Journal of Statistical Modelling: Theory and Applications
Vol. 1, No. 2, 2020, pp. 99-113
Yazd University Press 2020

JSMTA

Research Paper

Improved maximum likelihood estimation of parameters in the Maxwell distribution

Mohammad Mahdi Maghami, Mohammad Bahrami*
Department of Statistics, University of Isfahan, Isfahan, Iran

Received: October 26, 2019/ Revised: December 18, 2019 / Accepted: January 28, 2020

Abstract: Maximum likelihood estimators are usually biased. The first order bias term of the maximum likelihood estimators can be large for a small or medium sample size, and this bias may have a significant effect on distribution performance. Different methods may be used to reduce this bias. These methods have inspired many scholars to study this field over the past years, but the use of Bartlett's method requires the expected value of third power derivatives of the likelihood function. Consequently, because this quantity (the expected value of third power derivatives of the likelihood function) is not necessarily calculable in some distributions, in this paper we propose a new method based on algebraic approximation of the maximum likelihood estimator bias which needless the expected value of third power derivatives of the likelihood function. In addition, as an application of this method, we will consider a bias correction for estimating parameters of Maxwell distribution.

Keywords: Bias-corrected estimators; Bias prevention; Maximum likelihood estimator; Two-parameter Maxwell distribution.
Mathematics Subject Classification (2010): 62F15.

1 Introduction

Maxwell distribution was introduced for the first time by Maxwell (1860) and developed by Boltzman (1870). The random variable X has the Maxwell distribution with scale parameter θ which is shown by $M(\theta)$ if the probability density function (pdf) is

$$f(x;\theta) = \frac{4}{\sqrt{\pi}}\theta^{\frac{3}{2}}x^2e^{-\theta x^2}, \;\; x > 0, \, \theta > 0,$$

and the cumulative distribution function (cdf) is

$$F(x;\theta) = \frac{2}{\sqrt{\pi}}\,\Gamma\left(3/2, \theta x^2\right),$$

*Corresponding author: m.bahrami@sci.ui.ac.ir

where $\Gamma(a,x)=\int_0^x u^{a-1}e^{-u}du$ is the incomplete gamma function.

Tyagi and Bhattacharya (1989a) and Tyagi and Bhattacharya (1989b) calculated uniformly minimum-variance unbiased estimator, Bayes estimator of θ and its reliability function for this distribution. Chaturvedi and Rani (1998) introduced the generalized Maxwell distribution and calculated the classical and Bayes estimators for the parameters of this distribution. Podder and Roy (2003) calculated the parameter estimation of this distribution under the modified linear exponential loss function. Bekker and Roux (2005) calculated empirical Bayes estimator for the parameter of the Maxwell distribution. Krishna and Malik (2009) calculated ML and Bayes estimators of the reliability function under the type II censored data. Kazemi and Shahabi (2012) calculated the Bayes estimators for the parameters of a mixture distribution with the Maxwell components under type II censored data. Dey et al. (2013) studied the one-parameter Maxwell distribution and calculated the estimation of the parameter under different loss functions. In addition, recently Iriarte et al. (2017) introduced the new two-parameter gamma distribution as a generalization of the Maxwell distribution and some of its features using the G-gamma generator.

The remainder of this paper is organized as follows. In Section 2, we briefly discuss a corrective approach to derive modified MLEs. In Section 3, we discuss bias correction for one and two-parameter Maxwell distribution. In Section 4, we perform simulation studies and compare the proposed estimators in Section 3.

2 Preliminaries

2.1 Bartlett bias correction

One of the preliminary results for algebraic approximation estimate of the maximum likelihood bias was done by Bartlett (1953). He calculated the bias of MLE with order $O(n^{-1})$ of a one-parameter distribution. For Bartlett method, assume that we have a regular parametric model with $l(\theta)$ as the logarithm of likelihood function. For simplicity, we assume that the parameter θ is scalar and define $l'(\hat\theta)=\frac{\partial l(\theta)}{\partial\theta}|_{\theta=\hat\theta}$, $v(\theta)^a=(v(\theta))^a$ and $i_{rst}(\theta)=E\left[l'(\theta)^r l''(\theta)^s l'''(\theta)^t\right]$. Given the above notations, assume the regularity conditions as

$$i_{100}(\theta)=0, \qquad i_{010}(\theta)+i_{200}(\theta)=0, \qquad i_{001}(\theta)+3i_{110}(\theta)+i_{300}(\theta)=0,$$

In general, the MLE has a bias with order $O(n^{-1})$. (for example see Cox and Hinkley, 1979). Applying the Taylor's expansion for the second-order of the first derivative of $l(\theta)$ at $\theta=\hat\theta$ we have

$$l'(\hat\theta)=l'(\theta)+(\hat\theta-\theta)l''(\theta)+\frac{1}{2}(\hat\theta-\theta)^2 l'''(\theta)+O_p(n^{-2})=0,$$

where $l''(\theta)$ and $l'''(\theta)$ are the second and third derivatives of $l(\theta)$ respectively. If we obtain expected value for both sides of the above equation, we have:

$$\begin{aligned}E(\hat\theta-\theta)E(l''(\theta)) &+ cov((\hat\theta-\theta),l''(\theta))\\ &+ \frac{1}{2}E(\hat\theta-\theta)^2E(l'''(\theta))+\frac{1}{2}cov((\hat\theta-\theta)^2,l'''(\theta))\approx 0.\end{aligned}$$

So, we have an explicit form for $E(\hat{\theta}-\theta)$, that is the bias of the MLE with order $O(n^{-1})$. Using Bartlett's strategy we have

$$E(\hat{\theta}-\theta) = b(\theta) + O(n^{-2}),$$

where

$$\begin{aligned} b(\theta) &= \frac{1}{2} i_{200}(\theta)^{-2} \left\{ 2 i_{110}(\theta) + i_{001}(\theta) \right\} = O(n^{-1}), &(1)\\ b(\theta) &= -\frac{1}{2} i_{200}(\theta)^{-2} \left\{ i_{300}(\theta) + i_{110}(\theta) \right\} = O(n^{-1}). &(2) \end{aligned}$$

So the bias corrected estimator of θ is $\hat{\theta}_{BC} = \hat{\theta} - b(\hat{\theta})$ where, $E(\hat{\theta}_{BC} - \theta) = O(n^{-2})$.

Bartlett's method is also used for p-dimensional parameter vector, $\boldsymbol{\Theta}$. For simplicity and without loss of generality assume that $l(\boldsymbol{\Theta})$ be the log-likelihood function based on a sample of n observations for p-dimensional parameter vector $\boldsymbol{\Theta} = (\theta_1, \ldots, \theta_p)'$. The joint cumulants of derivatives of $l = l(\boldsymbol{\Theta})$ are given by the following formulas

$$\begin{aligned} &k_{ij} = E\left[\frac{\partial^2 l}{\partial\theta_i \partial\theta_j}\right], \quad k_{ijl} = E\left[\frac{\partial^3 l}{\partial\theta_i \partial\theta_j \partial\theta_l}\right], \\ &k_{ij,l} = E\left[\left(\frac{\partial^2 l}{\partial\theta_i \partial\theta_j}\right)\frac{\partial l}{\partial\theta_l}\right], \quad k_{ij}^{l} = \frac{\partial k_{ij}}{\partial\theta_l}, \quad i,j,l = 1,2,\ldots,p. \qquad (3) \end{aligned}$$

Assume that the log-likelihood is well-defined and regular function relative to all third order derivatives and all of the above equations are of order $O(n)$. Also, assume that $K = [-\kappa_{ij}]$ is the $p \times p$ information matrix of $\boldsymbol{\Theta}$ for $i,j = 1,2,\ldots,p$. Cox and Snell (1968) showed that when the sample is independent but not necessarily identically distributed, the bias of the estimator of the s^{th} element of $\boldsymbol{\Theta}$, namely $\hat{\theta}_s$ may be expressed in the form of

$$Bias(\hat{\theta}_s) = \sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{l=1}^{p} \kappa^{si}\kappa^{jl}\left[\frac{1}{2}k_{ijl} + k_{ij,l}\right] + O(n^{-2}), \ \ s = 1,2,\ldots,p, \qquad (4)$$

where κ^{ij} is the (i,j)th element of the inverse of the information matrix, K. Then, Cordeiro and Klein (1994) showed that when all equations in (3) are of order $O(n)$, even if the observations are not independent, equation (4) is satisfied. Thus they expressed a better form of $Bias(\hat{\theta}_s)$ as follow

$$Bias(\hat{\theta}_s) = \sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{l=1}^{p} \kappa^{si}\kappa^{jl}\left[k_{ij}^{(l)} - \frac{1}{2}k_{ijl}\right] + O(n^{-2}), \ \ s = 1,2,\ldots,p. \qquad (5)$$

By defining $a_{ij}^{(l)} = k_{ij}^{(l)} - \frac{1}{2}k_{ijl}$ for $i,j,l = 1,2,\ldots,p$ and matrix $A^{(l)} = \left\{a_{ij}^{(l)}\right\}$, and consideration $A = \left[A^{(1)}|A^{(2)}|\ldots.|A^{(p)}\right]$, the bias of $\hat{\boldsymbol{\Theta}}$ can be written as

$$Bias(\hat{\boldsymbol{\Theta}}) = K^{-1}A\ vec(K^{-1}) + O(n^{-2}).$$

Finally, the bias-correction of the MLE will be defined as $\tilde{\boldsymbol{\Theta}} = \hat{\boldsymbol{\Theta}} - \hat{K}^{-1}\hat{A}\,vec(\hat{K}^{-1})$ where $\hat{K} = K|_{\hat{\boldsymbol{\Theta}}}$ and $\hat{A} = A|_{\hat{\boldsymbol{\Theta}}}$. One advantage of this method is that these expressions can also be calculated when the likelihood equations do not have a closed-form solution. In such cases the bias-correction of the ML can be obtained by using numerical computation methods. In this case $\tilde{\boldsymbol{\Theta}}$ is unbiased of order $O(n^{-2})$.

One of the problems of the Bartlett's method is requirement the mean of the first and second derivatives of the likelihood function and also the mean of the $l'(\hat{\theta})l'(\hat{\theta})$ and $\left(l'(\hat{\theta})\right)^3$, (see equations (1)-(5)). As we see, these values are not necessarily available for all distributions. As example, consider the two-parameter Maxwell distribution, which we will discuss later. Therefore, in these cases, a new method is proposed for the bias correction of the MLE to hope that derivatives of lower order of the log-likelihood function $l(\theta)$ will be useful.

2.2 Firth's bias prevention

The bias correction depends on the finiteness of the maximum likelihood estimators. By definition, when the MLE is infinite, the bias correction of the MLE is not defined. Another method is the bias prevention which is proposed by Firth (1993) and has a different theory.

If $U(\theta) = l'(\theta)$ and $j(\theta) = -l''(\theta)$ are the score functions and the observed information, respectively, then by subtracting an appropriate part of the score function, we can remove bias of order $O(n^{-1})$ from the MLE. This idea is bias prevention of $b(\theta)$ by addition of $M^*(\theta)$ to score function. This method leads to a modified likelihood function as

$$U^*(\theta) = U(\theta) + M^*(\theta) = 0.$$

Thus $\hat{\theta}^*$ is obtained by solving equation $U^*(\theta) = 0$. The adjusted $M^*(\theta)$ is selected so that

$$E(\hat{\theta}^* - \theta) = O(n^{-2}). \tag{6}$$

Using Taylor series for $U^*(\theta) = 0$ around θ we can find an expansion of $(\hat{\theta}^* - \theta)$. So using condition (6), we have

$$M^*(\theta) = -b(\theta)E\left(j(\theta)\right). \tag{7}$$

2.3 Bootstrap approach

Jackknife bias reduction idea is the calculation of the estimate by deleting one or more observations at a time from the original data. An estimate for the bias of the statistic, can be calculated from this new set of replicates of the statistic. The idea of bootstrap's bias reduction is sampling with replacement from the main data and calculate the statistic in each cases. Therefore, it is possible to calculate the statistic for the bias of estimator by these samples. Although mathematically, there is a widespread theory about the difference between these two methods, in practice, for researchers the main difference is that by repeating these methods on a given data, bootstrap offers different results, while Jackknife is able to give the same results each time.

Here, we study Efron's resampling method (1979) to calculate the bias correction of MLE as another algebraic method. Assume $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ is a random sample of size n of random variable Y with distribution function F. Let $\eta = t(F)$ be a parameter of known function F and $\hat{\eta} = s(\boldsymbol{Y})$ denotes the estimation of η. We assume a large number of simulated samples with replacement of the sample $\mathbf{y}$ in Efron's bootstrap resampling and calculate $\hat{\eta}$ in each bootstrap samples $\mathbf{y}^* = (y_1^*, \dots, y_n^*)$, say $\hat{\eta}^* = s(\mathbf{y}^*)$. Then we use empirical distribution of $\hat{\eta}^*$ to estimate distribution function $\hat{\eta}$. If F belongs to a known parametric family with finite dimension of F_η, then we can obtain parametric estimation of F by consistence estimation $F_{\hat{\eta}}$. The bias of $\hat{\eta} = s(y)$ may be written as

$$B_F(\hat{\eta}, \eta) = E_F[s(y)] - \hat{\eta}(F),$$

where the index F shows that the expected value is obtained under F. The estimate of bootstrap bias-correction is calculated by replacing F with $F(\hat{\eta})$. So bias may be written as

$$B_{F(\hat{\eta})} = E_{F(\hat{\eta})}[\hat{\eta}] - \hat{\eta}.$$

For N bootstrap independent samples where are generated from main data y, we calculate bootstrap estimates corresponding to $\left(\hat{\eta}^{*(1)}, \dots, \hat{\eta}^{*(n)}\right)$. When N is large, the expected value of $E_{F(\hat{\eta})}[\hat{\eta}]$ may be approximated as

$$\hat{\eta}^{*(.)} = \frac{1}{N} \sum_{i=1}^{N} \hat{\eta}^{*(i)}. \tag{8}$$

The estimation of bootstrap of $\hat{\eta}$ by using N iteration is

$$B_{F(\hat{\eta})}(\hat{\eta}, \eta) = \hat{\eta}^{*(.)} - \hat{\eta}.$$

The MLE with a second order correction for arbitrary distribution can be calculated as follows

$$\eta^B = \hat{\eta} - B_{F(\hat{\eta})}(\hat{\eta}, \eta) = 2\hat{\eta} - \hat{\eta}^{*(.)},$$

where $\hat{\eta}^{*(.)}$, may be calculated by bootstrap samples and using (8).

2.4 Bias of MLE by using new method

Consider the definitions of Section 1. By Taylor's expansion to the first order of the first derivative of the log-likelihood function at $\theta = \hat{\theta}$, we have

$$0 = l'(\hat{\theta}) = l'(\theta) + (\hat{\theta} - \theta) l''(\theta) + O_p(n^{-1}).$$

So for any t, we can write

$$l'(t) + (\hat{\theta} - t) l''(t) + O_p(n^{-1}) = 0. \tag{9}$$

Now, we integrate both sides of equation (9) over $-\theta < t < \theta$. Note that this is possible only if the parameter space be the real line $\mathbb{R}$. So we have

$$\int_{-\theta}^{\theta} l'(t)dt + \int_{-\theta}^{\theta} (\hat{\theta} - t) l''(t)dt \approx 0.$$

Using integration by parts, we have

$$2l(\theta) - 2l(-\theta) + \hat{\theta}\left(l'(\theta) - l'(-\theta)\right) - \theta\left(l'(\theta) + l'(-\theta)\right) \approx 0.$$

So, we have $\hat{\theta} \approx \frac{2l(-\theta)-2l(\theta)+\theta\left(l'(\theta)+l'(-\theta)\right)}{l'(\theta)-l'(-\theta)}$, or $\hat{\theta}-\theta \approx \frac{2l(-\theta)-2l(\theta)+2\theta l'(-\theta)}{l'(\theta)-l'(-\theta)}$. Now, using Taylor series (first order) of $\frac{1}{l'(\theta)-l'(-\theta)}$ around $E\left(l'(\theta)-l'(-\theta)\right) = -E\left(l'(-\theta)\right)$, We have

$$\begin{aligned}
\frac{1}{l'(\theta)-l'(-\theta)} &= \frac{-1}{E(l'(-\theta))} - \frac{1}{E^2\left(l'(-\theta)\right)}\left(l'(\theta)-l'(-\theta)+E\left(l'(-\theta)\right)\right) + O_P\left(n^{-1}\right) \\
&= \frac{l'(-\theta)-l'(\theta)-2E\left(l'(-\theta)\right)}{E^2\left(l'(-\theta)\right)} + O_P(n^{-1}).
\end{aligned}$$

Therefore for sufficiently large n,

$$\hat{\theta}-\theta \approx \left(2l(-\theta)-2l(\theta)+2\theta l'(-\theta)\right)\frac{l'(-\theta)-l'(\theta)-2E\left(l'(-\theta)\right)}{E^2\left(l'(-\theta)\right)}. \tag{10}$$

Finally, by obtaining expectation of both sides of the equation (10), we will get the bias of MLE i.e. $E(\hat{\theta}-\theta)$. For simplicity we recommend using symbols $v(\theta)^a = \left(v(\theta)\right)^a$, $j_{rst}^{abc}(\theta) = E\left\{l(a\theta)^r l'(b\theta)^s l(c\theta)^t\right\}$ and $j_{rs}^{ab}(\theta) = E\left\{l(a\theta)^r l'(b\theta)^s\right\}$. We have

$$\begin{aligned}
E(\hat{\theta}-\theta) \approx\ & \frac{1}{j_{01}^{1-1}(\theta)^2}\left(2j_{11}^{-1-1}(\theta) - 2j_{11}^{-11}(\theta) - 4j_{10}^{-11}(\theta)j_{01}^{1-1}(\theta) - 2j_{11}^{1-1}(\theta)\right. \\
& \left. +2j_{11}^{11}(\theta) + 4j_{10}^{11}(\theta)j_{01}^{1-1}(\theta) + 2\theta j_{02}^{1-1}(\theta) - 2\theta j_{011}^{1-11}(\theta) - 4\theta j_{01}^{1-1}(\theta)^2\right)
\end{aligned} \tag{11}$$

It should be noted that in this method of bias estimation, we only need the expected value of the first order derivative of logarithm of likelihood or their product. Therefore, this method can be used for estimate bias of MLE for a larger family of distributions (Including the location parameter of Maxwell distribution). The new method proposed in one-parameter case is not simply generalizable to multi-parameters. In a two-parameters case, assume that we have a regular parametric model with logarithm of likelihood function, $l(\theta_1,\theta_2)$.

Assume that the MLE of parameters are $\hat{\theta}_1$ and $\hat{\theta}_2$, respectively. For simplicity we use the symbols of Section 2.1 and let $l_1'(\theta_1,\theta_2) = \frac{\partial l(\theta_1,\theta_2)}{\partial\theta_1}$, $l_2'(\theta_1,\theta_2) = \frac{\partial l(\theta_1,\theta_2)}{\partial\theta_2}$, $l_1''(\theta_1,\theta_2) = \frac{\partial^2 l(\theta_1,\theta_2)}{\partial^2\theta_1}$ and $l_2''(\theta_1,\theta_2) = \frac{\partial^2 l(\theta_1,\theta_2)}{\partial^2\theta_2}$. In this case, the functions $l_1'(\theta_1,\hat{\theta}_2)$ and $l_2'(\hat{\theta}_1,\theta_2)$ are one-variable functions equal to zero at $\theta_1=\hat{\theta}_1$ and $\theta_2=\hat{\theta}_2$, respectively. Therefore, using first order of Taylor expansion for $l_1'(\theta_1,\hat{\theta}_2)$, $l_2'(\hat{\theta}_1,\theta_2)$ at points $\theta_1=\hat{\theta}_1$ and $\theta_2=\hat{\theta}_2$, we have

$$\begin{aligned}
0 = l_1'(\hat{\theta}_1,\hat{\theta}_2) = l_1'(\theta_1,\hat{\theta}_2) + \left(\hat{\theta}_1-\theta_1\right)l_1''(\theta_1,\hat{\theta}_2) + O_p(n^{-1}), \\
0 = l_2'(\hat{\theta}_1,\hat{\theta}_2) = l_2'(\hat{\theta}_1,\theta_2) + \left(\hat{\theta}_2-\theta_2\right)l_2''(\hat{\theta}_1,\theta_2) + O_p(n^{-1}),
\end{aligned}$$

or, for each t_1 and t_2 we can write

$$\begin{aligned}
0 = l_1'(\hat{\theta}_1,\hat{\theta}_2) = l_1'(t_1,\hat{\theta}_2) + \left(\hat{\theta}_1-t_1\right)l_1''(t_1,\hat{\theta}_2) + O_p(n^{-1}), \\
0 = l_2'(\hat{\theta}_1,\hat{\theta}_2) = l_2'(\hat{\theta}_1,t_2) + \left(\hat{\theta}_2-t_2\right)l_2''(\hat{\theta}_1,t_2) + O_p(n^{-1}).
\end{aligned}$$

Now, to have lower orders derivatives, we can integrate the sides of the above equations on $-\theta_1 < t_1 < \theta_1$ and $-\theta_2 < t_2 < \theta_2$,

$$\begin{aligned}\int_{-\theta_1}^{\theta_1} l_1'(t_1,\hat{\theta}_2)dt_1 + \int_{-\theta_1}^{\theta_1} (\hat{\theta}_1 - t_1) l_1''(t_1,\hat{\theta}_2)dt_1 &\approx 0,\\ \int_{-\theta_2}^{\theta_2} l_2'(\hat{\theta}_1,t_2)dt_2 + \int_{-\theta_2}^{\theta_2} (\hat{\theta}_2 - t_2) l_2''(\hat{\theta}_1,t_2)dt_2 &\approx 0.\end{aligned}$$

Just like as one-parameter case and with some calculations, we have

$$\hat{\theta}_1 - \theta_1 \approx \frac{2l(-\theta_1,\hat{\theta}_2) - 2l(\theta_1,\hat{\theta}_2) + 2\theta_1 l'(-\theta_1,\hat{\theta}_2)}{l'(\theta_1,\hat{\theta}_2) - l'(-\theta_1,\hat{\theta}_2)}, \tag{12}$$

$$\hat{\theta}_2 - \theta_2 \approx \frac{2l(\hat{\theta}_1,-\theta_2) - 2l(\hat{\theta}_1,\theta_2) + 2\theta_2 l'(\hat{\theta}_1,-\theta_2)}{l'(\hat{\theta}_1,\theta_2) - l'(-\hat{\theta}_1,\theta_2)}. \tag{13}$$

Now for the approximate denominator of the (12) and (13), since $l'(\theta_1,\hat{\theta}_2)$ and $l'(\hat{\theta}_1,\theta_2)$ are not necessarily the derivative of the log-likelihood of a special density, so in this case, it is not possible to use Taylor's expansion around their expected values. So, using this method in two or more parameters is an open problem.

The method that we use to calculate the bias correction of the location and scale parameters of Maxwell distribution is first to remove scale parameter of the two-parameter Maxwell density function and then consider the estimate bias correction for the desired parameter (location parameter). In fact, in this case, a certain amount of parameters estimate placed in the profile loglikelihood function. Another example is provided in bias prevention, see Sartori (2006).

3 Maxwell distribution bias correction

3.1 One-parameter

To generate random numbers from the one-parameter Maxwell distribution, we use the following process:

If Y has gamma distribution with scale parameter $\frac{1}{\theta}$ and the shape parameter $\frac{3}{2}$, then $X = +\sqrt{Y}$ has $M(\theta)$ distribution (see Krishna and Malik, 2012). Moreover, the one-parameter Maxwell distribution moments can be calculated by using the following lemma (Ling and Giles, 2014).

Lemma 3.1. *The rth moment of $M(\theta)$ for $r > -3$ is $\mu_r = \frac{2}{\sqrt{\pi}}\Gamma\left((r+3)/2\right)\frac{1}{\theta^{r/2}}$.*

Ling and Giles (2014) studied the empirical term for the bias of order $O(n^{-1})$ of MLE for generalized Rayleigh distribution with the parameters (θ, k) and in a special case for the one-parameter Maxwell distribution (equivalent to generalized Rayleigh distribution with parameters $\left(\theta = \frac{1}{2\lambda^2}, k = \frac{1}{2}\right)$) using Bartlett's method. According to this paper, the bias-correction of MLE of the parameter θ is

$$Bias(\hat{\theta})) = \frac{\theta\{[2(k+1)(-\frac{1}{k^2} + \Psi_{(1)}(k)) - 3](-\frac{1}{k^2} + \Psi_{(1)}(k)) - (k+1)(\frac{2}{k^3} + \Psi_{(2)}(k))\}}{2n[(k+1)(-1/k^2 + \Psi_{(1)}(k)) - 1]^2}, \tag{14}$$

where $\Psi_{(1)}(k)$ and $\Psi_{(2)}(k)$ are the digamma and trigamma functions, respectively. By substituting $k = 1/2$ in (14), we obtain the bias correction for the parameter θ in one-parameter Maxwell distribution as

$$Bias(\hat{\theta}) = 0.67\frac{\theta}{n}. \tag{15}$$

In this section, we directly calculate the bias correction of θ by using Bartlett's method. According to the given the one parameter Maxwell density function we can write

$$l(x_i;\theta) = n\log\frac{4}{\sqrt{\pi}} + \frac{3n}{2}\log\theta + \sum_{i=1}^{n} 2\log x_i - \theta\sum_{i=1}^{n} x_i^2, \quad \theta > 0.$$

So, we have:

$$U(\theta) = l'(\theta) = \frac{\partial l}{\partial\theta} = \frac{3n}{2\theta} - \sum_{i=1}^{n} x_i^2,$$

$$l''(\theta) = \frac{\partial^2 l}{\partial\theta^2} = \frac{-3n}{2\theta^2}, \qquad l'''(\theta) = \frac{\partial^3 l}{\partial\theta^3} = \frac{3n}{\theta^3}.$$

Each of (1) and (2) can be used to calculate the bias correction of the parameter θ. Using (2), we need $i_{300}(\theta) = E\left(\frac{\partial l}{\partial\theta}\right)^3$, which its calculation is complex. Thus for calculating the bias of the estimator of the parameter θ, we use (1). So we have

$$i_{200} = -i_{010} = -E\left(l''(\theta)\right) = \frac{3n}{2\theta^2}, \qquad i_{001}(\theta) = E\left(l'''(\theta)\right) = \frac{3n}{\theta^3},$$

$$i_{110}(\theta) = E\left(l'(\theta)l''(\theta)\right) = E\left[\left(\frac{3n}{2\theta} - \sum_{i=1}^{n} x_i^2\right)\left(\frac{-3n}{2\theta^2}\right)\right] = 0.$$

Substituting calculated values in (1), we have:

$$b(\hat{\theta}) = \frac{1}{2} i_{200}(\theta)^{-2}\left\{2i_{110}(\theta) + i_{001}(\theta)\right\} = \frac{2\theta}{3n}. \tag{16}$$

By comparing the (15) and (16), we can conclude that the two equations are almost equivalent. On the other hand, we can use the bias obtained in (15) or (16) in bias prevention method by Firth. By using the (7), we have: $M(\theta) = -b(\theta)E\left(j(\theta)\right) = -\frac{2\theta}{3n}\frac{3n}{2\theta^2} = -\frac{1}{\theta}$. So Firth's bias prevention estimator is obtained by solving the following equation:

$$U^*(\theta) = U(\theta) + M(\theta) = \frac{3n}{2\theta} - \sum_{i=1}^{n} x_i^2 - \frac{1}{\theta} = 0.$$

3.2 Two-parameter

If a random variable X is defined as $X = M(\theta) + \mu$ with $\mu \in \mathbb{R}$, then $X \sim M(\theta,\mu)$ where θ and μ are the scale and location parameters of a two-parameter Maxwell distribution, respectively. Therefore, the probability density function (pdf) of X is

$$f(x;\theta,\mu) = \frac{4}{\sqrt{\pi}}\theta^{\frac{3}{2}}\left(x-\mu\right)^2 e^{-\theta(x-\mu)^2}, \; x > \mu, \theta > 0,$$

and the corresponding cumulative distribution function (cdf) is

$$F(x;\theta,\mu)=\frac{2}{\sqrt{\pi}}\Gamma\left(3/2,\theta\left(x-\mu\right)^2\right),\ x>\mu.$$

Dey et al. (2016) studied the problem of estimating location and scale parameters of the Maxwell distribution from both frequentist and Bayesian methods. They studied and compared several different methods for estimating the parameters of this distribution, namely, ML, method of moments, least square's, weighted least square and Bayesian method. After comparing these methods, the authors concluded that for small samples, the estimate of location parameter μ will have a higher bias in comparing to the estimate of scale parameter θ . Finally, they suggested for the parameters of this distribution, it is better to use the MLEs or the Bayes estimators. By examining the bias value of the MLE for the location parameter, it is considered that in this distribution the bias correction estimator of this parameter is useful.

Note that while $M(\theta,\mu)=\mu+M(\theta,0)$, so the moments of the distribution $M(\theta,\mu)$ are easily calculated by using Lemma 3.1.

Now we calculate the MLE of the parameters θ and μ using a random sample of size n from $M(\theta,\mu)$. If both parameters of $M(\theta,\mu)$ are unknown, the log-likelihood function for the two-parameter Maxwell distribution is

$$\log L\left(\theta,\mu\right)=\frac{3n}{2}\log\theta+2\sum_{i=1}^{n}\log\left(x_i-\mu\right)-\theta\sum_{i=1}^{n}\left(x_i-\mu\right)^2+C,$$

where C is a constant. Normal equations that are used to estimate distribution parameters are

$$\frac{\partial\log L}{\partial\theta} = \frac{3n}{\theta}-\sum_{i=1}^{n}\left(x_i-\mu\right)^2=0, \tag{17}$$

$$\frac{\partial\log L}{\partial\mu} = -\sum_{i=1}^{n}\frac{2}{\left(x_i-\mu\right)}+2\theta\sum_{i=1}^{n}\left(x_i-\mu\right)=0. \tag{18}$$

If we wish to use the equation (4) to calculate the bias correction of θ and μ, for example we need the expected value of

$$\frac{\partial\log L}{\partial\mu^3}=\frac{\partial}{\partial\mu}\left(-\sum_{i=1}^{n}\frac{2}{\left(x_i-\mu\right)^2}-2n\theta\right)=-\sum_{i=1}^{n}\frac{4}{\left(x_i-\mu\right)^3}. \tag{19}$$

But according to Lemma 3.1, since μ_r' can only be calculated for $r>-3$, the expected value of (19) is not available. As a result, using the Bartlett method, the bias calculation of the parameters of the location and the scale of this distribution are not available. So proposing a new method to estimate bias of MLE using the expected value of lower order derivatives will be useful.

According to conclusions of Dey et al. (2016), in Maxwell Distribution bias of MLE for estimate location parameter is more significant relative to the scale parameter, especially for small samples. So, estimation and bias correction of location parameter is useful in two-parameter Maxwell distribution. So, first the scale parameter should be removed from two-parameter Maxwell distribution. We will use the following process:

Assume X has $M(\theta,\mu)$, so $Y = X\sqrt{\theta}$ has a density of the form

$$f(y;\theta,\mu) = \frac{4}{\sqrt{\pi}}\left(y-\mu\sqrt{\theta}\right)^2 e^{-\left(y-\mu\sqrt{\theta}\right)^2},\ y>\mu,\theta>0.$$

So, we can write: If, $X \sim M(\theta,\mu)$, then $Y = X\sqrt{\theta} \sim M\left(1,\mu\sqrt{\theta}\right)$. Now considering the parameter $\mu' = \mu\sqrt{\theta}$ as a new location parameter in the Maxwell distribution, we try to estimate this parameter. We assume

$$f(y;\mu') = \frac{4}{\sqrt{\pi}}(y-\mu')^2 e^{-\left(y-\mu'\right)^2},\ y>\mu'. \tag{20}$$

In this case, the log-likelihood function of the form

$$l(\mu') = \log L(\mu') = n\log\frac{4}{\sqrt{\pi}} + 2\sum_{i=1}^{n}\log(y_i-\mu') - \sum_{i=1}^{n}(y_i-\mu')^2.$$

Now, we follow method of section 2.4 to estimate bias of MLE of μ' by solving the following equation

$$U(\mu') = \frac{\partial\log L}{\partial\mu'} = -\sum_{i=1}^{n}\frac{2}{(y_i-\mu')} + 2\sum_{i=1}^{n}(y_i-\mu') = 0,$$

and is calculated using numerical methods. To calculate the bias of MLE for μ' using equation (11), we should calculate functions which are simply calculated using numerical methods. For simplicity, we define the following functions:

$$\begin{aligned}
f(\mu',\mu_1,\mu_2) &= E_{\mu'}\left\{l(\mu_1)U(\mu_2)\right\} \\
&= E_{\mu'}\left\{\left(n\log\frac{4}{\sqrt{\pi}} + 2\sum_{i=1}^{n}\log(y_i-\mu_1) - \sum_{i=1}^{n}(y_i-\mu_1)^2\right)\right. \\
&\quad \left.\times\left(-\sum_{i=1}^{n}\frac{2}{(y_i-\mu_2)} + 2\sum_{i=1}^{n}(y_i-\mu_2)\right)\right\} \\
&= n\log\frac{4}{\sqrt{\pi}}\left[-2nE_{\mu'}\left(\frac{1}{Y-\mu_2}\right) + 2nE_{\mu'}(Y-\mu_2)\right] \qquad (21) \\
&\quad -4n\left\{E_{\mu'}\left(\frac{\log(Y-\mu_1)}{Y-\mu_2}\right)\right. \\
&\quad \left.+(n-1)E_{\mu'}(\log(Y-\mu_1))E_{\mu'}\left(\frac{1}{Y-\mu_2}\right)\right\} \\
&\quad +4n\left\{E_{\mu'}((Y-\mu_2)\log(Y-\mu_1))\right. \\
&\quad \left.+(n-1)E_{\mu'}(\log(Y-\mu_1))E_{\mu'}(Y-\mu_2)\right\} \\
&\quad +2n\left\{E_{\mu'}\left(\frac{(Y-\mu_1)^2}{(Y-\mu_2)}\right)\right.
\end{aligned}$$

$$
\begin{aligned}
& \quad +(n-1)E_{\mu'}\left(Y-\mu_1\right)^2 E_{\mu'}\left(\frac{1}{(Y-\mu_2)}\right)\Big\} \\
& \quad -2n\Big\{E_{\mu'}\left(\left(Y-\mu_1\right)^2\left(Y-\mu_2\right)\right) \\
& \quad +(n-1)E_{\mu'}\left(Y-\mu_1\right)^2 E_{\mu'}\left(Y-\mu_2\right)\Big\} \\
g(\mu',\mu_1,\mu_2) &= E_{\mu'}\left\{U\left(\mu_1\right)U\left(\mu_2\right)\right\} \\
&= E_{\mu'}\Bigg\{\left(-\sum_{i=1}^{n}\frac{2}{(y_i-\mu_1)}+2\sum_{i=1}^{n}\left(y_i-\mu_1\right)\right) \\
& \quad \times\left(-\sum_{i=1}^{n}\frac{2}{(y_i-\mu_2)}+2\sum_{i=1}^{n}\left(y_i-\mu_2\right)\right)\Bigg\} \qquad (22) \\
&= 4n\Big\{E_{\mu'}\left(\frac{1}{(Y-\mu_1)(Y-\mu_2)}\right) \\
& \quad +(n-1)E_{\mu'}\left(\frac{1}{(Y-\mu_1)}\right)E_{\mu'}\left(\frac{1}{Y-\mu_2}\right)\Big\} \\
& \quad -4n\left\{E_{\mu'}\left(\frac{(Y-\mu_2)}{(Y-\mu_1)}\right)+(n-1)E_{\mu'}\left(\frac{1}{(Y-\mu_1)}\right)E_{\mu'}\left(Y-\mu_2\right)\right\} \\
& \quad -4n\left\{E_{\mu'}\left(\frac{(Y-\mu_1)}{(Y-\mu_2)}\right)+(n-1)E_{\mu'}\left(\frac{1}{(Y-\mu_2)}\right)E_{\mu'}\left(Y-\mu_1\right)\right\} \\
& \quad +4n\{E_{\mu'}((Y-\mu_1)(Y-\mu_2))+(n-1)E_{\mu'}(Y-\mu_1)E_{\mu'}(Y-\mu_2)\}, \\
h(\mu',\mu_1) &= E_{\mu'}\left\{U\left(\mu_1\right)\right\} \\
&= E_{\mu'}\left(-\sum_{i=1}^{n}\frac{2}{(y_i-\mu_1)}+2\sum_{i=1}^{n}\left(y_i-\mu_1\right)\right) \qquad (23) \\
&= -2nE_{\mu'}\left(\frac{1}{Y-\mu_1}\right)+2nE_{\mu'}\left(Y-\mu_1\right), \\
m(\mu',\mu_1) &= E_{\mu}\left\{l\left(\mu_1\right)\right\} \\
&= E_{\mu'}\left(n\log\frac{4}{\sqrt{\pi}}+2\sum_{i=1}^{n}\log\left(y_i-\mu_1\right)-\sum_{i=1}^{n}\left(y_i-\mu_1\right)^2\right) \\
&= n\log\frac{4}{\sqrt{\pi}}+2nE_{\mu'}\left(\log\left(Y-\mu_1\right)\right)-nE_{\mu'}\left(Y-\mu_1\right)^2. \qquad (24)
\end{aligned}
$$

In functions (21)-(24) we calculate $E_{\mu'}$ based on pdf in (20) numerically. According to equation (11) and given functions (21)-(24) we have

$$
\begin{aligned}
bias(\hat{\mu}') &= E\left(\hat{\mu}'-\mu'\right) \\
&\approx \frac{1}{h(\mu',-\mu'^2)}(2f(\mu',-\mu',-\mu')-2f(\mu',-\mu',\mu') \\
& \quad -4m(\mu',-\mu')h(\mu',-\mu')-2f(\mu',\mu',-\mu')+2f(\mu',\mu',\mu') \\
& \quad +4m(\mu',\mu')h(-\mu',\mu')+2\mu' g(\mu',-\mu',-\mu') \\
& \quad -2\mu' g(\mu',-\mu',\mu')-4\mu' h(\mu',-\mu'^2)). \qquad (25)
\end{aligned}
$$

So, the bias correction of $\hat{\mu'}$ is calculated using equation

$$\tilde{\mu}' = \hat{\mu}' - bias\,(\hat{\mu}')\,|_{\hat{\mu}}. \tag{26}$$

In practice we need $\hat{\mu}'$ to estimate bias of μ' using equation (26). To calculate $\hat{\mu}'$ first we use ML method and solve equations (17) and (18) and calculate $\hat{\mu}$ and $\hat{\theta}$ and then estimate $\hat{\mu}'$ via $\hat{\mu}\sqrt{\hat{\theta}}$. Of course, for the more precise estimation of $\hat{\theta}$ we can also use Firth's bias prevention for $x - \hat{\theta}$. Finally, having calculated corrected estimate of μ' using equation (26), we calculate parameter μ using Equation $\tilde{\mu} = \hat{\mu} - \frac{\tilde{\mu}'}{\sqrt{\hat{\theta}}}$.

4 Simulation studies

To compare the estimates of the Maxwell distribution parameters, we use the Monte Carlo simulation. This simulation was performed using R software. We selected sample sizes of $n = 10, 20, 40, 60, 100$ with $\theta = 1$ and $\mu = 0, 1$. The two- parameter Maxwell distribution was simulated using lemma 1 and definition 1. The Monte Carlo sample repeat count and bootstrap repeat count per (n, θ, μ) are $M = 5000$ and $B = 1000$, respectively. In simulation to evaluate the mean bias and the mean squared error (MSE), per β^{est} parameter estimate of β we use the following relations, respectively:

$$Bias(\hat{\beta}^{est}) = \frac{1}{M}\sum_{i=1}^{M}(\hat{\beta}^{est} - \beta), \quad MSE(\hat{\beta}^{est}) = \frac{1}{M}\sum_{i=1}^{M}(\hat{\beta}^{est} - \beta)^2.$$

Table 1 and 2 represent the mean bias and the mean squared error percent. For example, the mean square error percent for β is obtained by relation $100 \times MSE(\hat{\beta}^{est})$. We use the method proposed in the Section 2 for each random and independent sample, so, for a sample of $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ we use five methods to estimate (θ, μ).
Method 1: Using the ML method, this is obtained using equations (17) and (18). In this case we use *nlm* function in *R* software. We show this estimate as $(\hat{\theta}, \hat{\mu})$.
Method 2: Using bootstrap bias correction for the MLE in the Section 2.3. The estimation is shown as $(\hat{\theta}^B, \hat{\mu}^B)$.
Method 3: Using the bias correction of MLE in the new method for μ. In this method first the MLEs $(\hat{\theta}, \hat{\mu})$ is calculated for $\boldsymbol{x}$. Then, using relation $\hat{\mu'} = \hat{\mu}\sqrt{\hat{\theta}}$, and using the method in Section 3.2 we calculate $\tilde{\mu}' = \hat{\mu}' - bias\,(\hat{\mu}')$ and then $\tilde{\mu} = \hat{\mu} - \tilde{\mu}'/\sqrt{\hat{\theta}}$. The estimates are shown as $(\hat{\theta}, \tilde{\mu})$.
Method 4: Using the bias correction of the MLE in the new method and the Bartlett bias correction method, first it is calculated the MLE of μ as $\hat{\mu}$. Then, using Bartlet's bias correction and using the equation 14 we calculate the corrected parameter estimate $\hat{\theta}$, namely $\hat{\theta}^{BC} = \hat{\theta} - \frac{2\hat{\theta}}{3n}$ where n is the sample size. Now, using relation $\hat{\mu'}^{BC} = \hat{\mu}\sqrt{\hat{\theta}^{BC}}$ and by using the section 2 method, we calculate $\tilde{\mu}^{BC} = \hat{\mu}'^{BC} - bias\left(\hat{\mu}'^{BC}\right)$. We show this estimate as $(\hat{\theta}^{BC}, \tilde{\mu}^{BC})$.
Method 5: Using the bias correction of the MLE in the new method and the Firth's bias prevention, first it is calculated the MLE of μ as $\hat{\mu}$. Then, using Firth's bias prevention and using equation 7 we calculate the prevention of the parameter estimate

$\hat{\theta}^F$. Then using $\hat{\mu'}^{BC} = \hat{\mu}\sqrt{\hat{\theta}^F}$ and using the previous section method we calculate $\tilde{\mu}^F = \hat{\mu}'^F - bias\left(\hat{\mu}'^F\right)$; these estimates are shown as $(\hat{\theta}^F, \tilde{\mu}^F)$.

Table 1 shows the bias percent of the parameters estimation of $M(1,0)$ using methods 1 to 5. Based on this table, the bias of maximum likelihood estimates of (θ, μ), namely $(\hat{\theta}, \hat{\mu})$ are always positive. Also, the bias and MSE are large for a small sample size and decrease with increasing sample size. On the other hand, for a constant sample size, the bias and MSE of $\hat{\theta}$ are larger than $\hat{\mu}$.

Table 1: Bias percent [MSE %] estimates of distribution parameters M(1,0).

n	$\hat{\theta}$	$\hat{\theta}^B$	$\hat{\theta}^{BC}$	$\hat{\theta}^F$	$\hat{\mu}$	$\hat{\mu}^B$	$\tilde{\mu}$	$\tilde{\mu}^{BC}$	$\tilde{\mu}^F$
10	24.55	−3.87	1.12	2.31	19.71	5.24	4.87	5.09	4.59
	[32.15]	[17.36]	[15.04]	[14.47]	[23.10]	[15.83]	[13.63]	[14.72]	[15.96]
20	12.02	−2.91	0.71	0.70	10.92	−4.11	3.92	3.07	3.82
	[18.96]	[12.02]	[11.93]	[10.08]	[14.10]	[13.16]	[12.74]	[12.67]	[13.75]
30	11.78	1.20	0.59	0.83	8.43	3.281	−2.19	3.02	−3.01
	[17.24]	[11.50]	[9.01]	[11.70]	[12.48]	[12.63]	[12.36]	[11.97]	[12.09]
40	9.54	0.19	−0.06	0.07	5.97	2.93	−2.02	−2.99	2.54
	[13.08]	[9.38]	[6.78]	[5.52]	[9.68]	[8.61]	[9.32]	[7.15]	[11.31]
50	7.94	0.15	0.04	0.07	5.02	−2.42	1.99	−2.35	−2.41
	[12.97]	[8.05]	[5.91]	[5.00]	[8.00]	[7.65]	[5.67]	[7.06]	[10.40]
60	5.37	−0.01	0.04	0.02	4.12	−2.01	1.83	−2.25	2.36
	[10.07]	[4.38]	[4.31]	[4.79]	[7.95]	[6.95]	[5.51]	[6.72]	[9.13]
70	4.87	0.09	−0.03	0.05	4.09	1.75	1.53	1.41	1.39
	[6.65]	[3.65]	[3.61]	[4.01]	[7.17]	[6.00]	[5.32]	[5.42]	[8.83]
80	4.54	−0.08	0.03	−0.04	3.65	−1.59	−0.91	1.27	−1.19
	[5.02]	[3.00]	[2.26]	[2.17]	[6.84]	[5.68]	[4.70]	[5.07]	[6.94]
90	3.92	0.05	−0.02	−0.02	2.79	0.94	0.89	−0.93	1.06
	[3.20]	[2.92]	[1.51]	[1.84]	[5.32]	[5.50]	[4.02]	[4.95]	[6.44]
100	3.59	0.03	0.01	0.02	2.06	0.64	0.57	0.42	−0.97
	[2.89]	[1.51]	[1.79]	[2.88]	[3.72]	[2.97]	[2.07]	[2.87]	[4.40]

Simulation studies on table 1 showed that bias of $(\hat{\theta}^B, \hat{\mu}^B)$ is sometimes positive and sometimes negative. But for constant sample these bias are smaller than the bias of $\hat{\theta}, \hat{\mu})$ in term of absolute value. This also applies to the MSE value of these estimates. In this case on the contrary of the previous case, bias of $\hat{\mu}^B$ especially on large sample size is bigger than the bias of $\hat{\theta}^B$. This result shows that the bootstrap's bias correction is useful for his distribution.

According to table 1 the bias of $(\hat{\theta}^{BC}, \tilde{\mu}^{BC})$ is also sometimes positive or negative. But the biasedness of $(\hat{\theta}^{BC}, \hat{\mu}^{BC})$ are smaller than the biasedness of $(\hat{\theta}^B, \hat{\mu}^B)$ in term of absolute for constant sample. But this is not true for MSE value of these estimates, especially for large sample. This result shows that the Bartlet's bias correction compared to bootstrap's bias correction is only logical in terms of reducing the biasedness of this distribution. Accordingly, the comparison between bias of $\tilde{\mu}$ and $\tilde{\mu}^{BC}$ also shows that $\tilde{\mu}$ has priority compared to $\tilde{\mu}^{BC}$ in reducing biasedness in most samples.

The comparison between bias and MSE of $\hat{\theta}^F$ and other estimates of θ shows that this estimate has priority in term of bias and MSE compared to $\hat{\theta}$ and $\hat{\theta}^B$, but has not priority relative to $\hat{\theta}^{BC}$. Finally, the comparison between bias and MSE estimates of $\tilde{\mu}^F$ and other estimates shows this estimate has not priority relative to other estimates in MSE but it has priority relative to other estimate of μ in biasedness.

Table 2 shows that the percent of biasedness in [MSE %] estimation of distribution parameters $M(1,1)$ using methods 1 to 5. The table 2 confirms the results of table 1, but the important point is that bias and MSE of parameters estimates of this distribution are considerably more than $M(1,0)$.

Generally, simulation shows that $(\hat{\theta}^{BC}, \tilde{\mu})$ is prefered to $(\hat{\theta}^{F}, \tilde{\mu})$ relative to biasedness criterian but not in MSE.

Table 2: Bias percent [MSE %] estimates of distribution parameters M(1,1).

n	$\hat{\theta}$	$\hat{\theta}^B$	$\hat{\theta}^{BC}$	$\hat{\theta}^F$	$\hat{\mu}$	$\hat{\mu}^B$	$\tilde{\mu}$	$\tilde{\mu}^{BC}$	$\tilde{\mu}^F$
10	26.13	5.98	3.32	4.64	20.40	8.97	6.07	6.37	5.99
	[35.72]	[17.30]	[16.24]	[15.33]	[26.83]	[22.60]	[19.41]	[18.73]	[17.04]
20	15.06	−4.37	2.97	−2.09	19.03	8.01	−5.80	5.50	−5.37
	[23.95]	[15.49]	[15.97]	[14.01]	[18.73]	[17.34]	[16.31]	[17.93]	[15.07]
30	13.41	4.05	−2.46	2.00	15.36	−7.64	−5.68	−5.43	5.27
	[19.15]	[14.62]	[14.02]	[13.73]	[17.93]	[12.06]	[12.67]	[13.23]	[14.02]
40	10.75	3.97	−2.34	−1.99	14.30	−6.93	5.01	−4.93	5.09
	[18.25]	[14.01]	[13.55]	[14.65]	[17.02]	[11.94]	[10.68]	[12.97]	[11.96]
50	9.00	3.61	2.10	−1.96	12.65	6.32	−4.82	−4.86	4.73
	[13.28]	[13.87]	[12.87]	[12.40]	[16.05]	[11.38]	[10.26]	[9.92]	[11.35]
60	7.43	−2.42	1.91	1.93	12.04	5.84	4.38	4.02	−3.93
	[12.02]	[10.86]	[12.03]	[12.01]	[15.32]	[10.62]	[9.66]	[9.86]	[10.27]
70	5.65	2.06	−1.89	1.93	11.61	−4.52	3.23	−3.11	3.51
	[10.38]	[9.67]	[11.93]	[11.32]	[15.07]	[9.41]	[9.51]	[9.39]	[9.30]
80	4.32	−1.97	1.66	−1.90	10.86	3.07	2.92	2.97	−2.83
	[7.11]	[9.00]	[11.04]	[10.95]	[14.90]	[9.00]	[9.35]	[8.92]	[8.28]
90	3.68	1.74	1.00	0.93	10.60	2.62	2.31	2.62	2.38
	[3.01]	[8.41]	[9.34]	[9.82]	[13.95]	[8.70]	[8.56]	[8.45]	[7.80]
100	3.51	−1.16	0.95	0.93	10.27	−1.80	−1.51	1.32	1.92
	[2.97]	[6.32]	[8.97]	[9.02]	[12.69]	[6.04]	[5.04]	[4.92]	[5.10]

References

Bartlett, M.S. (1953). Approximate confidence intervals. *Biometrika*, **40**(1/2), 12-19.

Bekker, A. and Roux, J.J.J. (2005). Reliability characteristics of the Maxwell distribution: A Bayes estimation study. *Communications in Statistics-Theory and Methods*, **34**(11), 2169-2178.

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.

Cordeiro, G.M. and Klein, R. (1994). Bias correction in ARMA models. *Statistics and Probability Letters*, **19**(3), 169-176.

Cox, D.R. and Hinkley, D.V. (1979). *Theoretical Statistics.* Chapman and Hall/CRC.

Cox, D.R. and Snell, E.J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society*: Series B (Methodological), **30**(2), 248-265.

Dey, S., Dey, T. and Maiti, S.S. (2013). Bayesian inference for Maxwell distribution under conjugate prior. *Model Assisted Statistics and Applications*, **8**(3), 193-203.

Dey, S., Dey, T., Ali, S. and Mulekar, M.S. (2016). Two-parameter Maxwell distribution: Properties and different methods of estimation. *Journal of Statistical Theory and Practice*, **10**(2), 291-310.

Efron, B. (1992). *Bootstrap Methods: Another Look at the Jackknife*. In Breakthroughs in statistics. Springer, New York, NY.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, **80**(1), 27-38.

Iriarte, Y.A., Astorga, J.M., Bolfarine, H. and Gamez, H.W. (2017). Gamma-Maxwell distribution. *Communications in Statistics-Theory and Methods*, **46**(9), 4264-4274.

Kazemi, L., and Shahabi, C. (2012, November). *Geocrowd: enabling query answering with spatial crowdsourcing*. In Proceedings of the 20th international conference on advances in geographic information systems, 189-198.

Krishna, H. and Malik, M. (2009). Reliability estimation in Maxwell distribution with Type-II censored data. *International Journal of Quality and reliability management*. **26**(2), 184-195.

Krishna, H. and Malik, M. (2012). Reliability estimation in Maxwell distribution with progressively type-II censored data. *Journal of Statistical Computation and Simulation*, **82**(4), 623-641.

Ling, X. and Giles, D. (2014). Bias reduction for the maximum likelihood estimator of the parameters of the generalized Rayleigh family of distributions. *Communications in Statistics-Theory and Methods*, **43**(8), 1778-1792.

Maxwell, J. C. (1860). V. Illustrations of the dynamical theory of gases.—Part I. On the motions and collisions of perfectly elastic spheres. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **19**(124), 19-32.

Podder, C.K. and Roy, M. K. (2003). Bayesian estimation of the parameter of Maxwell distribution under MLINEX loss function. *Journal of Statistical Studies*, **23**, 11-16.

Sartori, N. (2006). Bias prevention of maximum likelihood estimates for scalar skew normal and skew t distributions. *Journal of Statistical Planning and Inference*, **136**(12), 4259-4275.

Tyagi, R.K. and Bhattacharya, S.K. (1989). Bayes estimation of the Maxwell's velocity distribution function. *Statistica*, **29**(4), 563-567.

Tyagi, R.K., and Bhattacharya, S.K. (1989). A note on the MVU estimation of reliability for the Maxwell failure distribution. *Estadistica*, **41**(137), 73-79.