribution Revisited VIA Extension of Jeffrey's prior Information and a New Loss Function. Statistical Journal, 9(3): 214-226.
[12] Alkanani, I. H. and Sara, S. T. 2014. Bayesian and Non Bayesian Methods to Estimate the two Parameters of Logistic distribution. Int. J. of Applied Science and Technology, 2(1): 197-218.