Parameter estimation for the two-parameter Maxwell distribution under complete and censored samples

Authors: Talha Arslan
– Department of Econometrics, Van Yuzuncu Yil University, Turkey (mstalhaarslan@yyu.edu.tr)

Sukru Acitas
– Department of Statistics, Eskisehir Technical University, Turkey (sacitas@eskisehir.edu.tr)

Birdal Senoglu
– Department of Statistics, Ankara University, Turkey (senoglu@science.ankara.edu.tr)

Abstract:

- The Maxwell distribution is one of the basic distributions in Physics besides being popular in Statistics for modeling lifetime data. This paper considers the parameter estimation of the Maxwell distribution via modified maximum likelihood (MML) methodology for both complete and censored samples. The MML estimators for the location and scale parameters of the Maxwell distribution have explicit forms and they are robust against the plausible deviations from the assumed model. A Monte Carlo simulation study is conducted to compare the performances of the MML estimators with the corresponding maximum likelihood (ML), least squares (LS) and method of moments (MoM) estimators.

Key-Words:

- *Efficiency; Maxwell distribution; Modified Likelihood; Monte Carlo Simulation; Newton-Raphson; Type-II censoring.*

AMS Subject Classification:

- 62F10, 62F12, 62F35, 62N02, 62P30.

1. INTRODUCTION

The Maxwell distribution is widely used in many problems especially in Physics. For example, the speed of molecules in thermal equilibrium is modelled by using the Maxwell distribution (Maxwell, 1860; Mathai and Princy, 2017). Also note that there is a lot of literature about the Maxwell distribution in Statistics. It was firstly used by Tyagi and Battacharya (1989a,b) for modeling the lifetime data. They used Bayes method to estimate the scale parameter of the distribution and obtain the minimum variance unbiased estimator for the reliability function. Dey and Maiti (2010) obtained the Bayes estimators of the scale parameter of the Maxwell distribution under various different loss functions. Kazmi et al. (2011) obtained the maximum likelihood (ML) estimators of the location and scale parameters of the mixture of the Maxwell distribution under Type-I censoring. Al-Baldawi (2013) compared the efficiency of the ML estimator of the scale parameter of the Maxwell distribution with the corresponding Bayes estimator. Hossain and Huerta (2016) used the Maxwell distribution in analysing the different data sets taken from the literature. Li (2016) obtained the estimators of the scale parameter of the Maxwell distribution using the Minimax, Bayesian and ML methods. Fan (2016) considered the Bayesian method to estimate the loss and risk function for the scale parameter of the Maxwell distribution. Dey et al. (2016) obtained estimators of the location and scale parameters of the Maxwell distribution via different estimation methods. See also Arslan et al. (2017), where the modified maximum likelihood (MML) estimators for the location and scale parameters of the Maxwell distribution are obtained.

The ML methodology is used to obtain the estimators of the parameters of the Maxwell distribution in most of the studies. However, the ML estimators of the location and scale parameters of the Maxwell distribution cannot be obtained explicitly. Therefore, iterative methods should be used. It is known that using iterative methods causes various problems such as (i) non-convergence of iterations (ii) convergence to multiple roots and (iii) convergence to the wrong root; see e.g. Barnett (1966), Puthenpura and Sinha (1986), and Vaughan (1992).

The motivation of this study is to obtain the explicit estimators for the location and scale parameters of the Maxwell distribution. For this purpose, Tiku's (1967, 1968) MML methodology is used. The MML estimators are formulated for both complete and censored samples. An extensive Monte-Carlo (MC) simulation study is carried out to compare performances of the MML estimators with the well-known and widely-used ML, least squares (LS) and method of moments (MoM) estimators.

The rest of the paper is organized as follows. Maxwell distribution is reviewed in Section 2. Section 3 is reserved to the parameter estimation methodologies. The results of the MC simulation study are presented in Section 4. The ML and MML estimators are given under Type-II censoring scheme in Section 5. In Section 6, two real data sets are analyzed to show the implementation of the proposed methodology. The paper ends with some concluding remarks.

2. MAXWELL DISTRIBUTION

Traditionally, the probability density function (pdf) of the Maxwell distribution is given by

$$f(v) = 4\pi \left(\frac{m}{\pi 2kT}\right)^{3/2} v^2 \exp\left\{-\left(\frac{m}{2kT}v^2\right)\right\}, v > 0 \tag{2.1}$$

where m is the molecular weight in kg/mol, T is the temperature in Kelvin, k is the constant J/K and v denotes the speed of the molecule. If the reparametrization $\sigma = \sqrt{2kT/m}$ is used and a location parameter μ is added into the equation (2.1), then the resulting distribution is called as two-parameter Maxwell distribution.

The pdf and the corresponding cumulative distribution function (cdf) of the two-parameter Maxwell distribution are given by

$$f(x;\mu,\sigma) = \frac{4}{\sigma\Gamma(1/2)} \left(\frac{x-\mu}{\sigma}\right)^2 \exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^2\right\}; \mu \le x \le \infty, \sigma \ge 0 \tag{2.2}$$

and

$$F(x;\mu,\sigma) = \frac{1}{\Gamma(3/2)}\Gamma\left[\left(\frac{x-\mu}{\sigma}\right)^2, 3/2\right], \tag{2.3}$$

respectively. Here, μ is the location parameter and σ is the scale parameter. Also, $\Gamma(\cdot)$ and $\Gamma(\cdot,\cdot)$ stand for the gamma and incomplete gamma functions, respectively. See Figure 1 where the plots of the Maxwell distribution are illustrated for certain values of σ.

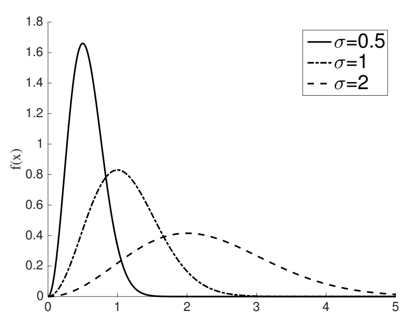

Figure 1: Plots of the Maxwell distribution for certain values of σ.

In the rest of the paper, we use the term Maxwell distribution instead of two-parameter Maxwell distribution for the sake of simplicity.

3. PARAMETER ESTIMATION UNDER COMPLETE SAMPLES

In this section, brief descriptions of the ML, MML, MoM and LS methodologies are provided.

3.1. The ML method

Let $X_1, X_2, \ldots, X_n$ be a random sample from the Maxwell distribution. Then, the log-likelihood ($\ln L$) function can be written as follows:

$$\ln L = n \ln C - n \ln \sigma + 2 \sum_{i=1}^{n} \ln z_i - \sum_{i=1}^{n} z_i^2 \tag{3.1}$$

where $C = 4/\Gamma(1/2)$ and $z_i = (x_i - \mu)/\sigma$ $(i = 1, 2, \ldots, n)$. The ML estimates of the parameters μ and σ are obtained as solutions of the following likelihood equations:

$$\frac{\partial \ln L}{\partial \mu} = -\frac{2}{\sigma} \sum_{i=1}^{n} g(z_i) + \frac{2}{\sigma} \sum_{i=1}^{n} z_i = 0 \tag{3.2}$$

and

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} - \frac{2}{\sigma} \sum_{i=1}^{n} z_i g(z_i) + \frac{2}{\sigma} \sum_{i=1}^{n} z_i^2 = 0 \tag{3.3}$$

where $g(z) = z^{-1}$. Equations (3.2) and (3.3) cannot be solved explicitly since they contain the nonlinear $g(z) = z^{-1}$ function. In this study a Newton-Raphson (NR) method is utilized to obtain the solutions of Equations (3.2) and (3.3) simultaneously. The Hessian matrix,

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 \ln L}{\partial \mu^2} & \frac{\partial^2 \ln L}{\partial \mu \partial \sigma} \\ \frac{\partial^2 \ln L}{\partial \sigma \partial \mu} & \frac{\partial^2 \ln L}{\partial \sigma^2} \end{bmatrix}, \tag{3.4}$$

is used in the NR method. The elements of the Hessian matrix and Fisher Information matrix ($\mathbf{I}$) are provided in the Appendix for the Maxwell distribution.

The following equations are used in the NR method to solve the likelihood equations in (3.2) and (3.3):

$$\begin{bmatrix} \frac{\partial^2 \ln L}{\partial \mu^2}(\mu^k, \sigma^k) & \frac{\partial^2 \ln L}{\partial \mu \partial \sigma}(\mu^k, \sigma^k) \\ \frac{\partial^2 \ln L}{\partial \sigma \partial \mu}(\mu^k, \sigma^k) & \frac{\partial^2 \ln L}{\partial \sigma^2}(\mu^k, \sigma^k) \end{bmatrix} \begin{bmatrix} \Xi \mu^k \\ \Xi \sigma^k \end{bmatrix} = \begin{bmatrix} \frac{\partial \ln L}{\partial \mu}(\mu^k, \sigma^k) \\ \frac{\partial \ln L}{\partial \sigma}(\mu^k, \sigma^k) \end{bmatrix}, \tag{3.5}$$

where k denotes the iteration number and Ξ stands for the incremental values. See also Arslan and Senoglu (2018), where a similar algorithm scheme has already been used for the one-way ANOVA model under Jones and Faddy's skew t distribution.

3.2. The MML method

As mentioned in the subsection 3.1, the ML estimators of the location and scale parameters cannot be obtained in closed forms because of the nonlinear function $g(\cdot)$ in Equations (3.2) and (3.3). We here propose to use non-iterative MML methodology developed by Tiku (1967, 1968) to avoid the computational difficulties and/or problems mentioned in Section 1. The MML methodology also allows us to obtain closed forms of the estimators. There are three steps to obtain the MML estimators of the location parameter μ and scale parameter σ. They are given step by step as follows:

Step 1 Standardized observations $z_i = (x_i - \mu)/\sigma$ $(i = 1, 2, \ldots, n)$ are ordered in ascending way, i.e. $z_{(1)} \leq z_{(2)} \leq \ldots \leq z_{(n)}$.

Step 2 The ordered observations are incorporated into likelihood equations, since complete sums are invariant to ordering, i.e. $\sum_{i=1}^{n} h(z_i) = \sum_{i=1}^{n} h(z_{(i)})$, where $h(\cdot)$ is any function.

Step 3 $g(z_{(i)})$ is linearized around the expected values of the standardized ordered observations , i.e. $t_{(i)} = E(z_{(i)})$, by using the first two terms of Taylor series expansion:

$$g\left(z_{(i)}\right) \cong \alpha_i - \beta_i z_{(i)}, \quad (i = 1, ..., n)\,. \tag{3.6}$$

After incorporating Equation (3.6) into the likelihood equations, we obtain the following modified likelihood equations:

$$\frac{\partial \ln L^*}{\partial \mu} = -\frac{2}{\sigma} \sum_{i=1}^{n} \left(\alpha_i - \beta_i z_{(i)}\right) + \frac{2}{\sigma} \sum_{i=1}^{n} z_{(i)} = 0 \tag{3.7}$$

and

$$\frac{\partial \ln L^*}{\partial \sigma} = -\frac{n}{\sigma} - \frac{2}{\sigma} \sum_{i=1}^{n} z_{(i)} \left(\alpha_i - \beta_i z_{(i)}\right) + \frac{2}{\sigma} \sum_{i=1}^{n} z_{(i)}^2 = 0. \tag{3.8}$$

The solutions of these equations are the following MML estimators:

$$\hat{\mu}_{MML} = \bar{x}_w - \frac{\Delta}{m} \hat{\sigma}_{MML} \quad \text{and} \quad \hat{\sigma}_{MML} = \frac{-B + \sqrt{B^2 + 4nC}}{2\sqrt{n(n-1)}} \tag{3.9}$$

where

$$\bar{x}_w = \sum_{i=1}^{n} \delta_i x_{(i)}/m, \quad m = \sum_{i=1}^{n} \delta_i, \quad \delta_i = \beta_i + 1, \quad \beta_i = t_{(i)}^{-2}, \quad \Delta = \sum_{i=1}^{n} \alpha_i,$$

$$\alpha_i = 2t_{(i)}^{-1}, \quad B = 2\sum_{i=1}^{n} \alpha_i (x_{(i)} - \bar{x}_w) \quad \text{and} \quad C = 2\sum_{i=1}^{n} \delta_i (x_{(i)} - \bar{x}_w)^2.$$

Here, $x_{(i)}$ represents the i-th ordered observation. It should be noted that $t_{(i)} = E(z_{(i)})$ can be obtained approximately using the following equality:

$$t_{(i)} = F^{-1}\left(\frac{i}{n+1}\right), \quad i = 1, 2, \ldots, n$$

where $F^{-1}(\cdot)$ is the quantile function of the standard Maxwell distribution. The use of these approximate values does not affect the efficiency of the MML estimators adversely. It should also be noticed that the denominator of $\hat{\sigma}_{MML}$ is $2n$, however it is replaced by $2\sqrt{n\,(n-1)}$ for bias correction.

The MML estimators are derived in closed form since they are expressed as functions of the sample observations. Furthermore, they are asymptotically equivalent to the ML estimators. The MML estimators are also almost fully efficient, i.e. they have minimum variance bounds (MVBs). They also have very small bias or no bias even for small sample sizes. It should also be mentioned that the MML methodology gives small weight(s) to the outlying observation(s) in the direction of the longer tail(s). Therefore, the MML estimators are robust to the outlier(s), see e.g. Acitas et al. (2013) and references given therein for further information. See also Figure 2 where plots of the weights for the Maxwell distribution, i.e. $\delta_i = t_{(i)}^{-2} + 1$, are illustrated.

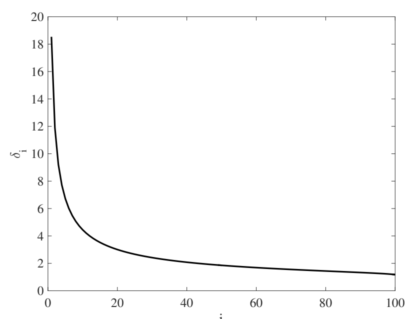

Figure 2: Plot of the weights for the Maxwell distribution; $n = 100$.

The asymptotic distributions of the $\hat{\mu}_{MML}$ and $\hat{\sigma}_{MML}$ are provided in Lemma 3.1 and Lemma 3.2.

Lemma 3.1. *$\hat{\mu}_{MML}$ is normally distributed with mean μ and variance σ^2/m for $n \to \infty$.*

Proof: The proof is done based on the following fact: The likelihood equation given in (3.2) and modified likelihood equation given in (3.7) are asymptotically equivalent. Furthermore, $\partial \ln L^*/\partial\mu$ can be written as

$$\begin{aligned}\frac{\partial \ln L^*}{\partial \mu} &= \frac{m}{\sigma^2}\left[\left(\bar{x}_w - \frac{\Delta}{m}\hat{\sigma}_{MML}\right) - \mu\right] \\ &= \frac{m}{\sigma^2}\left(\hat{\mu}_{MML} - \mu\right),\end{aligned} \tag{3.10}$$

see Kendall and Stuart (1961). $\hat{\mu}_{MML}$ is normally distributed since $\mathrm{E}(\partial^r \ln L^*/\partial\mu^r) = 0$ for all $r \geq 3$; see Bartlett (1953). □

Lemma 3.2. *Conditional on μ known, $n\hat{\sigma}^2_{MML}/\sigma^2$ is asymptotically chi-square distributed with n degrees of freedom.*

Proof: This follows from the fact that $B_0/\sqrt{nC_0} \cong 0$ and thus,

$$\frac{\partial \ln L^*}{\partial \sigma} = \frac{n}{\sigma^3}\left(\frac{C_0}{n} - \sigma^2\right) \tag{3.11}$$

where B_0 and C_0 are the same as B and C, respectively. See for example Tiku (1982) and Senoglu (2007) for further information. □

3.3. The MoM method

MoM estimators of the location and scale parameters of the Maxwell distribution are obtained by equating the first two theoretical moments to the first two sample moments. Therefore, MoM estimators of μ and σ are given by

$$\hat{\mu}_{MoM} = \bar{x} - \frac{2}{\sqrt{\pi}}\hat{\sigma}_{MoM} \quad \text{and} \quad \hat{\sigma}_{MoM} = s\sqrt{\frac{2\pi}{3\pi - 8}}, \tag{3.12}$$

respectively. Here,

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{and} \quad s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

It is clear that MoM estimators are functions of the sample observations as in MML estimators.

3.4. The LS method

LS estimators of μ and σ are obtained by minimizing the following function

$$\sum_{i=1}^{n}\left(F\left(x_{(i)}\right) - \frac{i}{n+1}\right)^2, \quad i = 1, 2, \ldots, n \tag{3.13}$$

with respect to the parameters of interest (Swain et al., 1988). Here, $F(\cdot)$ is the cdf of the Maxwell distribution. It is clear that explicit forms of the LS estimators are not available. Therefore, we use the "`fminunc`" function which exists in the optimization toolbox of MATLAB2017a to obtain the LS estimates of μ and σ.

4. SIMULATION STUDY

In this section, the results of the simulation study in which the performances of the MML estimators are compared with the ML, MoM and LS estimators are presented.

In the simulation setup, we use the sample sizes $n = 10$ (small), $n = 20$, $n = 50$ (moderate) and $n = 120$ (large). Without loss of generality, the location parameter μ and scale parameter σ are taken to be 0 and 1, respectively. All the simulations are carried out for $\lfloor 100,000/n \rfloor$ MC runs where $\lfloor \cdot \rfloor$ denotes the floor function (also known as the greatest integer function) that takes integer part of the number. We use the MATLAB2017a software for all computations. In the ML estimation procedure, the initial values for $\hat{\mu}$ and $\hat{\sigma}$ are taken as $\mu^0 = \hat{\mu}_{MML}$ and $\sigma^0 = \hat{\sigma}_{MML}$, respectively.

It should be noted that estimates of μ may sometimes be greater than the smallest order statistics $x_{(1)}$ due to the computational problems. These estimators are referred as impermissible estimators (Dubey, 1967). The problem is extinguished by reducing the impermissible estimators as $x_{(1)} - 10^{-4}$, see for example Kantar and Senoglu (2008).

The performances of the ML, MML, MoM and LS estimators are compared by using bias, variance, mean square error (MSE) and deficiency (Def) criteria. Def is a natural measure of the joint efficiency of the estimators $\hat{\mu}$ and $\hat{\sigma}$ and is defined by

$$Def(\hat{\mu}, \hat{\sigma}) = MSE(\hat{\mu}) + MSE(\hat{\sigma}), \tag{4.1}$$

see for example Akgul et al. (2016). The results of the simulation study are tabulated in Table 1. Following conclusions are drawn from Table 1.

Table 1: Simulated bias, variance, MSE and Def values of the ML, MML, MoM and LS estimators ($\mu = 0$ and $\sigma = 1$).

		$\hat{\mu}$			$\hat{\sigma}$			
	Estimators	Bias	Variance	MSE	Bias	Variance	MSE	Def
$n = 10$	ML	-0.108	0.050	0.062	0.094	0.042	0.051	0.113
	MML	-0.095	0.052	0.061	0.060	0.046	0.050	0.111
	MoM	-0.030	0.066	0.067	0.028	0.054	0.055	0.122
	LS	0.254	0.127	0.191	-0.338	0.131	0.245	0.436
$n = 20$	ML	-0.061	0.025	0.028	0.051	0.022	0.025	0.053
	MML	-0.061	0.026	0.029	0.038	0.023	0.025	0.054
	MoM	-0.018	0.033	0.034	0.016	0.028	0.028	0.061
	LS	0.193	0.058	0.096	-0.278	0.059	0.137	0.232
$n = 50$	ML	-0.030	0.009	0.010	0.022	0.009	0.009	0.019
	MML	-0.035	0.009	0.011	0.019	0.009	0.009	0.020
	MoM	-0.010	0.013	0.013	0.005	0.011	0.011	0.023
	LS	0.153	0.021	0.044	-0.241	0.020	0.078	0.122
$n = 120$	ML	-0.011	0.003	0.004	0.009	0.003	0.003	0.007
	MML	-0.015	0.004	0.004	0.009	0.003	0.003	0.007
	MoM	0.000	0.005	0.005	0.000	0.004	0.004	0.009
	LS	0.148	0.008	0.030	-0.232	0.008	0.062	0.092

Concerning the bias values, and for all sample sizes, the MoM estimator and LS estimator of μ have the smallest and the largest bias value, respectively. It can also be deduced from Table 1 that the bias values of the ML and MML estimators are very similar to each other as expected. The ML, MML and MoM estimators overestimate the location parameter μ while the LS estimator underestimates.

It is clear from Table 1 that the MoM estimator of σ has superiority over the ML, MML and LS estimators in terms of the bias criterion. For the small sample size, it is seen the MML estimator performs better than the ML estimator. However, the ML and MML estimators have more or less the same bias values for moderate and large sample sizes. The LS estimator of the σ has the biggest bias value among the all estimators.

Overall, all the estimators have negligible bias values except the LS estimators in what concerns the bias values of $\hat{\mu}$ and $\hat{\sigma}$.

Concerning the MSE values, the ML and MML estimator of μ have almost same the MSE values for all sample sizes. The LS estimator of the location parameter μ has the worst performance in terms of MSE among all other estimators.

Similar results are also obtained for the scale parameter σ. For example, the LS estimator does not perform well. The ML and MML estimators outperform the MoM estimator in most of the cases, however the MoM estimator has a considerably good performance. Table 1 also reveals that the ML and MML estimators are the most efficient.

To sum up, the ML and MML estimators are preferable among the other estimators according to the MSE criterion. The MSE values for $\hat{\mu}$ and $\hat{\sigma}$ decrease when the sample size n increases, as the theory says.

Concerning the Def values, the ML estimator has the smallest Def values among the other estimators for all cases. The Def values of the MML estimator are very close to those of the ML estimator except $n = 10$. The LS estimator shows the worst performance since it has the biggest Def values.

Finally, the ML and MML estimators are seen to be more efficient than the MoM and LS estimators. It is also clear that the performance of the ML and MML estimators are more or less the same as expected. As it is indicated previously, obtaining the ML estimates of the parameters requires iterative methods and this may cause some problems. On the other hand, the MML estimators are easily obtained from the sample observations without any iterative computations. As a result, the MML estimators may be preferable if our focus is to avoid the computational complexities besides having efficient estimators.

Robustness of the estimators

In this part of the simulation study, robustness properties of the ML, MML, MoM and LS estimators are investigated when there are plausible deviations from an assumed model. For this purpose, we assume that the underlying true model is Maxwell(μ=0, σ=1) and consider the following alternative models:

Outlier Model: $(n-r)Maxwell(0,1) + rMaxwell(0,2); r = \lfloor 0.5 + 0.1n \rfloor$.

Mixture Model: $0.80Maxwell(0,1) + 0.20Maxwell(0,2)$.

Contamination model: $0.90Maxwell(0,1) + 0.10Weibull(1, 0.8046)$.

Here, $Weibull(1, 0.8046)$ denotes the Weibull distribution with scale parameter $\sigma = 1$ and shape parameter $p = 0.8046$. Simulated mean, variance, MSE and Def values for the ML, MML, MoM and LS estimators of μ and σ under the alternative models are given in Table 2.

Table 2: Simulated mean, variance, MSE and Def values of the ML, MML, MoM and LS estimators under the alternative models.

		$\hat{\mu}$			$\hat{\sigma}$			
	Estimators	Mean	Variance	MSE	Mean	Variance	MSE	Def
				Model I: Outlier Model				
$n = 10$	ML	-0.032	0.090	0.091	1.142	0.108	0.128	0.220
	MML	-0.065	0.102	0.107	1.197	0.126	0.165	0.271
	MoM	-0.215	0.185	0.231	1.289	0.188	0.271	0.502
	LS	-0.215	0.178	0.326	1.535	0.198	0.484	0.810
$n = 20$	ML	-0.083	0.045	0.052	1.191	0.055	0.092	0.144
	MML	-0.097	0.049	0.059	1.219	0.061	0.109	0.168
	MoM	-0.242	0.099	0.158	1.315	0.099	0.198	0.356
	LS	-0.242	0.077	0.173	1.459	0.084	0.295	0.468
$n = 50$	ML	-0.113	0.019	0.032	1.219	0.023	0.071	0.103
	MML	-0.116	0.019	0.033	1.231	0.025	0.078	0.111
	MoM	-0.264	0.045	0.115	1.334	0.044	0.156	0.271
	LS	-0.264	0.029	0.094	1.402	0.030	0.191	0.285
				Model II: Mixture Model				
$n = 10$	ML	-0.101	0.114	0.124	1.307	0.179	0.274	0.398
	MML	-0.142	0.129	0.150	1.372	0.206	0.344	0.494
	MoM	-0.321	0.231	0.334	1.483	0.288	0.521	0.856
	LS	-0.321	0.354	0.705	1.841	0.483	1.191	1.896
$n = 20$	ML	-0.175	0.058	0.089	1.380	0.094	0.239	0.328
	MML	-0.194	0.062	0.100	1.415	0.102	0.274	0.374
	MoM	-0.383	0.126	0.272	1.541	0.154	0.446	0.719
	LS	-0.383	0.143	0.370	1.715	0.198	0.710	1.081
$n = 50$	ML	-0.208	0.023	0.066	1.408	0.037	0.204	0.270
	MML	-0.211	0.023	0.068	1.422	0.038	0.216	0.284
	MoM	-0.410	0.050	0.218	1.561	0.060	0.375	0.593
	LS	-0.410	0.048	0.212	1.632	0.066	0.465	0.677
				Model III: Contamination Model				
$n = 10$	ML	-0.096	0.167	0.177	1.095	0.217	0.226	0.402
	MML	-0.114	0.184	0.197	1.138	0.250	0.269	0.466
	MoM	-0.221	0.351	0.400	1.197	0.377	0.416	0.816
	LS	-0.221	0.194	0.363	1.478	0.199	0.428	0.791
$n = 20$	ML	-0.167	0.103	0.131	1.157	0.135	0.160	0.291
	MML	-0.167	0.109	0.137	1.177	0.151	0.182	0.319
	MoM	-0.266	0.255	0.326	1.236	0.261	0.317	0.643
	LS	-0.266	0.078	0.193	1.400	0.076	0.236	0.429
$n = 50$	ML	-0.215	0.049	0.096	1.207	0.069	0.111	0.207
	MML	-0.209	0.051	0.094	1.214	0.075	0.121	0.215
	MoM	-0.313	0.162	0.259	1.282	0.158	0.237	0.496
	LS	-0.313	0.029	0.113	1.354	0.028	0.154	0.266

It can be seen from the Table 2 that the ML and MML estimators outperform the MoM and LS estimators according to the MSE and Def criteria. This result implies that the ML and MML estimators of parameters μ and σ are more robust to the data anomalies given above.

5. PARAMETER ESTIMATION UNDER THE TYPE-II CENSORING

Analysis of censored samples are usually encountered in different fields of science such as agriculture, social sciences, medicine, and so on (Senoglu and Tiku, 2004). Therefore, we consider a Type-II censoring scheme. Type-II censoring arises if a predetermined number of lower and upper observations are censored (Senoglu and Tiku, 2004; Arslan and Senoglu, 2018).

According to the simulation results related with the robustness issue in Section 4, we concentrated on the ML and MML estimators of μ and σ under censoring. Let,

$$z_{(r_1)} \leq z_{(r_1+1)} \leq \ldots \leq z_{(n-r_2-1)} \leq z_{(n-r_2)}$$

be a Type-II censored samples where r_1 and r_2, with $r_1, r_2 \geq 0$ and $0 < r_1 + r_2 < n$, stand for the number of censored observations from the below and above, respectively. Then, the likelihood (L) function of the Maxwell distribution under the Type-II censored sample can be written as

$$L = \left[1 - F\left(z_{(r_1+1)}\right)\right]^{r_1} \prod_{i=r_1+1}^{n-r_2} f(z_{(i)}) \left[F\left(z_{(n-r_2)}\right)\right]^{r_2} \tag{5.1}$$

where $f(\cdot)$ and $F(\cdot)$ are the pdf and cdf of the Maxwell distribution given in Equations (2.2) and (2.3), respectively.

5.1. The ML method

The ML estimates of the parameters μ and σ under the Type-II censored samples are obtained by solving the following likelihood equations:

$$\frac{\partial \ln L}{\partial \mu} = -\frac{r_1}{\sigma} g_1\left(z_{r_1+1}\right) - \frac{2}{\sigma} \sum_{i=r_1+1}^{n-r_2} g_2\left(z_i\right) + \frac{2}{\sigma} \sum_{i=r_1+1}^{n-r_2} z_i + \frac{r_2}{\sigma} g_3\left(z_{n-r_2}\right) = 0 \tag{5.2}$$

and

$$\begin{aligned} \frac{\partial \ln L}{\partial \sigma} = &-\frac{n - r_1 - r_2}{\sigma} - \frac{r_1}{\sigma} z_{r_1+1} g_1\left(z_{r_1+1}\right) - \frac{2}{\sigma} \sum_{i=r_1+1}^{n-r_2} z_i g_2\left(z_i\right) + \frac{2}{\sigma} \sum_{i=r_1+1}^{n-r_2} z_i^2 \\ &+ \frac{r_2}{\sigma} z_{n-r_2} g_3\left(z_{n-r_2}\right) = 0 \end{aligned} \tag{5.3}$$

where $g_1\left(z_{r_1+1}\right) = \dfrac{f\left(z_{r_1+1}\right)}{F\left(z_{r_1+1}\right)}$, $g_2\left(z_i\right) = z_i^{-1}$ and $g_3\left(z_{n-r_2}\right) = \dfrac{f\left(z_{n-r_2}\right)}{1 - F\left(z_{n-r_2}\right)}$.

Similar to the complete sample case, the likelihood equations in (5.2) and (5.3) are nonlinear functions of the unknown parameters. Therefore, they cannot be obtained explicitly. The NR algorithm is also used here to solve the likelihood equations simultaneously.

5.2. The MML method

The MML estimators for the location μ and scale σ parameters of the Maxwell distribution are obtained under the Type-II censored samples by using an algorithm similar to the one given in subsection 3.2.

Nonlinear functions are linearized around the expected values of the standardized ordered observations , i.e. $t_{(i)} = E(z_{(i)})$, by using the first two terms of a Taylor series expansion:

$$\begin{aligned} g_1\left(z_{(r_1+1)}\right) &\cong \alpha_{1r_1+1} - \beta_{1r_1+1} z_{(r_1+1)}, \quad && g_2\left(z_{(i)}\right) \cong \alpha_{2i} - \beta_{2i} z_{(i)}, \\ g_3\left(z_{(n-r_2)}\right) &\cong \alpha_{3n-r_2} - \beta_{3n-r_2} z_{(n-r_2)}, \quad && \left(i = r_1+1, ..., n-r_2\right). \end{aligned} \tag{5.4}$$

After replacing nonlinear functions with their linearized versions in the likelihood equations, the following MML estimators are obtained:

$$\hat{\mu}_{MML} = \bar{x}_w - \frac{\Delta}{m}\hat{\sigma}_{MML} \quad \text{and} \quad \hat{\sigma}_{MML} = \frac{-B + \sqrt{B^2 + 4AC}}{2\sqrt{A(A-1)}} \tag{5.5}$$

where

$$m = r_1\beta_{1r_1+1} + 2\sum_{i=r_1+1}^{n-r_2}(\beta_{2i}+1) - r_2\beta_{3n-r_2}, \quad A = n - r_1 - r_2,$$

$$\bar{x}_w = \frac{r_1\beta_{1r_1+1}x_{(r_1+1)} + 2\sum_{i=r_1+1}^{n-r_2}(\beta_{2i}+1)\,x_{(i)} - r_2\beta_{3n-r_2}x_{n-r_2}}{m},$$

$$\Delta = r_1\alpha_{1r_1+1} + 2\sum_{i=r_1+1}^{n-r_2}(\alpha_{2i}) - r_2\alpha_{3n-r_2},$$

$$B = r_1\beta_{1r_1+1}(x_{(r_1+1)} - \bar{x}_w)^2 + 2\sum_{i=r_1+1}^{n-r_2}(\beta_{2i}+1)\,(x_{(i)} - \bar{x}_w)^2 - r_2\beta_{3n-r_2}(x_{n-r_2} - \bar{x}_w)^2,$$

$$C = r_1\alpha_{1r_1+1}(x_{(r_1+1)} - \bar{x}_w)^2 + 2\sum_{i=r_1+1}^{n-r_2}(\alpha_{2i}+1)\,(x_{(i)} - \bar{x}_w)^2 - r_2\alpha_{3n-r_2}(x_{n-r_2} - \bar{x}_w)^2,$$

$$\alpha_{1r_1+1} = g_1\left(t_{(r_1+1)}\right) + \beta_{1r_1+1}t_{(r_1+1)}, \quad \beta_{1r_1+1} = \frac{f'\left(t_{(r_1+1)}\right)}{F\left(t_{(r_1+1)}\right)} - \left[\frac{f\left(t_{(r_1+1)}\right)}{F\left(t_{(r_1+1)}\right)}\right]^2,$$

$$\alpha_{2i} = 2t_{(i)}^{-1}, \quad \beta_{2i} = t_{(i)}^{-2},$$

$$\alpha_{3n-r_2} = g_3\left(t_{(n-r_2)}\right) + \beta_{3n-r_2}t_{(n-r_2)}, \quad \beta_{3n-r_2} = \frac{f'\left(t_{(n-r_2)}\right)}{1 - F\left(t_{(n-r_2)}\right)} - \left[\frac{f\left(t_{(n-r_2)}\right)}{1 - F\left(t_{(n-r_2)}\right)}\right]^2.$$

It should be noticed that the denominator $2A$ is replaced by $2\sqrt{A\,(A-1)}$ in $\hat{\sigma}_{MML}$ as a bias correction.

We conducted a MC simulation study for this case and obtained similar results with those obtained in the complete sample case. Therefore, we would not give the results here for the sake of brevity. However, they can be provided upon request from the authors.

6. Applications

In this section, two real data sets are modelled by using the Maxwell distribution. The unknown parameters are estimated via the ML and MML methods since the MoM and LS methods fail to exhibit a good performance (see Section 4).

6.1. Example 1: Breaking Stress of Carbon Fibres Data

In this subsection, observations on the breaking stress of carbon fibres (in Gba) are used to show the implementation of the proposed methodology. The data set is given in Table 3. Further information about the data set can be found in Nicolas and Padgett (2006). See also Qian (2012) and Al-Sobhi and Soliman (2016), where the breaking stress of carbon fibres data are modelled using the exponentiated exponential (EE) and exponentiated Weibull (EW) distributions.

Table 3: Observations on breaking stress of carbon fibres, $n = 100$.

0.39	0.81	0.85	0.98	1.08	1.12	1.17	1.18	1.22	1.25	1.36	1.41	1.47	1.57
1.57	1.59	1.59	1.61	1.61	1.69	1.69	1.71	1.73	1.8	1.84	1.84	1.87	1.89
1.92	2.00	2.03	2.03	2.05	2.12	2.17	2.17	2.17	2.35	2.38	2.41	2.43	2.48
2.48	2.5	2.53	2.55	2.55	2.56	2.59	2.67	2.73	2.74	2.76	2.77	2.79	2.81
2.81	2.82	2.83	2.85	2.87	2.88	2.93	2.95	2.96	2.97	2.97	3.09	3.11	3.11
3.15	3.15	3.19	3.19	3.22	3.22	3.27	3.28	3.31	3.31	3.33	3.39	3.39	3.51
3.56	3.6	3.65	3.68	3.68	3.68	3.70	3.75	4.2	4.38	4.42	4.7	4.9	4.91
5.08	5.56												

In this study, Maxwell distribution is considered for modelling purposes. The modelling performance of the Maxwell distribution is compared with the performances of EE and EW distributions using well-known criteria such as Akaike Information Criterion (AIC) and corrected AIC (AICc). The smaller value of the AIC and AICc imply better fitting.

The parameter estimates along with $\ln L$, AIC and AICc values are given in Table 4. The results show that the Maxwell distribution performs a better modeling performance than its rivals in terms of considered criteria.

Table 4: Parameter estimates for breaking stress of carbon fibres data.

		$\hat{\mu}$	$\hat{\sigma}$	$\ln L$	AIC	$AICc$
Maxwell Distribution	ML	0.1402	2.1869	-141.6621	287.3242	287.4479
	MML	0.1816	2.1636	-141.7226	287.4452	287.5689
	$\hat{\alpha}_{ML}$	$\hat{\beta}_{ML}$	$\hat{\sigma}_{ML}$	lnL	AIC	$AICc$
Exponentiated Weibull	1.3169	2.4091	2.6824	-141.3320	288.6640	288.9140
Exponentiated Exponantial	7.7883	—	0.9870	-146.1823	296.3646	296.4883

It is also clear from the $\ln L$ values given in Table 4 that the ML estimates are preferable over the MML estimates. However, the ML estimates are obtained via the iterative method. On the other hand, the MML estimates are obtained easily since they are formulated explicitly. Furthermore, $\ln L$ values based on the ML and MML estimates do not differ so much. Therefore, the MML estimates can also be preferable for this data. It should be also noted that the Maxwell distribution provides better modelling performance than the EW distribution in spite of the fact that it has a lower number of parameters.

6.2. Example 2: Windmill data

The windmill data, in Table 5, was first considered by Joglekar et al. (1989). See also Kotb and Raqab (2017), where the modified Weibull distribution is used for modelling this data set.

Table 5: Observations on windmill data, $n = 25$.

0.123	0.5	0.558	0.653	1.057	1.137	1.144	1.194	1.501	1.562
1.582	1.737	1.800	1.822	1.866	1.930	2.088	2.112	2.166	2.179
2.236	2.294	2.303	2.310	2.386					

In this study, the Maxwell distribution is used to model the windmill data. Its modelling performance is also compared with the modelling performance of the modifed Weibull distribution. The results are given in Table 6.

Table 6: Parameter estimates for windmill data.

		$\hat{\mu}$	$\hat{\sigma}$	$\ln L$	AIC	$AICc$
Maxwell Distribution	ML	-0.1640	1.5393	-25.9676	55.9351	56.4806
	MML	-0.0905	1.5103	-26.0949	56.1898	56.7353
	$\hat{\alpha}_{ML}$	$\hat{\beta}_{ML}$	$\hat{\theta}_{ML}$	lnL	AIC	$AICc$
Modifed Weibull	0.2249	6.4644	0.0080	-25.7511	57.5022	58.6451

It can be concluded from Table 6 that the Maxwell distribution is preferable over the modified Weibull distribution according to the AIC and AICc criteria. The MML estimates can also be used as an alternative to the ML estimates here since the results are similar. Furthermore, the MML estimators have closed forms unlike the ML estimators.

7. Conclusion

In this study, estimation of the location and scale parameters of the Maxwell distribution is considered. Since the ML estimators cannot be obtained explicitly, the MML estimators having closed forms are derived. The MML estimators are asymptotically equivalent to the ML estimators. They are also fully efficient. We conducted a MC simulation study to compare the performance of the MML estimators with the ML, MoM and LS estimators. Simulation results show that the performance of the ML estimators is better than the other estimators. Furthermore, the MML and ML estimators have more or less the same performance. However, the ML estimators are obtained based on iterative methods. It is well known that using iterative methods causes some problems as mentioned in the text. On the other hand, the MML estimators are easily obtained from the sample observations without any iterative computations. It is concluded that the MML estimators may be preferable as an alternative to the ML estimators, if our focus is to avoid the computational complexities whilst high efficiency.

References

Acitas, S., Kasap, P., Senoglu, B., Arslan, O. (2013). One-step M-estimators: Jones and Faddy's skewed t-distribution, *Journal of Applied Statistics*, **40**,**7**, 1545–1560.

Akgul, F.G.., Senoglu, B., Arslan, T. (2016). An alternative distribution to Weibull for modeling the wind speed data: Inverse Weibull distribution, *Energy Conversion and Management*, **114**,**7**, 1234-240.

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.

Al-Sobhi, M.M., Soliman, A.A. (2016). Estimation for the exponentiated Weibull model with adaptive Type-II progressive censored schemes, *Applied Mathematical Modelling*, **40**, 1180–1192.

Arslan, T., Senoglu, B. (2018). Type II Censored Samples in Experimental Design Under Jones and Faddy's Skew t Distribution, *Iranian Journal of Science and Techonolgy Transaction A: Science*, **42**, 2145–2157.

Arslan, T., Acitas, S., Senoglu, B. (2017). *Estimating the location and scale parameters of the Maxwell distribution*, Data Science, Statistics and Visualisation Conference, Lisboa.

Barnett, V.D. (1966). Evaluation of the maximum likelihood estimator when the likelihood equation has multiple roots, *Biometrika*, **53**, 151–165.

Bartlett, M.S. (1953). Approximate confidence intervals, *Biometrika* , **40**, 12–19.

Dey, S., Maiti, S.S. (2010). Bayesian estimation of the parameter of Maxwell distribution under the different loss functions, *Journal of Statistical Theory and Practice* , **40**,**2**, 279–287.

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**, 291–300.

Dubey, S.D. (1967). On some permissible estimators of the location parameter of the Weibull and certain other distributions, *Technometrics* **9**,**2**, 293307.

Fan, G. (2016). Estimation of the loss and risk functions of parameter of Maxwell distribution, *Science Journal of Applied Mathematics and Statistics*, **4**,**4**, 129–133.

Hossain, A.M., Huerta, G. (2016). Bayesian estimation and prediction for the Maxwell failure distribution based on Type II censored data, *Open Journal of Statistics*, **6**, 49–60.

Joglekar, G., Schuenemeyer, J.H. and LaRiccia, V. (1989) Lack-of-fit testing when replicates are not available, *Amer. Statist.* **43** , 135–143.

Kantar, Y.M. and Senoglu, B. (2008). A comparative study for the location and scale parameters of the Weibull distribution with given shape parameter, *Computers & Geosciences*, **34**, 1900–1909.

Kazmi, S.M.A., Aslam, M. and Ali, S. (2011). A note on the maximum likelihood estimators for the mixture of Maxwell distributions using Type-I censored scheme, *The Open Statistics and Probability Journal*, **3** , 31–35.

Kotb, M.S., Raqab, M.Z. (2017). Inference and prediction for modified Weibull distribtion based on doubly censored samples, *Mathematics and Computers in Simulation*, **132**, 195–207.

Kendall, M.G., Stuart, A. (1961). *The Advanced Theory of Statistics*, Vol. 2, Charles Griffin and Co., London.

Li, L. (2016). Minimax estimation of the parameter of Maxwell distribution under the different loss functions, *American Journal of Theoretical and Applied Statistics*, **5**,**4**, 202–207.

Mathai,A.M.,Princy,T. (2017). Multivariate and matrix-variate analogues of Maxwell-Boltzmann and Rayleigh densities, *Pyhsica A*, **468**, 668–676.

Maxwell,J.C. (1860). Illustrations of the dynamical theory of gases. Part I. On the motions and collisions of perfectly elastic spheres, *Philosophical Magazine*, 4th series, 19, 19–32.

Nicolas, M.D. and Padgett, W.J. (2006). A bootstrap control chart for Weibull percentiles, *Qual. Rehab. Eng. Int.*, **22**, 141–151.

Puthenpura, S., Sinha, N.K. (1986). Modified maximum likkelihood method for the robust estimation of system parameters from very noise data, *Automatica*, **22**, 231–235.

Senoglu, B. (2007). Estimating parameters in one-way analysis of covarinace model with short-tailed symmetric error distributions, *Journal of Computational and Applied Mathematics*, **201**, 275–283.

Senoglu, B. and Tiku, M.L. (2004). Censored and truncated samples in experimental design under non-normality, *Statistical Methods* **6**, **2**, 173–199.

Swain, J., Venkatraman, S. and Wilson, J. (1988). Least squares estimation of distribution function in Jonhson's translation system, *Journal of Statistical Computation and Simulation*, **29**, 271–297.

Tiku, M. L. (1967.) Estimating the mean and standard deviation from a censored normal sample, *Biometrika*, **54**, 155–165.

Tiku, M. L. (1968). Estimating the parameters of Normal and Logistic distributions form censored samples, *Aust. J. Stat.*, **10**, 64–74.

Tiku, M.L. (1982). Testing linear contrast of means in experimental design without assuming normality and homogeneity of variances, *Biometrical Journal*, **24**, **6**, 613–627.

Tyagi, R.K., Battacharya, S.K. (1989a). Bayes Estimation of the Maxwell's Velocity Distribution Function, *Statistica*, **4**, 563–567.

Tyagi, R.K., Battacharya, S.K. (1989b). A Note on the MVU Estimation of the Maxwell's Failure Distribution, *Estadistica*, **41**, 73–79.

Qian L. (2012). The Fisher information matrix for three-parameter exponentiated Weibull distribution under type II censoring, *Statistical Methodology*, **9**, 320–329.

Vaughan, D.C. (1992). On the Tiku-Suresh method of estimation, *Commun. Stat.-Theory Meth.*, **21**,**2**, 451–469.

Appendix

1 *Elements of the Hessian matrix*

$$
\begin{aligned}
\frac{\partial^2 \ln L}{\partial \mu^2} &= -\frac{2n}{\sigma^2} - \frac{2}{\sigma^2} \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma} \right)^{-2}, \\
\frac{\partial^2 \ln L}{\partial \mu \partial \sigma} &= -\frac{4}{\sigma^2} \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma} \right), \\
\frac{\partial^2 \ln L}{\partial \sigma^2} &= \frac{3n}{\sigma^2} - \frac{6}{\sigma^2} \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma} \right)^{2}.
\end{aligned}
$$

2 *Fisher Information ($\mathbf{I}$) matrix of the Maxwell distribution*

$$
\begin{aligned}
\mathbf{I} &= \begin{bmatrix} -\mathbf{E}\left(\dfrac{\partial^2 \ln L}{\partial \mu^2} \right) & -\mathbf{E}\left(\dfrac{\partial^2 \ln L}{\partial \mu \partial \sigma} \right) \\ -\mathbf{E}\left(\dfrac{\partial^2 \ln L}{\partial \sigma \partial \mu} \right) & -\mathbf{E}\left(\dfrac{\partial^2 \ln L}{\partial \sigma^2} \right) \end{bmatrix} \\
&= \frac{n}{\sigma^2} \begin{bmatrix} 6 & \dfrac{8}{\sqrt{\pi}} \\ \dfrac{8}{\sqrt{\pi}} & 6 \end{bmatrix}.
\end{aligned}
$$