Parameters estimation methods of the Weibull distribution: A comparative study

Osabuohien-Irabor Osarumwense\(^1\) and Nwosu Chinwe Rose\(^2\)

\(^1\)Department of Mathematics/Statistics, Ambrose Alli University, Nigeria.
\(^2\)Department of Applied Statistics, Nnamdi Azikiwe University, Nigeria.

ABSTRACT

In this study, a comparison of performance among four different methods (Probability Weighted Moments (PWM), Maximum Likelihood Methods (MLM), the Methods of Moments (MOM), and the least Square Method (LSM)) of estimating the parameters of the Weibull distribution was done, using different values of \(\alpha\) and \(\beta\). This study contains two different experiments; experiment I and experiment II. Values of \(\alpha > \beta\), \(\alpha < \beta\), \(\alpha = \beta\) were used in experiment I, and another varied values of \(\alpha < \beta\), \(\alpha < \beta\), \(\alpha = \beta\) in line with some modeled distribution of the Weibull distribution, were used in experiment II. These values were generated by simulation. The aim is to find the best methods of estimating the two-parameter Weibull distribution. Based on the sample sizes and parameters considered, the method which gave the best estimate for the two-parameter Weibull distribution, is the method of Moments (MOM), taking into consideration the total deviation as a measurement for comparison. This study will help to estimate parameters associated with various models of a component or system such as reliability problems, maintainability and predictions such as MTTF, MTTFF, MTTR, MTBF, Hazard Rate etc.

© 2014 Elixir All rights reserved.
\[ f(x) = \alpha \beta t^{\beta-1} e^{-\alpha t^\beta} \quad t > 0, \alpha > 0, \beta > 0 \] (3)

Where \( t \) is the time to failure (life length of a component), and \( \alpha \) and \( \beta \) are the scale and shape parameter respectively. [6], also put the two-parameter Weibull distribution which has been used in this study as;

\[ f(x) = \left( \frac{\beta}{\alpha} \right) t^{\beta-1} e^{-\frac{t^\beta}{\alpha}} \quad t > 0, \alpha > 0, \beta > 0 \] (4)

The cummulation distribution function (CDF) of the two-parameter Weibull distribution also known in reliability study as Unit unreliability written as \( F(t) \), given by [7] as,

\[ F(t; \alpha, \beta) = 1 - e^{-\frac{t^\beta}{\alpha}} \quad t > 0, \alpha > 0, \beta > 0 \] (5)

Where \( F(t; \alpha, \beta) \) is the unreliability unit and \( e^{-\frac{t^\beta}{\alpha}} \) is the reliability unit or survival function of the Weibull distribution equating to \( I \).

The instantaneous failure rate at any point in line of the Weibull distribution, which is also known as the hazard function \( h(t) \) is given by [8],

\[ h(t) = \left( \frac{\beta}{\alpha} \right) t^{\beta-1} \quad t > 0, \alpha > 0, \beta > 0 \] (6)

and the mean and variance of the two-parameter Weibull distribution from (4) are given as respectively as,

\[ E(T) = \alpha \beta \Gamma \left( 1 + \frac{1}{\beta} \right) \] (7)

\[ Var(T) = \left\{ \alpha^2 \beta^2 \Gamma \left( 1 + \frac{2}{\beta} \right) - \Gamma^2 \left( 1 + \frac{1}{\beta} \right) \right\} \] (8)

The shape parameter \( (\beta) \), gives the Weibull distribution its flexibility, and as shown by [9], if \( \beta = 1 \), the Weibull distribution is identical to the exponential distribution, if \( \beta = 2 \), it is identical to the Rayleigh distribution and if \( \beta = 3.5 \) and 2.5, the Weibull approximates the normal and lognormal distribution respectively. These and other unknown parameters of the Weibull distribution have to be estimated, hence, the need for this paper which aims at finding the best method of estimating these parameters.

Many papers have been published in the 1960s and 1970s concerning parameter estimation of the Weibull distribution based on complete and censored sample. [10], described the Maximum likelihood estimation method and derived the variance-Covariance matrix of the parameter of a two-parameter Weibull distribution based on complete and censored sample.

[11], developed an iterative procedure to find the Maximum likelihood estimate for a three-parameter Weibull distribution and illustrated numerical examples for; one-, two-, and three-parameter Weibull model. [12], derive the elements of the information matrix of the maximum likelihood estimate for a three-parameter Weibull distribution.

[13], derived the confidence limits of a two-parameter Weibull distribution.

[14], applied a two-parameter Weibull model for regional flood quantile estimation based on the index flood assumption and compared the method of moments (MOM), probability Weighted Moments (PWM) and Maximum likelihood (ML) by simulation experiment.

[15], derived the asymptotic variance of the quantile obtained from regional analysis for three estimation techniques for a two-parameter Weibull Model. They also applied the regional Weibull Model to fit the frequency distribution of annual flood data.

[9], gave the confidence interval for the shape, scale and location parameters based on the method of Maximum likelihood and linear estimation.

[16], proposed a method that depends on two-step iteration procedure. In the first step, he assumed that the location parameter is known, and the scale and shape parameters are estimated using the graphical methods. In the second step, the shape parameter is assumed known and the scale and location parameter are estimated by transforming the data using law transformation so that the transformation data can be modeled to a two parameter exponential distribution.
[17], presented methods for estimating the parameters of the two-parameter Weibull distribution. He put his computational experiments in two categories; Graphical methods (Weibull probability plotting and Hazard plotting technique) and analytical methods (Maximum likelihood Estimator, Methods of Moments, and the Least squares Methods). Based on the computational results, he concluded that the best estimating methods for two-parameter Weibull distribution, is the Method of Moments.

[18], estimated the parameter of the Weibull distribution graphically using the Weibull probability plot to transform the Weibull transformation into x

[5], compared the three well-known methods (MLM, MOM, LSM) of parameter estimation of the Weibull distribution. A scale ($\alpha$) and shape ($\beta$) parameter of 1.0 and 3.0 were used respectively for the two-parameter Weibull distribution and values of 1, 2, and 3 were used for the three parameter Weibull distribution respectively.

[19], estimated the parameters of the modified Weibull distribution.

**Material and methods**

In this study, the data used were generated by simulation with the help of statistical software known as Minitab. Sets of parameters used for this study were put in two different experiments. In experiment I, a case where the $\alpha > \beta(3.0, 0.8)$, $\alpha < \beta(0.5, 4.0)$ and $\alpha = \beta(= 2)$. Experiment II, which is in line with the values of some modeled distributions of the Weibull distribution as shown in [9]. ($\alpha = 1, \beta = 1$), ($\alpha = 1, \beta = 2$), and ($\alpha = 1, \beta = 3.5$). These six sets of parameters were used in seven different sample sizes of 10, 20, 30, 50, 60, 70 and 100 giving a total of fifteen different sample data. With the help of Microsoft Excel 2005, the probability density functions of these four different estimating methods were compared. And the data generated plus the assumed parameters were used in evaluating these estimators.

Greenwood et al, [20], explained that a probability distribution function, $F(t) = \text{Prob.}(T \leq t)$ can be characterized by the probability weighted moments (PWM) given as;

$$M_{i,j,k} \equiv E(t^i F^j (1 - F)^k) = \int_0^1 (t (F)^i F^j (1 - F)^k) dF(t) \quad (9)$$

Where $i, j$ and $k$ are real numbers, but if $j$ and $k$ are non-negative integers, then;

$$M_{i,0,k} = \sum_{j=0}^{k} \binom{k}{j} (-1)^j M_{i,j,0} \quad (10)$$

$$M_{i,j,0} = \sum_{k=0}^{j} \binom{j}{k} (-1)^k M_{i,0,k} \quad (11)$$

Where, if $M_{i,0,k}$ exists and $t$ is a continuous function of $F$, $M_{i,j,0}$ exists. And if, $t = t(F)$ is in inverse form, then,

$$t(F) = \alpha \ln \left( \frac{1}{1 - F(t; \alpha, \beta)} \right)^{\frac{1}{\beta}} \quad (12)$$

If $M_k = M_{1,0,k}$, then the probability weighted moments (PWM) of two-parameters Weibull distribution, without loss of generality is given as,

$$M_k = \frac{\alpha \Gamma \left( 1 + \frac{1}{\beta} \right)}{\left( 1 + k \right)^{\left( 1 + \frac{1}{\beta} \right)}} \quad (13)$$

Therefore,

$$\hat{\alpha} = \frac{M_0}{\Gamma \left( \frac{\ln(M_0/M_1)}{\ln(2)} \right)} \quad (14)$$

$$\hat{\beta} = \frac{\ln(2)}{\ln(M_0/2M_1)} \quad (15)$$

Where;
\[\hat{M}(0) = \frac{1}{n} \sum_{i=1}^{n} X_i \Rightarrow \sum_{i=1}^{n} X_i \]  \hspace{1cm} (16)

\[\hat{M}(1) = \frac{1}{n(n-1)} \sum_{i=1}^{n-1} X_i (n-i) \Rightarrow \sum_{i=1}^{n} X_i (F_i) \]  \hspace{1cm} (17)

\[\hat{M}(2) = \frac{1}{n(n-1)(n-2)} \sum_{i=1}^{n-1} X_i (n-i)(n-i-1) \Rightarrow \sum_{i=1}^{n} X_i (F_i)^2 \]  \hspace{1cm} (18)

In the maximum Likelihood method, if \(t_1, t_2 \ldots \ldots t_n\) is a random sample of \(n\) independent observation whose probability density is \(F(t; \alpha, \beta)\), then the likelihood can be written as,

\[\mathcal{L}(t_1, t_2 \ldots \ldots t_n; \alpha, \beta) = \prod_{i=1}^{n} f(t; \alpha, \beta) \]  \hspace{1cm} (19)

Then, the estimators are;

\[\hat{\alpha} = \frac{\sum_{i=1}^{n} t^\beta \ln t_i}{\sum_{i=1}^{n} t^\beta \ln t_i} \]  \hspace{1cm} (20)

\[\alpha = \frac{1}{n} \sum_{i=1}^{n} t^\beta \]  \hspace{1cm} (21)

Another analytical method of estimation used was the method of moments with estimators as,

If the \(k^{th}\) moment of the sample is:

\[W_k = E(t^k) = \alpha \frac{k}{\beta} \Gamma(1 + \frac{k}{\beta}) \]  \hspace{1cm} (22)

\[\Gamma(n) = \int_0^\infty t^{n-1} e^{-t} dt \]  \hspace{1cm} (23)

Therefore, the estimators are;

\[\hat{\alpha} = \left( \frac{\frac{1}{n} \sum_{i=1}^{n} t^\beta}{\Gamma(1 + \frac{k}{\beta})} \right)^{\frac{1}{\beta}} \]  \hspace{1cm} (24)

\[\alpha = \left( \frac{\frac{1}{n} \sum_{i=1}^{n} t^\beta}{\Gamma(1 + \frac{k}{\beta})} \right)^{\frac{1}{\beta}} \]  \hspace{1cm} (25)

For the least square estimator, double natural logarithms were “taken” to put the function (relationship between the cdf and the parameters) in linear order.

\[F(i) = 1 - e^{-\frac{i \beta}{\alpha}} \]  \hspace{1cm} (26)

Where \(F(i) = \frac{i}{n+1}\) is the mean rank failure [17], \(i\) is the number of failure, \(n = \text{number of data points}\), and \(\ln = \text{Natural logarithms}\).

Therefore,
Each of the different estimators for the four different methods was resolved simultaneously using the error predictor numerical iteration as described by [7]. The iteration modifies the initial value of the shape parameter (β) and starting with a modification of ±0.1, the values of α = ̂α decreases in steps of 0.1, 0.01, 0.001…………etc as its converges. The iteration in this study used an accuracy of 0.01l with a termination criterion of ξ where ξ is the desired level of accuracy and α = ̂α is the difference between the two estimators. And this set β = ̂β and ̂α = α + ̂α 

And the stirling approximation formula for gamma function was used in expressing values associated with gamma, and given by [21], as,

\[ \Gamma(\alpha + 1) \approx \sqrt{2\pi\alpha} \left( \frac{\alpha}{e} \right)^\alpha \]  

(29)

\[ \alpha! \approx \alpha^\alpha e^{-\alpha} \sqrt{2\pi\alpha} \left( 1 + \frac{1}{12\alpha} \right) \]  

(30)

(30) is the modification of (29) which was used in this study.

In statistical analysis, the total deviation (TD) and the mean square error (MSE) are two measurements that can give an indication of accuracy of parameter estimation [5]. The total deviation for the two-parameter Weibull distribution can be calculated as given by [17],

\[ TD = \left| \frac{\alpha - \bar{\alpha}}{\alpha} \right| + \left| \frac{\beta - \bar{\beta}}{\beta} \right| \]  

(31)

The total deviation (TD) was used in this study as a measurement to indicate the degree of accuracy. The smaller the variance (Total deviation), the better the method.

**Results and discussion**

The different estimators obtained from the four different methods of estimating the two parameter Weibull distribution were evaluated and the different estimates at different sample sizes calculated.

In experiment I (see tables 1, 2, and 3), the values of α and β were varied, while in experiment II, as shown in tables 4, 5, and 6, the values of α remain unchanged while β varied in line with some modeled distributions from the Weibull distribution. The computational analysis of experiment I for the four compared estimation methods (Maximum likelihood Method, Method of Moments and the least Square Method) for the two-parameter Weibull distribution revealed that the method of moments (MOM) is the best estimating method with a total deviation (TD) of 0.190 and 0.432 when a sample size of 20 and 100 were used for values of α = 3.0 and β = 0.8 respectively as shown in Table I.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>β</td>
<td>n</td>
<td>̂α</td>
<td>̂β</td>
</tr>
<tr>
<td>3.0</td>
<td>0.8</td>
<td>20</td>
<td>3.621</td>
<td>0.889</td>
</tr>
<tr>
<td>60</td>
<td>3.809</td>
<td>0.943</td>
<td>0.447</td>
<td>1.706</td>
</tr>
<tr>
<td>100</td>
<td>4.000</td>
<td>0.716</td>
<td>0.438</td>
<td>1.857</td>
</tr>
</tbody>
</table>

Source: Researcher’s Calculation
Table 2: Comparison of MLM, MOM and LSM when $\alpha = 0.5$ and $\beta = 4.0$

<table>
<thead>
<tr>
<th>Parameters/ Sample sizes</th>
<th>Sample sizes</th>
<th>Prob. Weighted Moments</th>
<th>Max. Likelihood Method (MLM)</th>
<th>Moments Method (MOM)</th>
<th>Least Square Method (LSM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$, $\beta$, n</td>
<td></td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
</tr>
<tr>
<td>0.5, 4.0, 30</td>
<td>0.777</td>
<td>6.174, 1.097</td>
<td>0.081, 4.000, 0.383</td>
<td>0.082, 4.000, 0.386</td>
<td>1.141, 0.308, 2.205</td>
</tr>
<tr>
<td></td>
<td>0.816</td>
<td>5.410, 0.984</td>
<td>0.066, 4.000, 0.368</td>
<td>0.066, 4.000, 0.367</td>
<td>1.761, 1.004, 3.271</td>
</tr>
<tr>
<td>20, 2.0, 70</td>
<td>1.161</td>
<td>2.011, 0.166</td>
<td>0.975, 1.090, 0.115</td>
<td>0.969, 1.070, 0.101</td>
<td>1.540, 1.011, 0.551</td>
</tr>
<tr>
<td></td>
<td>1.097</td>
<td>0.875, 0.222</td>
<td>0.857, 0.960, 0.183</td>
<td>0.884, 1.010, 0.126</td>
<td>5.231, 1.004, 4.235</td>
</tr>
</tbody>
</table>

Source: Researcher’s Calculation

However, the Probability Weighted moments (PWM) with a deviation of 0.354 was best when a sample size of 10 was used. Also, for case where $\beta > \alpha$, where sample sizes of 30 and 50 were used as shown in Table 2, the method of Moments (MOM) had the least deviation of 0.836 and 0.867 respectively and therefore consider as the best estimating method for this sample sizes. However, the MLM came close with a total deviation of 0.838 and 0.868 respectively. Results from Table 1-3, show that the least Squares method is consistently the worst estimating method for the sample sizes and parameter values. When a sample size of 70 was applied to a case of $\alpha=\beta=2$, the method of moments had the least total deviation (TD) of 1.696.

Table 4: Comparison of MLM, MOM and LSM when $\alpha = 1.0$ and $\beta = 1.0$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$, $\beta$, n</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
</tr>
<tr>
<td>1.0, 1.0, 10</td>
<td>2.067, 0.886, 1.181</td>
<td>1.149, 2.070, 2.129</td>
<td>1.127, 2.120, 1.127</td>
<td>3.010, 1.653, 2.663</td>
</tr>
<tr>
<td></td>
<td>1.121, 0.977, 0.144</td>
<td>0.975, 1.090, 0.115</td>
<td>0.969, 1.070, 0.101</td>
<td>1.540, 1.011, 0.551</td>
</tr>
<tr>
<td></td>
<td>1.097, 0.875, 0.222</td>
<td>0.857, 0.960, 0.183</td>
<td>0.884, 1.010, 0.126</td>
<td>5.231, 1.004, 4.235</td>
</tr>
</tbody>
</table>

Source: Researcher’s Calculation

Table 5: Comparison of MLM, MOM and LSM when $\alpha = 1.0$ and $\beta = 2.0$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$, $\beta$, n</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
</tr>
<tr>
<td>1.0, 2.0, 10</td>
<td>1.246, 2.216, 0.354</td>
<td>1.429, 2.330, 0.594</td>
<td>1.415, 2.300, 0.565</td>
<td>2.940, 1.407, 2.236</td>
</tr>
<tr>
<td></td>
<td>1.161, 2.011, 0.166</td>
<td>1.126, 1.910, 0.171</td>
<td>1.151, 1.990, 0.156</td>
<td>1.840, 1.017, 1.331</td>
</tr>
<tr>
<td></td>
<td>1.096, 2.321, 0.256</td>
<td>0.859, 1.920, 0.181</td>
<td>0.861, 1.980, 0.149</td>
<td>3.452, 1.005, 2.959</td>
</tr>
</tbody>
</table>

Source: Researcher’s Calculation

Table 6: Comparison of MLM, MOM and LSM when $\alpha = 1.0$ and $\beta = 3.5$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$, $\beta$, n</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
<td>$\bar{\alpha}$, $\beta$, TD</td>
</tr>
<tr>
<td>1.0, 3.5, 10</td>
<td>0.098, 3.713, 0.152</td>
<td>1.246, 3.800, 0.211</td>
<td>1.077, 3.400, 0.105</td>
<td>1.765, 2.400, 1.079</td>
</tr>
<tr>
<td></td>
<td>1.239, 3.255, 0.409</td>
<td>0.954, 5.100, 0.503</td>
<td>0.973, 3.510, 0.022</td>
<td>2.643, 1.018, 2.352</td>
</tr>
<tr>
<td></td>
<td>1.144, 3.866, 0.248</td>
<td>0.913, 3.350, 0.129</td>
<td>0.906, 3.200, 0.179</td>
<td>3.119, 1.005, 2.831</td>
</tr>
</tbody>
</table>

Source: Researcher’s Calculation

Table 4, 5 and 6 shows the computational analysis of experiment II, as described by [9]. In Table 4, where values of $\alpha=\beta=1$, the method of Moments had the least minimum variance of 1.127, 0.101 and 0.126 for sample sizes of 10, 20, and 60 respectively. The MLM was close with a TD of 1.219 and 0.183 when a sample size of 10 and 60 were used. The least squares method has again shown to be poor estimator for the two-parameter Weibull distribution. For $\alpha=1$, $\beta=1$ which modeled the Rayleigh distribution, the method of moments (MOM) was also the best estimating method for sample sizes of 20 and 60 with a total deviation (TD) of 0.156 and 0.149 respectively as shown in Table 5. But for $\alpha=1$, $\beta=3.5$ as shown in Table 6, with least deviation of 0.105, and 0.022 the Method of Moments (MOM) came best when a sample sizes of 10 and 20 were used respectively. However, the MLM was the best method for estimating the parameter of the Weibull distribution when a sample size of 60 was used.
Conclusion

Based on the computational results from the sets of parameters and sample sizes considered in this study, and as used in the two different experiments as shown in Tables 1, 2, 3, 4, 5 and 6, the following results shows that:

(i) The best method for estimating the two-parameter Weibull distribution, taking into consideration the total deviation as a measurement for comparison, is the method of Moments (MOM).

(ii) The Method of Moments (MOM) is better than the Greenwood’s Probability Weighted Moments (PWM) as a Method of estimating the two-parameter Weibull distribution.

(iii) Results also show that the Maximum likelihood method (MLM) and the probability Weighted Moments (PWM) can be a close substitute when estimating the two-parameter Weibull distribution.

(iv) While the least squares method (LSM) is a poor estimating method for the two-parameter Weibull distribution as revealed in the computational analysis.

References


**Biographies**

*Osabuohien-Irabor Osarumwense*, holds a Bachelor degree in Statistics and Masters degree in Applied Statistics from Ambrose Alli University, Nigeria and Nnamdi Azikiwe University, Nigeria, respectively. Currently, he is pursuing a doctorate degree in Statistics with specialty in Time Series Analysis. He is a Lecturer at Ambrose Alli University, Ekpoma, Nigeria.

*Dr. Nwosu Chinwe Rose*, holds a B.Sc Statistics, M.Sc Statistics from the prestigious University of Nigeria, Nsukka, and a PhD Statistics from Abia State University, Nigeria. She is currently the Head of Department and a Senior Lecturer in department of Statistics, Nnamdi Azikiwe University, Nigeria. She has held high administrative positions in and outside the University system including the position of the Registrar, Federal Polytechnic, Oko, Nigeria.