JOURNAL OF STATISTICAL THEORY AND PRACTICE
2016, VOL. 10, NO. 2, 291–310
http://dx.doi.org/10.1080/15598608.2015.1135090

Two-parameter Maxwell distribution: Properties and different methods of estimation

Sanku Dey[a], Tanujit Dey[b], Sajid Ali[c], and Madhuri S. Mulekar[d]

[a]Department of Statistics, St. Anthony's College, Shillong, Meghalaya, India; [b]Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, Ohio, USA; [c]Department of Decision Sciences, Bocconi University, Milan, Italy; [d]Department of Mathematics and Statistics, University of South Alabama, Mobile, Alabama, USA

ABSTRACT

In this article we consider the problem of estimating location and scale parameters of the Maxwell distribution from both frequentist and Bayesian points of view. Additionally, some properties of the distribution, namely, stochastic ordering, Rényi and Shannon entropies, and order statistics, are derived. Behavior of the estimators from different frequentist approaches, namely, maximum likelihood, method of moments, least square's, and weighted least square as well as Bayes estimators of parameters, is compared with respect to bias, mean squared errors, and the coverage percentage extracted from bootstrap confidence intervals. The existence and uniqueness of the maximum likelihood estimators are also discussed. The Bayes estimators and the associated credible intervals are obtained using importance sampling technique under squared error loss function. A gamma prior is used for the scale parameter and a uniform prior for the location parameter. An example with flood-level data is used to illustrate applicability of procedures discussed.

ARTICLE HISTORY
Received 13 April 2015
Accepted 18 December 2015

KEYWORDS
Entropy; maximum likelihood estimators; method of moment estimators; least squares estimators; Bayes estimators

AMS SUBJECT CLASSIFICATION
60E05; 62F10; 62F15

1. Introduction

Maxwell (1867) derived a mathematical formulation in three-dimensional space to describe the distributions of speeds of molecules in thermal equilibrium and it came to be known as the Maxwell distribution. Not all molecules move at the same speed and a few molecules move at faster speeds, resulting in a leptokurtic distribution, that is, unimodal with a longer right tail. A major characteristic of the Maxwell distribution is that it has a smooth increasing failure rate, because of which it is useful in those life-testing and reliability studies in which the assumption of constant failure rate, such as in the case of exponential distribution, is not realistic.

In the statistics literature, we come across hundreds of continuous univariate distributions (see Johnson et al. 1994). Several distributions have been extensively used over the past decades for modeling data in varied fields such as engineering, actuarial, environmental, and medical sciences, biological studies, demography, economics, finance, and insurance. However, in several situations with lifetime, rainfall, flood, or earthquake data, the popular distributions do not fit well to data (see Kumaraswamy 1976; 1978) or the estimation of distribution parameters is not tractable (see Chattopadhyay et al. 2014). In such situations, other not-as-popular distributions might provide a better fit and/or have a smaller number of

CONTACT Madhuri S. Mulekar mmulekar@southalabama.edu Department of Mathematics and Statistics, University of South Alabama, 411 University Blvd North, ILB 325, Mobile, AL, 36688-0002, USA.

parameters to be estimated. The Maxwell distribution, which plays an important role in physics, reliability analysis, and other applied sciences, provides an alternative option to be used in such cases and needs to be studied further.

Typically, the Maxwell distribution has been described using only one parameter, a scale parameter, which limits its applicability in practice (see Feller 1971). Using various techniques, it is possible to introduce additional parameters and expand the family of distributions for added flexibility (see Gupta and Kundu 1999; Adamidis et al. 2005; Kus 2007; Barreto-Souza et al. 2011; Ristic and Balakrishnan 2012). In this study, we are interested in expanding the Maxwell to a two-parameter distribution by incorporating a location parameter and then developing point and interval estimators for both location and scale parameters.

Tyagi and Bhattacharya (1989a; 1989b) for the first time considered one parameter (scale) Maxwell distribution as a model for the distribution of lifetimes. They obtained the minimum variance unbiased and the Bayes estimators of the scale parameter, and the reliability function of this distribution. Chaturvedi and Rani (1998) generalized the Maxwell distribution through some transformation on a gamma distributed random variable. They also obtained the classical and Bayes estimators of the parameters. Howlader and Hossain (1998) derived the highest posterior density (HPD) intervals for the unknown scale parameter, as well as for a future observation considering an asymptotically locally invariant prior proposed by Hartigan (1964). Podder and Roy (2003) estimated the parameter of this distribution under the Modified Linear Exponential Loss Function (MLINEX). Bekker and Roux (2005) studied the maximum likelihood estimator (MLE), as well as the Bayes estimators of the truncated first moment and hazard function of the Maxwell distribution. Krishna and Malik (2009) obtained the MLE and the Bayes estimators of the reliability characteristics under a type II censoring scheme. Dey and Maiti (2010) considered one-parameter Maxwell distribution with a scale parameter and obtained Bayes estimators using noninformative and conjugate priors under symmetric as well as asymmetric loss functions, namely, the quadratic loss function, squared-log error loss function, and modified linear exponential loss function. Performances of all these estimators were compared on the basis of their estimated risk. Krishna and Malik (2012) compared the MLE and the Bayes estimators of the scale parameter and the reliability function under a progressive type II censoring scheme. Recently, Dey et al. (2013) studied the one-parameter Maxwell distribution under different loss functions, namely, the squared error loss function and precautionary loss function, and compared the performances of these estimators. They also obtained predictive density and HPD prediction interval for a future observation.

Let X follow a two-parameter Maxwell distribution, $M(\theta, \mu)$; then the probability density function (PDF) of X is

$$f(x;\theta,\mu) = \frac{4}{\sqrt{\pi}}\,\theta^{\frac{3}{2}}\,(x-\mu)^2;\ \ x>\mu;\theta>0 \tag{1}$$

where θ and μ are the scale and the location parameters, respectively. The corresponding cumulative distribution function (CDF) is

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

The objective of this study is to describe some parametric estimation methods for the two-parameter Maxwell distribution and to identify the most efficient estimators. Some useful structural properties of the Maxwell distribution, specifically, stochastic ordering, entropies, and order statistics, are described in the appendix.

Frequentist methods, also referred to as traditional methods, such as the maximum likelihood estimator (MLE) and the method of moments (ME) estimator, and less commonly used techniques such as the least-square estimator (LSE) and weighted least-square estimator (WLSE) are presented in this article. The performance of each of these methods is studied using simulations for different sample sizes and compared in terms of their resulting biases, mean squared errors (MSE), and the coverage percentage of 95% bootstrap confidence intervals. The Bayes estimators are obtained using importance sampling technique under squared error loss function and compared with frequentist estimators. Different estimation methods appeal differently to their users. With computational advances, the need to have an estimator with closed form has decreased substantially. Thus, a user may prefer to employ the uniformly minimum variance estimation method although the estimator does not have a closed-form expression.

Different estimation methods were compared for the generalized Rayleigh distribution by Kundu and Raqab (2005); for the generalized logistic distribution by Alkasasbeh and Raqab (2009); for the Weibull distribution by Teimouri et al. (2013); and for a two-parameter Rayleigh distribution by Dey et al. (2014). To our knowledge, so far no attempt has been made to study statistical properties of a two-parameter Maxwell distribution and to compare different estimation methods for it.

The rest of the article is organized as follows: section 2 describes the maximum likelihood, the method of moment, the least square, and the weighted least square estimators. The Bayes estimators are presented in section 3. Section 4 introduces both frequentist and Bayesian interval estimation for the parameters of the Maxwell distribution. In section 5, the performance of several estimation procedures based on coverage percentages of bootstrap confidence intervals using frequentist and Bayesian approaches is provided. The methodology developed in this article and the usefulness of the two-parameter Maxwell distribution are illustrated by using flood-level data. Some concluding remarks are provided in section 6, and results about three theoretical properties of a two-parameter Maxwell distribution are given in the appendix.

2. Maxwell parameter estimation using frequentist approaches

Here we describe how to obtain the maximum likelihood (MLE), method of moments (ME), ordinary least square (LSE), and the weighted least squares (WLSE) estimators of the parameters θ and μ using a random sample of size n from $M(\theta, \mu)$ population.

2.1. *Maximum likelihood estimation*

When both θ and μ are unknown, the log-likelihood function for the Maxwell distribution is

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

where C is an additive constant. The normal equations to estimate unknown parameters θ and μ are

$$\frac{\partial \log L}{\partial \theta} = \frac{3n}{2\theta} - \sum_{i=1}^{n} (x_i - \mu)^2 = 0 \tag{4}$$

and

$$\frac{\partial \log L}{\partial \mu} = -\sum_{i=1}^{n} \frac{2}{(x_i - \mu)} + 2\theta \sum_{i=1}^{n} (x_i - \mu) = 0. \tag{5}$$

As the closed-form solution for θ and μ from Eqs. (4) and (5) is not possible, use of iterative methods to find the numerical solutions is recommended to obtain the MLEs of θ and μ. The existence of unique MLEs of parameters is shown in Theorem 2.1.1.

Theorem 2.1.1. *The unique MLEs of θ and μ exist for $(\mu, \theta) \in [0, x_{(1)}) \times (0, \infty)$.*

Proof. First we show the existence of MLEs of θ and μ, and then their uniqueness. Let $\Omega = (0, \infty) \times (-\infty, x_{(1)})$ and $\Omega_1 = (0, \infty) \times [0, x_{(1)})$. Over the domain of the log-likelihood function $(0, \infty) \times [0, x_{(1)})$, both second derivatives of the log-likelihood function are negative. Hence for a fixed $\theta\,(or\,\mu)$, $log\,L(\theta, \mu)$ is a strictly concave function of μ (or θ). Now,

- For a fixed θ, $\underset{\mu \to -\infty}{lim}$ log $L(\theta, \mu) = -\infty$ and $\lim_{\mu \to x_1}$ log $L(\theta, \mu) = -\infty$.
- For a fixed μ, $\lim_{\theta \to 0}$ log $L(\theta, \mu) = -\infty$. and $\lim_{\theta \to \infty}$ log $L(\theta, \mu) = -\infty$.

Therefore, for a fixed θ (or μ), $\log \mathrm{L}(\theta, \mu)$ is a unimodal function with respect to μ (or θ). For $(\theta_{0,}\ \mu_0) \in\ \Omega$ consider a set

$$\mathcal{A} = \{(\theta, \mu) : (\theta, \mu) \in \Omega, \log L(\theta, \mu) \geq \log\, L(\theta_0, \mu_0)\}.$$

Then $\mathcal{A}$ is a closed, bounded set, and hence compact. Since $log\,L(\theta, \mu)$ is a continuous function of (θ, μ), it has a maximum at some point in $\mathcal{A}$, say (θ_1, μ_1). In the case $\mu_1 \geq 0$, for any $(\theta, \mu) \in \Omega$, note that

$$\log L(\theta_1, \mu_1) > \log L(\theta, \mu_1) > \log L(\theta, \mu).$$

Hence (θ_1, μ_1) is the unique MLE of (θ, μ). Also for $\mu_1 < 0$, and any $(\theta, \mu) \in \Omega_1$,

$$\log L(\theta_1, \mu_1) > \log L(\theta_1, 0) > \log L(\theta_1, \mu) > \log L(\theta, \mu).$$

Hence, $(\theta_1, 0)$ is the MLE of (θ, μ) and it is unique.

2.2. *Moment estimators*

The MEs of θ and μ, namely, $\hat{\theta}_{ME}$ and $\hat{\mu}_{ME}$, respectively, are obtained as

$$\hat{\mu}_{ME} = \bar{X} - \hat{\theta}_{me}^{-1/2} \frac{2}{\Gamma(1/2)} \tag{6}$$

and

$$\hat{\theta}_{ME} = \left(\frac{1}{S^2}\right)\left[\frac{3}{2} - \frac{4}{(\Gamma(1/2))^2}\right] \tag{7}$$

by solving

$$\bar{X} = \mu + \frac{1}{\sqrt{\theta}}\frac{2}{\Gamma(1/2)} \quad \text{and} \quad S^2 = \frac{1}{\theta}\left[\frac{3}{2} - \frac{4}{(\Gamma(1/2))^2}\right]$$

where

$$S^2 = (n-1)^{-1}\sum\nolimits_{i=1}^{n}(x_i - \bar{x})^2$$

2.3. *Least squares estimators*

The least square estimators of the unknown location and scale parameters of the Maxwell distribution can be obtained by following the method used by Swain et al. (1988) to estimate the parameters of beta distribution. Let $G(X_{(j)})$ denotes the distribution function of the ordered random variable $X_{(j)}$. It is well known that

$$E[G(X_{(j)})] = \frac{j}{(n+1)} \quad \text{and} \quad V[G(X_{(j)})] = \frac{j(n-j+1)}{(n+1)^2(n+2)}$$

if $X_1, X_2, \ldots, X_n$ is a random sample of size n from a distribution function $G(\cdot)$. Then the LSEs of θ and μ are obtained by minimizing

$$\sum_{j=1}^{n}\left[G(X_{(j)}) - \frac{j}{n+1}\right]^2$$

with respect to unknown parameters θ and μ. Thus, the LSEs of θ and μ, say, $\hat{\theta}_{LSE}$ and $\hat{\mu}_{LSE}$, respectively, can be obtained by minimizing

$$\sum_{j=1}^{n}\left[\left\{1 - \frac{2}{\sqrt{\pi}}\Gamma\left(\frac{3}{2}, \theta(x_{(j)} - \mu)^2\right)\right\} - \frac{j}{n+1}\right]^2 \tag{8}$$

with respect to θ and μ.

2.4. *Weighted least squares estimators*

The WLSEs of θ and μ, say, $\hat{\theta}_{WLSE}$ and $\hat{\mu}_{WLSE}$, respectively, can be obtained by minimizing

$$\sum_{j=1}^{n} w_j\left[G(X_{(j)}) - \frac{j}{n+1}\right]^2 \tag{9}$$

with respect to θ and μ. The weights used in WLSEs are $w_j = \left[V\left(G(X_{(j)})\right)\right]^{-1}$.

3. Bayes estimation of Maxwell parameters

Consider the squared error loss function (SELF) to estimate the unknown parameters θ and μ of the Maxwell $M(\theta, \mu)$ distribution. Note that estimators using other loss functions can be obtained similarly. If all the parameters of the model are unknown, a joint conjugate prior for the parameters does not exist. Thus, we consider piecewise independent priors. Note that for a known μ, θ has a conjugate gamma prior as follows:

$$g_1(\theta) \propto \theta^{a-1} e^{-b\theta}; \theta > 0, a, b > 0. \tag{10}$$

For μ, we consider a nonproper uniform prior as follows:

$$g_2(\theta\mu) \propto d\mu; \quad 0 < \theta\mu < \infty.$$

The joint posterior distribution of θ and μ, after some simplification, is obtained as follows:

$$\pi(\theta, \mu|Data) = \frac{\theta^{\frac{3n}{2}+a-1} \prod_{i=1}^{n} (x_i - \theta\mu)^2 e^{-\theta\left(b+\sum_{i=1}^{n} (x_i-\mu)^2\right)}}{\int_0^\infty \int_0^{x_{(1)}} \theta^{\frac{3n}{2}+a-1} \prod_{i=1}^{n} (x_i - \mu)^2 e^{-\theta\left(b+\sum_{i=1}^{n} (x_i-\mu)^2\right)} d\mu d\theta} \tag{11}$$

Therefore, the Bayes estimator of some function of θ and μ, say $\zeta(\theta, \mu)$, under SELF, should be the posterior mean

$$\hat{\zeta}(\theta, \mu|Data) = \frac{\int_0^\infty \int_0^{x_{(1)}} \zeta(\theta, \mu) \theta^{\frac{3n}{2}+a-1} \prod_{i=1}^{n} (x_i - \mu)^2 e^{-\theta\left(b+\sum_{i=1}^{n} (x_i-\mu)^2\right)} d\mu d\theta}{\int_0^\infty \int_0^{x_{(1)}} \theta^{\frac{3n}{2}+a-1} \prod_{i=1}^{n} (x_i - \mu)^2 e^{-\theta\left(b+\sum_{i=1}^{n} (x_i-\mu)^2\right)} d\mu d\theta}, \tag{12}$$

provided it is finite.

The estimators are in the form of a ratio of two integrals for which closed-form solutions are unavailable. Therefore, one may apply Lindley's approximation (obtained using some adjustment to the MLE) to derive the Bayes estimator. Even though it is achievable, the associated credible interval is unattainable. Thus, we propose the importance sampling method to obtain the Bayes estimates and associated credible intervals. Chen and Shao (1999), and Kundu and Pradhan (2009) have shown that by proper importance sampling method, a simulation consistent Bayes estimator and the associated credible interval can be constructed. To implement the importance sampling procedure, using

$$k = \left(\frac{3n}{2} + a - 1\right) \left(n x_{(1)}^2 - 2x_{(1)} \sum_{i=1}^{n} x_i + b + \sum_{i=1}^{n} x_i^2\right)^{\frac{3n}{2}+a-1},$$

we rewrite Eq. (11) as

$$\pi(\theta, |\mu|Data) = \frac{f_1(\theta|\mu, Data) f_2(\mu|Data) h(\mu)}{\int_0^\infty \int_0^{x_{(1)}} (\theta|\mu, Data) f_2(\mu|Data) h(\mu) d\mu d\theta}, \tag{13}$$

where

$$f_1(\theta|\mu, Data) = \frac{\left(b + \sum_{i=1}^{n} (x_1 - \mu)^2\right)^{\frac{3n}{2}+a}}{\Gamma\left(\frac{3n}{2} + a\right)} \theta \frac{3n}{2} + a - 1_e$$
$$- \theta\left(b + \sum_{i=1}^{n} (x_i - \mu)^2\right); \theta > 0, \tag{14}$$

$$f_2(\mu|Data) = \frac{k\,2\left(\sum_{i=1}^{n} x_i - n\mu\right)}{\left(n\mu^2 - 2\mu \sum_{i=1}^{n} x_i + b + \sum_{i=1}^{n} x_i^2\right)^{\frac{3n}{2}+a}}; \; 0 < \mu < x_{(1)} \tag{15}$$

and

$$h(\mu) = \begin{cases} \left(\prod_{i=1}^{n} (x_i - \mu)^2\right) / \left(2\left(\sum_{i=1}^{n} x_i - n\mu\right)\right) & \text{for } \mu < x_{(1)} \\ 0 & \text{for } \mu \geq x_{(1)} \end{cases}. \tag{16}$$

Note that $f_1(\theta|\mu, Data)$ is a gamma density function with shape parameter $(3n/2) + a$ and scale parameter $b + \sum_{i=1}^{n} (x_i - \mu)^2$, respectively. Also, for $\mu < x_{(1)}$, $f_2(\mu|Data)$ is a proper density function with an easily invertible distribution function, namely,

$$F_2(\mu|Data) = \frac{\left(nx_{(1)}^2 - 2x_{(1)} \sum_{i=1}^{n} x_i + b + \sum_{i=1}^{n} x_i^2\right)^{\frac{3n}{2}+a-1}}{\left(n\mu^2 - 2\mu \sum_{i=1}^{n} x_i + b + \sum_{i=1}^{n} x_i^2\right)^{\frac{3n}{2}+a-1}}. \tag{17}$$

A simple procedure can be used to generate a random sample from Eq. (17) based on the samples $x_1, x_2, \ldots, x_n$ from $M(\mu, \theta)$ as follows. Generate a random variate u from a uniform (0, 1) distribution. If we let

$$y = \frac{A - \sqrt{A^2 - Bn}}{n},$$

where

$$A = \sum_{i=1}^{n} x_i, \;\; B = \sum_{i=1}^{n} x_i^2 + b - C,$$

and

$$C = \frac{\left(nx_{(1)}^2 - 2x_{(1)} \sum_{i=1}^{n} x_i + b + \sum_{i=1}^{n} x_i^2\right)}{u^{1/\left(\frac{3n}{2}+a-1\right)}},$$

then y is a random variate generated from Eq. (17).

The following procedure is proposed to compute the Bayes estimator of $\zeta(\theta, \mu)$:

- Step 1: Generate μ from $f_2(\mu|Data)$ using Eq. (17).
- Step 2: Generate $\theta|\mu$ from $gamma\left(3n/2 + a, b + \sum_{i=1}^{n} (x_i - \mu)^2\right)$.
- Step 3: Repeat steps 1–2 N times to obtain $(\theta_1, \mu_1), \cdots, (\theta_N, \mu_N)$.
- Step 4: Then a simulation consistent estimator of Eq. (12) can be obtained as

$$\frac{\sum_{i=1}^{N} \zeta(\theta_i, \mu_i) h(\mu_i)}{\sum_{i=1}^{N} h(\mu_i)}. \tag{18}$$

4. Interval estimators for Maxwell parameters

Previously we have proposed several point estimation procedures for the parameters of the Maxwell distribution, and here we describe the interval estimation procedures.

4.1. *A large-sample frequentist approach to interval estimation*

As shown earlier, the MLE of the vector of unknown parameters $\eta = (\theta, \mu)$ cannot be derived in closed form. As it is taxing to derive the exact distributions of the MLEs, the interval estimators are not easy to obtain. However, a large-sample approximation to the probability bounds for Maxwell parameters can be obtained. It is known from Lawless (1982) that the asymptotic distribution of the MLE of η, that is, $\hat{\eta}$, is

$$(\hat{\eta} - \eta) \to N_2(0, I^{-1}(\eta)),$$

where $I^{-1}(\eta)$ is the inverse of the observed information matrix of the unknown parameters $\eta = (\theta, \mu)$. Therefore,

$$\begin{bmatrix} var(\hat{\theta}) & cov(\hat{\theta}, \hat{\mu}) \\ cov(\hat{\mu}, \hat{\theta}) & var(\hat{\mu}) \end{bmatrix} = I^{-1}(\eta) = \begin{bmatrix} -\frac{\partial^2 \log L}{\partial \theta^2} & -\frac{\partial^2 \log L}{\partial \theta \partial \mu} \\ -\frac{\partial^2 \log L}{\partial \mu \partial \theta} & -\frac{\partial^2 \log L}{\partial \mu^2} \end{bmatrix}^{-1} \Bigg|_{(\theta,\mu)=(\hat{\theta},\hat{\mu})}$$

$$= \begin{bmatrix} 3n\big/2\hat{\theta}^2 & -2\sum_{i=1}^{n}(x_i - \hat{\mu}) \\ -2\sum_{i=1}^{n}(x_i - \hat{\mu}) & 2n\hat{\theta} + \sum_{i=1}^{n}\left(2/(x_i - \hat{\mu})^2\right) \end{bmatrix}^{-1} \tag{19}$$

Then the approximate $100(1-\alpha)\%$ confidence intervals for the parameters θ and μ, respectively, are given by

$$\hat{\theta} \pm_{z\alpha/2} \sqrt{Var(\hat{\theta})} \quad \text{and} \quad \hat{\mu} \pm_{z\alpha/2} \sqrt{Var(\hat{\mu})}$$

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

4.2. *Bayes approach to interval estimation*

The HPD credible interval of $\zeta(\theta, \mu)$ using the importance sampling procedure can be constructed as follows. Suppose ζ_ρ is such that $P(\zeta \le \zeta_p | Data) = p$, for $0 < p < 1$. In other words, ζ_ρ is the estimated pth quantile using a frequentist approach. Consider the function

$$g(\theta, \mu) = \begin{cases} 1 \ \text{if} \ \zeta \le \zeta_p \\ 0 \ \text{if} \ \zeta > \zeta_{p.} \end{cases} \tag{20}$$

such that $E(g(\theta,\mu)|Data) = p$. To obtain a simulation-consistent Bayes estimate of ζ_ρ under SELF from the generated sample $\{(\theta_1,\mu_1),\cdots,(\theta_N,\mu_N)\}$, let weights be

$$w_1 = \frac{h(\mu_i)}{\sum_{i=1}^{N} h(\mu_i)},$$

and $\zeta_i = \zeta(\theta_i,\mu_i)$ for $i = 1,2,\ldots,N$. Order pairs (ζ_i, w_i), $i = 1,2,\ldots,N$ with respect to values of ζ to give $\{(\zeta_{(1)}, w_{[1]}),\cdots,(\zeta_{(N)}, w_{[N]})\}$. That is, $\zeta_{(1)}<\zeta_{(2)}<\cdots<\zeta_{(N)}$, and $w_{[i]}$ is the weight associated with the ith ordered value of ζ, namely, $\zeta_{(i)}$. It is worth mentioning that these weights do not play any role in ordering of pairs (ζ_i, w_i). Then a simulation-consistent Bayes estimate of ζ_p can be obtained as $\hat{\zeta}_p = \zeta_{(N_p)}$, where

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

A $100(1-\alpha)\%$ credible interval for ζ can be obtained as $(\hat{\zeta}_\delta, \hat{\zeta}_{\delta+1-\alpha})$ where

$$\delta = w_{[1]}, w_{[1]} + w_{[2]},\cdots,\sum_{i=1}^{N_{1-\alpha}} w_{[i]}.$$

Therefore, a $100(1-\alpha)\%$ HPD credible interval of ζ is given by $(\hat{\zeta}_{\delta*}, \hat{\zeta}_{\delta*+1-\alpha})$, where δ^* is that value of δ for which the following inequality holds for all values of δ:

$$\left(\hat{\zeta}_{\delta*+1-\alpha} - \hat{\zeta}_{\delta*}\right) \le \left(\hat{\zeta}_{\delta+1-\alpha} - \hat{\zeta}_{\delta}\right).$$

5. A comparison of different estimators

Here we compare the effectiveness of different estimation methods discussed earlier for Maxwell parameters. In these comparisons, bias and the mean squared errors are used as measures of effectiveness of the estimation method. A simulation study was designed for this purpose. All computations were performed using Intel Core 2 Quad Processor and all programs were coded in R software.

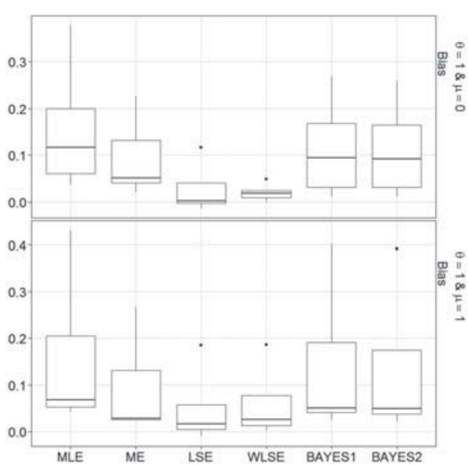
Two different priors were used to compare the Bayes estimates: (i) a close to a noninformative prior (Prior 0) with $a = b = 0.001$, where small values close to zero were used for a and b so that the posterior always becomes integrable, and (ii) an informative prior (Prior 1) with $a = 2$ and $b = 1$. We considered two conditions: (1) $\theta = 1$, $\mu = 0$, and (2) $\theta = 1$, $\mu = 1$. In total, 10,000 independent random samples, of sizes n = 10, 20, 40, 60, and 100, were obtained using the procedure described in section 3. Six different estimates of θ and μ were computed from each sample, namely, MLE, ME, LSE, WLSE, BAYES1, and BAYES2. Here BAYES1 refers to the Bayes estimate using the noninformative prior and BAYES2 refers to the Bayes estimate using the informative prior. The bias, that is, $\left(E(\hat{\theta}) - \theta\right)$, and MSE were computed for six different estimators for θ and μ using samples of five different sizes. The bootstrap confidence intervals based on percentiles (Efron 1982) were also computed using the following algorithm:

(1) From the sample $x_1, x_2, \ldots, x_n$, compute the estimators $\hat{\theta}$ and $\hat{\mu}$.
(2) Generate a bootstrap sample $x_1^*, x_2^*, \ldots, x_n^*$, using $\hat{\theta}$ and $\hat{\mu}$ as parameters.
(3) Compute the estimators $\hat{\theta}_1^*$ and $\hat{\mu}_1^*$ based on the bootstrap sample.
(4) Repeat steps 2 and 3 N times, and compute $\hat{\theta}_i^*$ and $\hat{\mu}_i^*$ for $i = 1, 2, \ldots, N$.
(5) Arrange $\hat{\theta}_i^*$ in ascending order and obtain θ_U and θ_L the upper and the lower limits, respectively, of a $100(1-\gamma)\%$ bootstrap confidence interval for θ, using $(\gamma/2)$-th and $(1-\gamma/2)$-th percentiles of ordered values of $\hat{\theta}_i^*$.
(6) Arrange $\hat{\mu}_i^*$ in ascending order and obtain μ_U and μ_L the upper and the lower limits, respectively, of a $100(1-\gamma)\%$ bootstrap confidence interval for μ, using $(\gamma/2)$-th and $(1-\gamma/2)$-th percentiles of ordered values of $\hat{\mu}_i^*$.

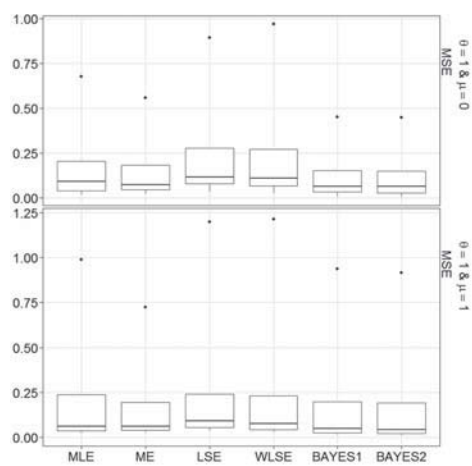
The Bayes estimates in each run are computed based on 10,000 importance samples.

Performance of different estimators was compared using bias, MSE, and the coverage percentage of bootstrap confidence intervals. For parameters θ and μ, respectively, Figures 1 and 3 provide comparison of distributions of bias and MSE for six different estimators studied using samples of size 10, 20, 40, 60, and 100. Similarly, Figures 2 and 4 provide information about the effect of sample size on bias and MSE for θ and μ, respectively. The coverage percentages of bootstrap confidence intervals for θ and μ using samples of size 10, 20, 40, 60, and 100 for these six methods are summarized in Figures 5 and 6.

From Figure 1 and Figure 3, comparison of performance of different estimators leads to the conclusion that overall LSE resulted in lower bias for θ, whereas ME and WLSE methods resulted in lower bias for μ, compared to the remaining methods studied here. All estimators of θ are biased and mostly right-skewed, but their bias distributions have different spreads. On the other hand, the distributions of MSE of θ seem to have fairly similar spreads in almost all cases. Overall, LSE and WLSE resulted in slightly less biased estimates of θ but with slightly higher MSE than the other estimators. All

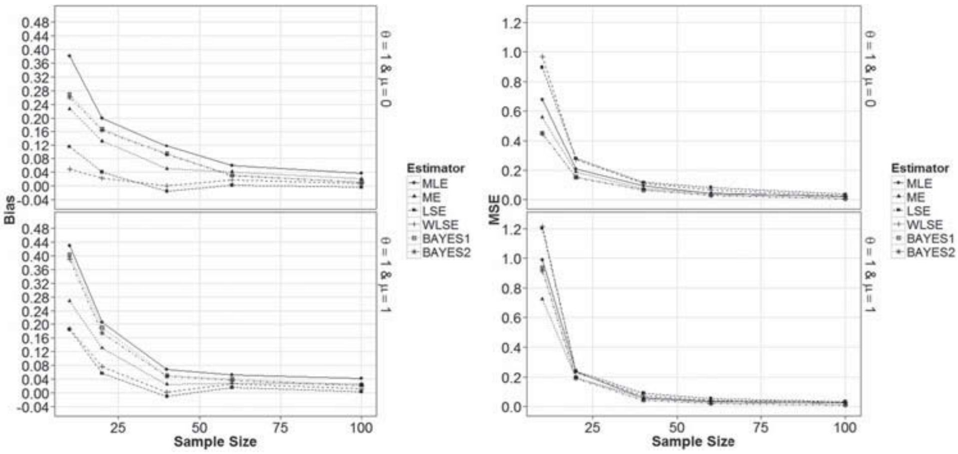


(a) Bias



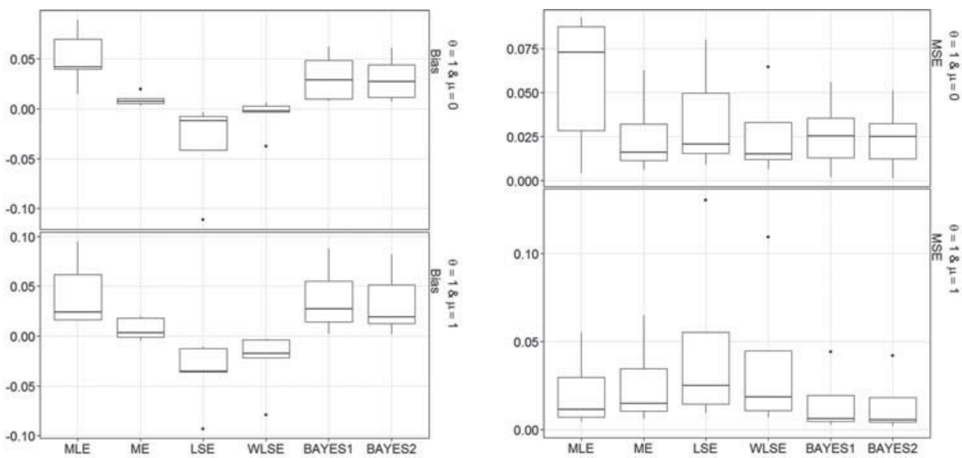
(b) MSE

Figure 1. Comparison of distributions of (a) average bias and (b) MSE for the estimates of θ using different estimation methods for $\theta = 1$ and $\mu = 0$ and 1.



(a) Bias (b) MSE

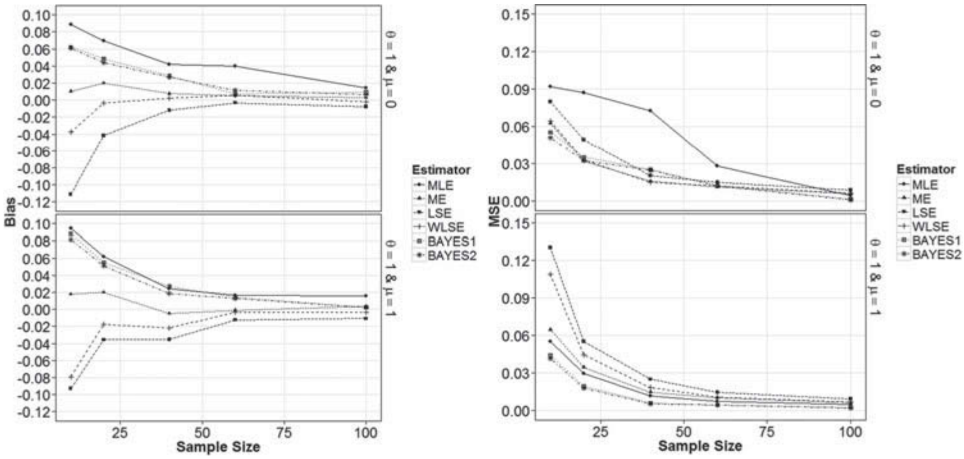
Figure 2. Comparison of (a) average bias and (b) MSE for the estimates of θ using different estimation methods as a function of sample sizes ($n = 10$, 20, 40, 60, and 100) considered in the simulation study for $\theta = 1$ and $\mu = 0$ and 1.



(a) Bias (b) MSE

Figure 3. Comparison of distributions of (a) average bias and (b) MSE for the estimates of using different estimation methods for $\theta = 1$ and $\mu = 0$ and 1 for the sample sizes considered in the simulation study.

estimators of μ are also biased but with different patterns; some tend to overestimate while the others tend to underestimate μ. The shapes and spreads of distributions of MSE(μ) are also different from estimator to estimator, as well as between two different values of μ studied. Overall, ME and WLSE resulted in lowest amount of bias among six estimators studied, but they do not necessarily have the lowest MSE. The Bayes estimators for both θ and μ resulted in slightly larger bias than other estimators except for MLE.



(a) Bias (b) MSE

Figure 4. Comparison of (a) average bias and (b) MSE for the estimates of μ using different estimation methods as a function of sample sizes considered in the simulation study for $\theta = 1$ and $\mu = 0$ and 1.

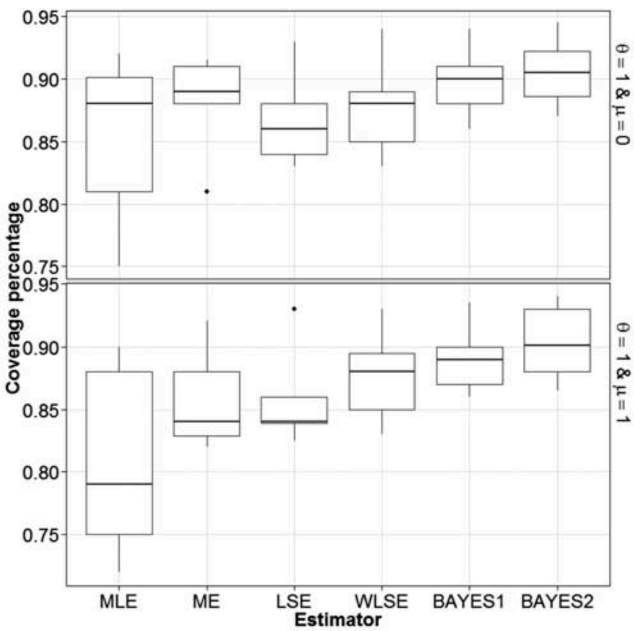


Figure 5. Comparison of distributions of coverage percentages for the 95% bootstrap confidence intervals based on different estimation procedures for θ.

Although Bayes estimates and MLE have similar distributions, the MLE has larger MSE. The performance of the Bayes estimator with informative and noninformative priors is almost similar.

Figures 2 and 4 show that all estimators of θ and μ are biased for small sample sizes with fairly large MSE. However, for the configurations studied, both bias and MSE approach zero rapidly as sample size increases. By sample size 40, estimators of θ are

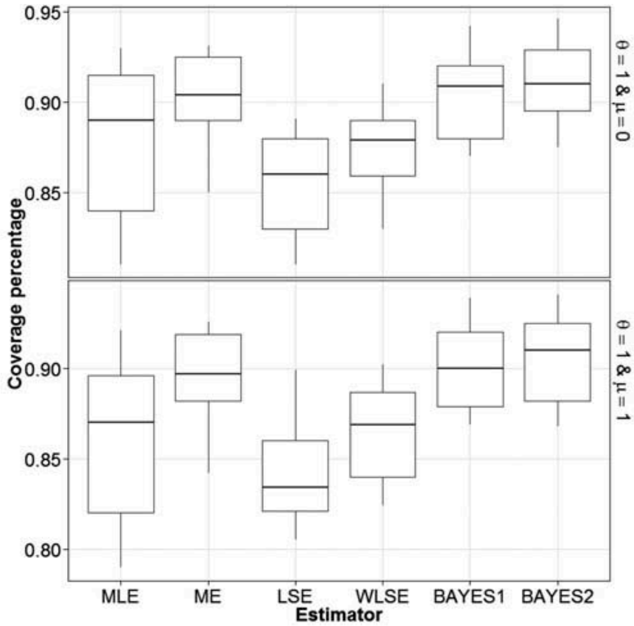


Figure 6. Comparison of distributions of coverage percentages for the 95% bootstrap confidence intervals based on different estimation procedure for μ.

essentially unbiased for all practical purposes. However, the estimators of μ need slightly larger sample sizes to become almost unbiased. A similar trend is observed for MSE with respect to the sample size used.

Figures 5 and 6 present the performance of estimators studied based on coverage percentage of 95% bootstrap confidence intervals for θ and μ, respectively. It is clear that MLEs produced lower coverage compared to the remaining estimators. Both Bayes estimators performed consistently well along with MEs. Similar patterns were observed from smaller to larger sample sizes.

Example 1. To illustrate the performance of estimation procedures discussed here, a data set obtained from Dumonceaux and Antle (1973) is used. This data set gives the maximum flood levels (in millions of cubic feet per second per 4-year period cycle) of the Susquehenna River at Harrisburg, PA, for the period of 1890–1969.

From summary statistics in Table 1, we see that the distribution is positively skewed.

To check the shape of the empirical hazard function, the scaled TTT (total time on test) transform plot (Aarset 1987) is presented in Figure 7. Its concave shape indicates that the hazard function of the distribution will be an increasing function. Therefore, the Maxwell distribution is a possible option along with gamma, lognormal, and Rayleigh distributions

Table 1. Summary statistics of the flood level data.

Sample size	Minimum	Q_1	Median	Q_3	Maximum	Mean	SD
20	0.2650	0.3345	0.4070	0.4578	0.7400	0.4232	0.1253

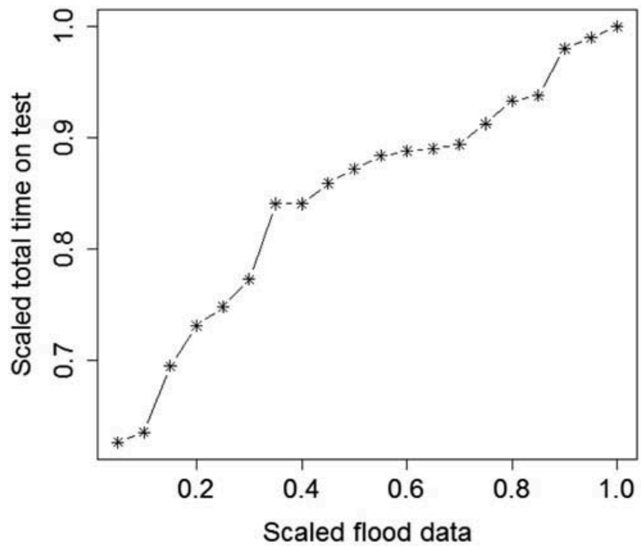


Figure 7. Scaled total time on test (TTT) transform plot of flood level data.

to analyze this data set. All four of these distributions were fitted to the data. The resulting goodness-of-fit test outcomes are listed in Table 2. From Table 2 we see that the Maxwell model is a possible option along with gamma, log-normal, and Raleigh distributions. Although all of these distributions provide acceptable fits, the Maxwell distribution resulted in the least opposing evidence, followed by the log-normal distribution.

Assuming the underlying distribution is a two-parameter Maxwell, the estimates of θ and μ were obtained using different methods presented earlier, which are listed in Table 3.

- MLE: To obtain MLEs, the contour plot of log-likelihood function for the flood-level data was used (see Figure 8). The maximum of the log-likelihood function is

Table 2. Goodness-of-fit test results.

	KS statistic	p value
Gamma (three-parameter)	0.176	0.530
Lognormal (two-parameter)	0.155	0.713
Maxwell (two-parameter)	0.152	0.742
Rayleigh (two-parameter)	0.179	0.506

Table 3. Estimates of Maxwell parameters from flood-level data.

Estimation method	$\hat{\theta}$	$\hat{\mu}$
MLE	18.2109	0.1635
ME	14.4529	0.1263
LSE	20.6898	0.1618
WLSE	19.6799	0.1571
Bayes	19.4258	0.1698
95% confidence interval	(7.0716, 29.3501)	(0.0847, 0.2421)
95% credible interval	(17.1876, 22.4571)	(0.1657, 0.2145)

indicated by a dot in the innermost contour. The coordinates of this point provide the MLEs of θ and μ, that is, $\hat{\theta} \approx 18.2109$ and $\hat{\mu} \approx 0.1634$, respectively.

- ME: The MEs of θ and μ, respectively, were computed using Eqs. (6) and (7).
- LSE: The LSEs of θ and μ, respectively, were obtained by minimizing Eq. (8) with respect to θ and μ by using R software.
- WLSE: The WLSEs of θ and μ, respectively, were obtained by minimizing Eq. (9) with respect to θ and μ by using R software.
- Bayes estimators: For computing Bayes estimates, we have used the hyperparameter values of (10) as $a = b = 0.001$, so that the posterior distribution becomes integrable. In total, 10,000 importance samples were used.

From Table 3, we observe that all procedures except ME resulted in fairly similar estimates. Both $\hat{\theta}_{ME}$ and $\hat{\mu}_{ME}$ are considerably lower than estimates using other methods. The confidence intervals for both θ and μ are wider than the credible intervals.

Figure 9a depicts that different estimation procedures discussed in this study for Maxwell model resulted in CDFs similar to each other when applied to the flood data set. Moreover, we also observe a minor overestimation in the upper tail of the empirical CDF for all methods.

Figure 9b presents the performance of Maxwell distribution as compared to some competing models fitted to the flood data (listed in Table 2). Parameters were estimated using the MLE procedure. As seen earlier, goodness-of-fit scores of log-normal and Maxwell models are very competitive. Both of these models perform quite similarly along with the Rayleigh model whereas the Gamma model exhibits some overestimation in the beginning and underestimation later for the flood data.

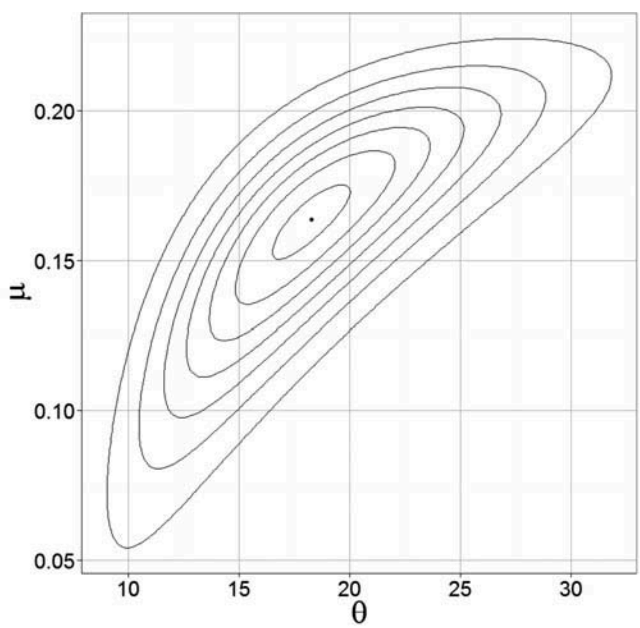


Figure 8. Contour plot of log likelihood function for different values of θ and μ for flood-level data.

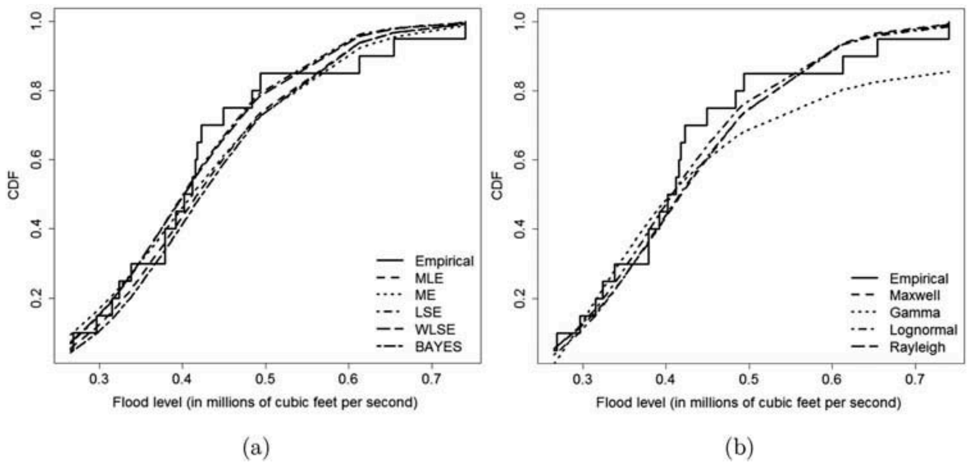


Figure 9. (a) The empirical and fitted CDF of Maxwell model (using different estimation procedures) of the flood data. (b) The empirical and fitted CDF of different models (based on MLE) of the flood data.

6. Conclusion

In this article we have considered several estimation techniques for estimating the unknown parameters of a two-parameter Maxwell distribution, namely, the maximum likelihood estimators, the method of moments estimators, the least square estimators, the weighted least square estimators, and the Bayes estimators. Results of a simulation study to compare these methods are presented, which show that the Bayes estimators under informative prior work well in terms of biases and mean squared errors; even the performance of Bayes estimators under the informative prior is as competitive as that under a noninformative prior. As the choice of hyperparameters of the prior distribution plays a significant role in Bayesian methodology, users are advised to be cautious about using an informative prior in real-life applications. The performance of maximum likelihood estimators is fairly reasonable and competitive. We recommend using the maximum likelihood estimators or the Bayes estimators in practice. We also compared this model with some existing competing models and the Maxwell model performed reasonably well.

Acknowledgments

The authors thank the reviewers and the editors for their comments, which helped improve the article.

References

Aarset, M. V. 1987. How to identify a bathtub shaped hazard rate? *IEEE Transactions on Reliability* 36:106–108.

Adamidis, K., T. Dimitrakopoulou, and S. Loukas. 2005. On an extension of the exponential-geometric distribution. *Statistics and Probababilty Letters* 73:259–269.

Alkasasbeh, M. R., and M. Z. Raqab. 2009. Estimation of the generalized logistic distribution parameters: Comparative study. *Statistical Methodology* 6:262–279.

Arnold, B. C., N. Balakrishnan, and H. N. Nagaraja. 1998. *A first course in order statistics*. New York, NY: Wiley.

Barreto-Souza, W., A. L. de Morais, and G. M. Cordeiro. 2011. The Weibull-geometric distribution. *Journal of Statistical Computation and Simulation* 81:645–657.

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

Chattopadhyay, S., C. A. Murthy, and S. K. Pal, 2014. Fitting truncated geometric distributions in large scale real world net works. *Theoretical Computer Science* 551:2238.

Chaturvedi, A., and U. Rani. 1998. Classical and Bayesian reliability estimation of the generalized Maxwell failure distribution. *Journal of Statistical Research* 32:113–120.

Chen, M. H., and Q. M. Shao. 1999. Monte Carlo estimation of Bayesian credible and HPD intervals. *Journal of Computational and Graphical Statistics* 8 (1): 69–92.

Dey, S., T. Dey, and D. Kundu. 2014. Two-parameter Rayleigh distribution: Different methods of estimation. *American Journal of Mathematical and Management Sciences* 33:55–74.

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

Dey, S., and S. S. Maiti. 2010. Bayesian estimation of the parameter of Maxwell distribution under different loss functions. *Journal of Statistical Theory and Practice* 4 (2):279–287.

Dumonceaux, R., and C. E. Antle. 1973. Discrimination between the lognormal and Weibull distribution. *Technometrics* 15:923–926.

Efron, B. 1982. The jackknife, the bootstrap, and other resampling plans. *Society of Industrial and Applied Mathematics CBMS-NSF Monographs* 38.

Feller, W. 1971. *An introduction to probability theory and its applications*, Vol. 2, 2nd ed. New York, NY: Wiley.

Gupta, R. D., and D. Kundu. 1999. Generalized exponential distribution. *Australian and New Zealand Journal of Statistics* 41:173–188.

Hartigan, J. 1964. Invariant prior distributions. *Annals of Mathematical Statistics* 35:836–845.

Howlader, H. A. and A. Hossain. 1998. Bayesian prediction intervals for Maxwell parameters. *Metron* LVI (1–2):97–105.

Johnson, N., S. Kotz, and N. Balakrishnan. 1994 *Continuous univariate distributions*, Vol. 1, 2nd ed. New York, NY: Wiley.

Krishna, H., and M. Malik. 2009. Reliability estimation in Maxwell distribution with type-II censored data. *International Journal of Quality & Reliability Management* 26 (2):184–195.

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

Kumaraswamy, P. 1976. Sinepower probability density function. *Journal of Hydrology* 31:181–184.

Kumaraswamy, P. 1978. Extended sinepower probability density function. *Journal of Hydrology* 37:81–89.

Kundu, D., and M. Z. Raqab. 2005. Generalized Rayleigh distribution: Different methods of estimations. *Computational Statistics and Data Analysis* 49:187–200.

Kundu, D., and B. Pradhan. 2009. Bayesian inference and life testing plans for generalized exponential distribution. *Science in China, Series A: Mathematics* 52:1373–1388.

Kus, C. 2007. A new lifetime distribution. *Computational Statistics and Data Analysis* 51:4497–509.

Lawless, J. F. 1982. *Statistical models and methods for lifetime data*. New York, NY: John Wiley & Sons.

Maxwell, J. C. 1867. On the dynamical theory of gases. *Philosophical Transactions of the Royal Society of London* 157:49–88.

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

Rényi, A. 1961. On measures of entropy and information. *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, Vol. I, 547–561. Berkeley, CA: University of California Press.
Ristic, M. M., and N. Balakrishnan. 2012. The gamma-exponentiated exponential distribution. *Journal of Statistical Computation and Simulation* 82 (8):1191–1206.
Shaked, M., and J. G. Shanthikumar. 1994. *Stochastic orders and their applications*. Boston, MA: Academic Press.
Shannon, C. E. 1951. Prediction and entropy of printed English. *Bell System Technical Journal* 30:50–64.
Swain, J. J., S. Venkataraman, and J. R. Wilson. 1988. Least squares estimation of distribution functions in Johnson's translation system. *Journal of Statistical Computation and Simulation* 29:271–297.
Teimouri, M., S. M. Hoseini, and S. Nadarajah. 2013. Comparison of estimation methods for the Weibull distribution. *Statistics* 47 (1):93–109.
Tyagi, R. K., and S. K. Bhattacharya. 1989a. Bayes estimation of the Maxwell's velocity distribution function. *Statistica* XLIX (4): 563–567.
Tyagi, R. K., and S. K. Bhattacharya. 1989b. A note on the MVU estimation of reliability for the Maxwell failure distribution. *Estadistica* 41 (137):73–79.

Appendix

A few structural properties of the Maxwell distribution, specifically, stochastic ordering, entropies, and order statistics are described here.

A.1. Stochastic ordering

Stochastic ordering of positive continuous random variables is an important tool for judging the comparative behavior. There are different types of stochastic orderings, which are useful in ordering random variables in terms of different properties. Here we consider four different stochastic orders, namely, the usual, the hazard rate, the mean residual life, and the likelihood ratio order for two independent Maxwell random variables under a restricted parameter space.

If X and Y are independent random variables with CDFs F_X and $F_{Y,}$ respectively, then X is said to be smaller than Y in the

- Usual stochastic (*st*) order (i.e, $X \leq_{st} Y$) if $F_X(x) \geq F_Y(y)$ for all x.
- Hazard rate (*hr*) order (i.e., $X \leq_{hr} Y$) if $h_X(x) \geq h_Y(y)$ for all x.
- Mean residual life (*mrl*) order (i.e., $X \leq_{mrl} Y$) if $m_X(x) \geq m_Y(y)$ for all x.
- Likelihood ratio (*lr*) order (i.e., $X \leq_{lr} Y$) if $f_X(x)/f_Y(y)$ decreases in x.

Theorem A.1.1 shows that the Maxwell distribution is ordered with respect to the strongest, that is, the likelihood ratio ordering under the restricted space of

$$\{\theta_1 = \theta_2 = \theta \text{ and } x > \mu_2 > \mu_1\} \quad \text{or} \quad \{\theta_1 > \theta_2 \text{ and } \mu_1 = \mu_2 = \mu\}. \tag{21}$$

This shows the flexibility of two-parameter Maxwell distribution.

Theorem A.1.1. *If* X *and* Y *are two independent two-parameter Maxwell distributions, then all four stochastic orderings exist under the restricted space of Eq. (21).*

Proof. Let $X \sim M(\theta_1, \mu_1)$ and $Y \sim M(\theta_2, \mu_2)$. The log-likelihood ratio of X to Y is

$$\log\frac{f_X(x)}{f_Y(x)} = \log(\theta_1/\theta_2)^{1.5} + 2\left[\log(x-\mu_1) - \log(x-\mu_2)\right] - \theta_1(x-\mu_1)^2 + \theta_2(x-\mu_2)^2$$

and the derivative of log-likelihood ratio with respect to x is

$$\frac{d}{dx}\log\frac{f_X(x)}{f_Y(x)} = -2\left[\frac{\mu_2-\mu_1}{(x-\mu_1)(x-\mu_2)}\right] - 2\left[x(\theta_1-\theta_2) - \theta_1\mu_1 + \theta_2\mu_2\right].$$

- Consider $\theta_1 = \theta_2 = \theta$. Then the derivative of log-likelihood ratio with respect to x is negative for $x > \mu_2 > \mu_1$ and positive for $x > \mu_1 > \mu_2$. This implies that $X \leq_{lr} Y$ provided $x > \mu_2 > \mu_1$.
- Now consider $\mu_1 = \mu_2 = \mu$. Then the derivative of log-likelihood ratio with respect to x is negative for $\theta_1 > \theta_2$ and positive for $\theta_1 < \theta_2$. This implies that $X \leq_{lr} Y$ provided $\mu_1 = \mu_2 = \mu$ and $\theta_1 > \theta_2$.

Hence we conclude that $X \leq_{lr} Y$ under the restricted space of Eq. (21). Shaked and Shanthikumar (1994) have shown that the following relation exists among four stochastic orderings of distributions listed earlier:

$$\begin{array}{c} X \leq_{lr} Y \Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{mrl} Y \\ \Downarrow \\ X \leq_{st} Y \end{array} \tag{22}$$

Therefore, using Eq. (22) and the result of likelihood ratio ordering, we can conclude that in addition to the likelihood ratio ordering, the usual stochastic, the hazard rate, and the mean residual life ordering also exists, for the two-parameter Maxwell distributions under the restricted parameter space of Eq. (21).

A.2. *Entropies*

There are many different entropy measures available in the literature. Two popular entropy measures are the Rényi entropy and Shannon entropy. Entropy of a random variable X is a measure of variation of the uncertainty or disorder in a population. For a two-parameter Maxwell distribution, the Rényi entropy (as defined by Rényi [1961]) is given by

$$\Upsilon_\eta(X) = \frac{1}{1-\eta}\log\left[\frac{2^{2\eta-1}\theta^{0.5(\eta-1)}}{\pi^{\frac{\eta}{2}}\eta^{\eta+0.5}}\Gamma(0.5+\eta)\right] \tag{23}$$

where $\eta > 0$ and $\eta \neq 1$. The Shannon entropy (as defined by Shannon [1951]) for the two-parameter Maxwell distribution is given by

$$E[-\log f(x)] = \gamma + 0.5(\log(\pi) - \log(\theta) - 1) \tag{24}$$

where γ is an Euler gamma constant.

A.3. *Order statistics*

Let $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ denote the order statistics of a random sample $X_1, X_2, \ldots, X_n$ from the continuous population with the PDF in Eq. (1); then the PDF of the k^{th} order statistics $X_{(k)}$, following the results from Arnold et al. (1998), is given by

$$f_{X_{(k)}}\left(x_{(k)}\right) = \frac{n!\theta^{1.5}}{(k-1)!(n-k)!}\sum_{m=0}^{n-k}\binom{n-k}{m}(-1)^m\frac{2^{k+m+1}}{\pi^{(k+m)/2}} \times \left[\Gamma\left(1.5\,,\theta\left(x_{(k)}-\mu\right)^2\right)\right]^{k-1+m}\left(x_{(k)}-\mu\right)^2 e^{-\theta\left(x_{(k)}-\mu\right)^2}.$$