Evaluation of Some Methods for Estimating the Weibull Distribution Parameters

J. A. Labban[1,*], H. H. Depheal[2]

[1]*Department of Statistics and Informatics, University of Al-Qadisiyah, Al-Qadisiyah, Iraq*

[2]*General Directorate of Al-Qadisiyah, Al-Qadisiyah, Iraq*

[*]*Corresponding author:* jubran.Labban@qu.edu.iq

Received: 30-7-2017, Accepted: 5-10-2017, Published: 19-10-2017

DOI: 10.18081/2226-3284/017-10/6-12

Abstract: This paper some of different methods to estimate the parameters of the 2-Paramaters Weibull distribution such as (Maximum likelihood Estimation, Moments, Least Squares, Term Omission). Mean square error will be considered to compare methods fits in case to select the best one. There by simulation will be implemented to generate different random sample of the 2-parameters Weibull distribution, those contain (n=10, 50, 100, 200) iteration each 1000 times.

***Keywords*:** Maximum likelihood estimation, Method of moments, least squares method, Mean-standard deviation method, Power density method, Term omission method, Mean squared error.

1. Introduction

In statistics, there are many applications for the Weibull distribution. The Weibull Distribution has been widely studied since its introduction in 1951 by Professor Wallodi Weibull[5]. Although it was first identified by Frechet in 1927, it is named after Waalobi Weibull and is a cousin to both the Frechet and Gumbel distributions. In section 3, we discuss the four methods; the maximum likelihood estimation (MLE), the moments estimation (MOE), the least squares method (LSM) or Rank Regression Method (RRM), Mean-Standard Deviation Method (MSD), Power Density Method (PDM), and Term Omission Method (TOM). Moreover, in section 4, these methods are compared, using the mean square error (MSE).

2-The Weibull Distribution:

The form of a two-parameter Weibull probability density function (pdf) with shape β and scale α parameters is given by

$$f(x\,|\,\alpha,\beta)=\frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1}\exp\left\{-\left(\frac{x}{\alpha}\right)^{\beta}\right\}$$

$$x\ge 0;\ \alpha,\beta>0 \qquad (1)$$

and the cumulative distribution function (cdf) of our distribution is

$$F(x\,|\,\alpha,\beta)=1-\exp\left\{-\left(\frac{x}{\alpha}\right)^{\beta}\right\},x\ge 0; \qquad (2)$$

with mean and variance $E(x)=\alpha\Gamma(1+\frac{1}{\beta})$,

$V(x)=\alpha^2\left[\Gamma(1+\frac{2}{\beta})-\Gamma^2(1+\frac{1}{\beta})\right]$ respectively,

where Γ is a gamma function.

2. Estimation Procedures

2.1. Maximum Likelihood Estimator (MLE)

Treating our parameters as unknown values is the maximum likelihood estimator (MLE) which is defined by and finding the joint density of all observations of a data set, which are assumed to be independent and identically distributed (iid). This estimator is important in statistics because of is asymptotic unbiasedness and minimal variance [9].

Let x_1, x_2, …, x_n be a random sample of size n drawn from a population with probability density function $f(x\,|\,\underline{\theta})$ where

$\underline{\theta} = (\alpha, \beta)$ is a vector of parameters of Weibull distribution, so that the joint density of the

$$L(\underline{\theta}, \underline{x}) = \prod_{i=1}^{n} f(x_i; \underline{\theta}) \tag{3}$$

$\underline{\theta} = (\alpha, \beta)$ maximizes L, or equivalently
The logarithm of L when

$$\frac{\partial \ln L}{\partial \underline{\theta}} = 0 \tag{4}$$

So, the likelihood function of Weibull pdf is given as:

$$L(x_1, x_2, ..., x_n; \alpha, \beta) = \prod_{i=1}^{n} \left(\frac{\beta}{\alpha}\right)\left(\frac{x_i}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{x_i}{\alpha}\right)^{\beta}\right]$$

$$= \left(\frac{\beta}{\alpha}\right)^{n} (\alpha)^{n-n\beta} \left(\prod_{i=1}^{n} x_i^{\beta-1}\right) \exp\left[-\sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta}\right]$$

$$= \left(\frac{\beta}{\alpha^{\beta}}\right)^{n} \left(\prod_{i=1}^{n} x_i^{\beta-1}\right) \exp\left[-\sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta}\right] \tag{5}$$

Taking the natural logarithm yields

$$\ln L = \left(\frac{\beta}{\alpha^{\beta}}\right)^{n} \left(\prod_{i=1}^{n} x_i^{\beta-1}\right) \exp\left[-\sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta}\right]$$

$$\ln L = n \ln \beta - n\beta \ln \alpha + (\beta - 1)\sum_{i=1}^{n} \ln x_i - \alpha^{-\beta} \sum_{i=1}^{n} (x_i)^{\beta} \tag{6}$$

With differentiating (6) partially with respect to β and α and equating to zero, we obtain the estimating equations as follows

$$\frac{\partial}{\partial \beta} \ln L = \frac{n}{\beta} + \sum_{i=1}^{n} \ln x_i - \frac{1}{\alpha} \sum_{i=1}^{n} x_i^{\beta} \ln x_i = 0 \tag{7}$$

and

$$\frac{\partial}{\partial \alpha} \ln L = -\frac{n}{\alpha} + \frac{1}{\alpha^2} \sum_{i=1}^{n} x_i^{\beta} = 0 \tag{8}$$

From eq. (8) we have the estimator of α as

$$\hat{\alpha}_{mle} = \left[\frac{1}{n} \sum_{i=1}^{n} x_i^{\hat{\beta}}\right]^{1/\beta} \tag{9}$$

and on substitution eq. (9) in (7) we obtain

$$\frac{1}{\beta} + \frac{1}{n} \sum_{i=1}^{n} \ln x_i - \frac{\sum_{i=1}^{n} x_i^{\beta} \ln x_i}{\sum_{i=1}^{n} x_i^{\beta}} = 0 \tag{10}$$

which may be solved to get the estimate of β using Newton-Raphson method. Which can be written in the form

$$x_{n+1} = x_n + \frac{f(x_n)}{f'(x_n)} \tag{11}$$

likelihood function is defined by the product of the densities of each data point [4].

The maximum likelihood of

where

$$f(\beta) = \frac{\sum_{i=1}^{n} x_i^{\beta} \ln x_i}{\sum_{i=1}^{n} x_i^{\beta}} - \frac{1}{\beta} - \frac{1}{n} \sum_{i=1}^{n} \ln x_i \tag{12}$$

And

$$f'(\beta) = \frac{\sum_{i=1}^{n} x_i^{\beta} \sum_{i=1}^{n} x_i^{\beta} (\ln x_i)^2 - \left(\sum_{i=1}^{n} x_i^{\beta} \ln x_i\right)^2}{\left[\sum_{i=1}^{n} x_i^{\beta}\right]^2} + \frac{1}{\beta^2} \tag{13}$$

Once $\hat{\beta}_{mle}$ is obtained, the estimate of α follows from (9).

2.2. Method of Moments Estimator (MME)

In many cases the method of moments estimator (MME), which in many cases can be derived by hand.

Let x_1, x_2, …, x_n be a dataset for which we seek an unbiased estimator for the k^th moment. The MME is defined by computing the sample moments

$$\hat{M}_r = \frac{1}{n} \sum_{i=1}^{n} x_i^r \tag{14}$$

and by setting them equal to the theoretical moments from the moment generating function of Weibull distribution $M_r(t)$, where the k^th moment for our distribution is

$$\mu_k = \left(\frac{1}{\alpha^k}\right)^{-\frac{k}{\beta}} \Gamma\left(1 + \frac{k}{\beta}\right) \tag{15}$$

By equating two equations (14), (15) we can find the 1^st and 2^nd moments about zero as follows:

$$\hat{M}_1 = \hat{\mu}_1 = \left(\frac{1}{\alpha}\right)^{\frac{1}{\beta}} \Gamma\left(1 + \frac{1}{\beta}\right) \tag{16}$$

and

$$\hat{M}_2 = \hat{\mu}^2 + \hat{\sigma}^2 = \left(\frac{1}{\alpha}\right)^{\frac{2}{\beta}} \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma\left(1 + \frac{1}{\beta}\right)^2\right] \tag{17}$$

where,

$$\frac{\hat{M}_1^2}{\hat{M}_2^2}=\frac{\Gamma\left(1+\frac{2}{\beta}\right)-\Gamma^2\left(1+\frac{1}{\beta}\right)}{\Gamma^2\left(1+\frac{1}{\beta}\right)} \tag{18}$$

is a function with β only.

The coefficient of variation (CV) can found by taking the square root of eq. (18)

$$CV=\frac{\sqrt{\Gamma\left(1+\frac{2}{\beta}\right)-\Gamma^2\left(1+\frac{1}{\beta}\right)}}{\Gamma\left(1+\frac{1}{\beta}\right)} \tag{19}$$

This method can be used as an alternative to the maximize likelihood method that discussed in previously section. So, we can easily determine the value of α and β by the following equations:

$$\hat{\beta}_{mme}=\left(\frac{0.9874}{\frac{\sigma}{\bar{x}}}\right)^{1.0983} \tag{20}$$

and

$$\hat{\alpha}_{mme}=\frac{\bar{x}}{\Gamma\left(1+\frac{1}{\hat{\beta}}\right)} \tag{21}$$

2.3. Rank Regression Method (RRM) [1]

It known as the least-squares method and it is commonly applied in engineering and mathematics problems. The least square principle minimizes the vertical distance between the data points and the straight line fitted to the data[4]. We can get a linear relation between the two parameters. Recall equation (2) we have the cumulative density function of two parameters Weibull distribution:

$$F(x\mid\alpha,\beta)=1-\exp\left\{-\left(\frac{x}{\alpha}\right)^{\beta}\right\}, x\geq 0; \tag{22}$$

This function can be written as

$$\{1-F(x)\}^{-1}=\exp\left\{\left(\frac{x}{\alpha}\right)^{\beta}\right\} \tag{23}$$

If we take the natural logarithm of equation (23) we obtain

$$-\ln\{1-F(x)\}=\left(\frac{x}{\alpha}\right)^{\beta} \tag{24}$$

Again, take the natural logarithm of eq. (24) we get the following equation

$$\ln\left[-\ln\{1-F(x)\}\right]=\beta\ln x-\beta\ln\alpha \tag{25}$$

Here, eq. (25) can be written as $Y=bX+a$ where
$Y=\ln\left[-\ln\{1-F(x)\}\right]$, $X=\ln x$, $a=-\beta\ln\alpha$, $b=\beta$

By Linear regression formula

$$b=\frac{n\sum_{i=1}^{n}X_iY_i-\sum_{i=1}^{n}X_i\sum_{i=1}^{n}Y_i}{n\sum_{i=1}^{n}X_i^2-\left(\sum_{i=1}^{n}X_i\right)^2} \tag{26}$$

and

$$\hat{\alpha}=\exp\left(-\frac{c}{\hat{\beta}}\right) \tag{27}$$

$c=\bar{y}-\hat{\beta}\bar{x}$, $\hat{\beta}=b$

2.4. Mean-Standard Deviation Method (MSD)

This method can be estimating the Weibull parameters α and β from the mean and standard deviation σ of wind data, consider the expression for average and standard deviation can be written as follows:

$$\left.\begin{aligned}\mu&=E(x)=\alpha\Gamma(1+\frac{1}{\beta})\\ \sigma&=\alpha\sqrt{\Gamma(1+\frac{2}{\beta})-\left(\Gamma(1+\frac{1}{\beta})\right)^2}\end{aligned}\right\} \tag{28}$$

From the above equations we have[6,12]

$$\left(\frac{\sigma_x}{\bar{x}}\right)^2=\frac{\Gamma(1+\frac{2}{\beta})}{\Gamma^2(1+\frac{1}{\beta})}-1 \tag{29}$$

where $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ and $\sigma^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ which can calculated for a given data set, and in a simple approach with an acceptable approximation for β is [2]:

$$\hat{\beta}_{msd}=\left(\frac{\sigma_x}{\bar{x}}\right)^{-1.086} \tag{30}$$

The parameter α can determined by

$$\hat{\alpha}_{msd} = \frac{\bar{x}}{\Gamma\left(1+\frac{1}{\hat{\beta}}\right)} \quad (31)$$

2.5. Power Density Method (PDM)

Akdag and Ali (2009) suggested a new method. It is used to estimate the two-Weibull parameters, which depends on the energy pattern factor method; it is related to the averaged data of wind speed. Recalling the mean of Weibull probability distribution, wind speed of eq. (30) can be written as follows [10,11]:

$$\bar{x} = \alpha\Gamma\left(1+\frac{1}{\beta}\right) \quad (32)$$

To obtain the shape and scale factor of Weibull distribution through this method, we firstly compute the energy pattern factor. The energy pattern factor is related to the averaged data of wind speed and is defined as a ratio between mean of cubic wind speed to cube of mean wind speed.

The cubic mean wind speed is given as:

$$\bar{x}_{cub} = \alpha^3\Gamma\left(1+\frac{3}{\beta}\right) \quad (33)$$

Then the energy pattern factor (E_{pf}) can expressed as[3]:

$$E_{pf} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^3}{\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^3} = \frac{\overline{x^3}_{cub}}{(\bar{x})^3} = \frac{\Gamma\left(1+\frac{3}{\beta}\right)}{\Gamma^3\left(1+\frac{1}{\beta}\right)} \quad (34)$$

By using energy pattern factor (E_{pf}) in eq. (34) we can estimate the Weibull parameters from the follows formulas:

$$\hat{\beta}_{pd} = 1+\frac{3.69}{(E_{pf})^2}, \qquad \hat{\alpha}_{pd} = \frac{\bar{x}}{\Gamma\left(1+\frac{1}{\hat{\beta}}\right)} \quad (35)$$

2.6. Term Omission Method (TOM) [7]

This method has been discovered by Labban J.A., it was developed in 2005. This method involves omission terms of equation to get a simplified formula containing unknowns parameters that we want to estimate.

Recalling the probability density function

$$f(x\,|\,\alpha,\beta) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1}\exp\left\{-\left(\frac{x}{\alpha}\right)^{\beta}\right\}$$

and the cumulative distribution function (cdf) of our distribution is

$$F(x\,|\,\alpha,\beta) = 1-\exp\left\{-\left(\frac{x}{\alpha}\right)^{\beta}\right\}$$

with shape β and scale α parameters.

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n drawn from a population with probability density function sample. Here in our method we need to use the cumulative distribution function of Weibull distribution.

$$x_1 \qquad y_1 = 1-\exp\left[-\left(\frac{x_1}{\alpha}\right)^{\beta}\right]$$

$$x_2 \qquad y_2 = 1-\exp\left[-\left(\frac{x_2}{\alpha}\right)^{\beta}\right]$$

$$\vdots$$

$$x_n \qquad y_n = 1-\exp\left[-\left(\frac{x_n}{\alpha}\right)^{\beta}\right]$$

To obtain the estimation of parameters we have to do the following steps.

a. Subtracting 1 from y_i, i=1,2,…,n, so we obtain

$$y_i^1 = \exp\left[-\left(\frac{x_i}{\alpha}\right)^{\beta}\right]$$

b. Taking the natural logarithm

$$y_i^2 = -\left(\frac{x_i}{\alpha}\right)^{\beta}$$

c. Multiplying y_i^2 by (-1)

$$y_i^3 = \left(\frac{x_i}{\alpha}\right)^{\beta}$$

d. Calculate $y_j^4 = \frac{y_{i+1}^3}{y_i^3}$, j=1,2,…n-1, to obtain

$$y_j^4 = \frac{y_{i+1}^3}{y_i^3} = \frac{-\left(\frac{x_{i+1}}{\alpha}\right)^{\beta}}{-\left(\frac{x_i}{\alpha}\right)^{\beta}} = \left(\frac{x_{i+1}}{x_i}\right)^{\beta}$$

e. Again, on taken natural logarithm to y_j^4 we get

$$y_j^5 = \ln(y_j^4) = \beta\ln\left(\frac{x_{i+1}}{x_i}\right)$$

f. Finally, divided y_j^5 by $\ln\left(\frac{x_{i+1}}{x_i}\right)$ we get estimation of β as follow:

$$\beta^j = \frac{y_j^5}{\ln\left(\frac{x_{i+1}}{x_i}\right)}, \quad j=1,2,\ldots,n-1\ldots \qquad (36)$$

and

$$\hat{\beta} = \underset{1\le j\le n-1}{Min}\left[\sum_{i=1}^{n}\left\{F(t_i;\beta^j) - F(t_i;\beta)\right\}^2\right] \qquad (37)$$

and from step (c) for any i=1,2,…,n we get [8]:

$$y_i^1 = \exp\left[-\left(\frac{x_i}{\alpha}\right)^{\hat{\beta}}\right] \Rightarrow \hat{\alpha} = \frac{x_i}{\left(-\ln y_i^1\right)^{\frac{1}{\hat{\beta}}}}$$

$$i = 1,2,\ldots,n \qquad (38)$$

3. Simulation Results

In this section we will use simulation studies to evaluate the performance of the proposed methods and compare between them.

Random samples were generated of 2-Parameter Weibull distribution with two specific parameters values, where α>0 (scale parameters) and β>0 (shape parameter) (1, 1 and 1.5, 2.1) respectively, where the sample was (n=10, 50, 100, 200). The tables below show the summary of estimation of parameters for all six methods in this study and performance between them using the test criterion Mean Squared Error (MSE) as follows:

$$MSE = \frac{\sum_{i=1}^{n}\{\hat{F}(t_i) - F(t_i)\}^2}{n} \qquad (39)$$

4. Conclusion

In the previous tables 1 and 2, and according to Mean Square Error we can conclude the following conclusions:

In table 1 we observe that:

1. The preference TOM and RRM over other methods.
2. The preference of the rest methods is MLE, MME, MSD and PDM

respectively according to Mean Square Error.

While in table 2 we observe that:

1.The preference of TOM over other methods.

2.The preference of the rest methods is MSD, RRM, PDM, MME, and MLE respectively according to Mean Square Error.

Table (1): The estimate value of parameters that found by six methods for 2-Parameter Weibull distribution with (α,β)=(1,1) with (n=10, 50, 100, 200).

Method	N	α	β	MSE
		Estimated	Estimated	
MLE	10	1.30617001	0.376973	0.022504
	50	1.222645	1.375659	0.005598
	100	0.918627	1.054843	0.000269
	200	0.945383	0.939041	7.21E-05
MME	10	1.470573	1.666277	0.030062
	50	1.250265	1.304994	0.009666
	100	0.886699	1.041207	0.000225
	200	0.849048	0.88235	0.000621

RRM	10	1	1	0
	50	1	1	0
	100	1	1	0
	200	1	1	0
MSD	10	1.484538	1.668303	0.030569
	50	1.264434	1.308291	0.010101
	100	1.05	0.900205	0.00018
	200	0.862399	0.890857	0.000524
PDM	10	1.582975	1.680244	0.034093
	50	1.30614	1.317203	0.011399
	100	1.070907	1.134387	0.00071
	200	1.033217	0.97345	1.3E-05
TOM	10	1	1	0
	50	1	1	0
	100	1	1	0
	200	1	1	0

Table (2): The estimate value of parameters that found by six methods for 2-Parameter Weibull distribution with (α,β)=(2.1,1.5) with (n=10, 50, 100, 200).

Method	N	α	β	MSE
		Estimated	Estimated	
MLE	10	1.986616	1.132828	0.009734
	50	3.446153	1.590979	0.045521
	100	2.09844	1.281535	0.001822
	200	2.788626	1.45551	0.004062
MME	10	1.838974	1.149549	0.005226
	50	2.330993	1.652145	0.002535
	100	1.79157	1.30186	0.002646
	200	2.017377	1.448499	9.71E-05
RRM	10	2.1	1.5	0
	50	1.468894	1.524193	0.015135
	100	2.074487	1.530132	9E-06
	200	1.810674	1.522195	0.000802
MSD	10	1.845422	1.16367	0.004866
	50	2.332464	1.665661	0.002636
	100	1.795395	1.316018	0.00251
	200	2.020022	1.462503	8.27E-05
PDM	10	1.87497	1.237592	0.003302
	50	2.332987	1.670627	0.002674
	100	1.795189	1.315238	0.002518
	200	2.019584	1.460142	8.5E-05
TOM	10	2.1	1.5	0
	50	2.1	1.5	0
	100	2.1	1.5	0
	200	2.1	1.5	0

References

[1] Akdağ, S.A.; Ali, D.; 2009, "A new method to estimate Weibull parameters for wind energy applications", Energy Convers. Manag. 50:1761-1766.

[2] Akhlaque, A.M.; Firoz, A.; and Wasim, A.M.; 2006, "Assessment of wind power potential for coastal areas of Pakistan", Turk. J. Phys. 30:127-135.

[3] Rocha, C.; de Sousa, P.A.; de Andrade, R.C.; da Silva, C.F.; 2012, "Comparison of seven numerical methods for determining Weibull parameters for wind energy generation in the northeast region of Brazil", *Appl. Energy*, *89*, 395–400.

[4] Egwim, K.C.; Eke C.N.; Onuoha, D.O.; Igbo, C.A.; 2015, "Comparative Analysis of Methods of Estimating 2-Parameter Weibull Distribution", Vol 5(2), 10-16.

[5] Felix, N.N.; Chukwudi A.; 2014, "A Comparison of Methods for the Estimation of Weibull Distribution Parameters", Metodološki zvezki, Vol 11(1), 65-78.

[6] Fung, CH.; Yim, HL.; Lau, KH.; Kot, SC.; 2007, "Win energy potential in Guangdong providence", J. Geophys. Res. 112:1-12.

[7] Labban, J.A.; 2005, "Modified and Simplified Method for Finite Difference (Divided)", Al-Qadisiya Journal for Science. Vol. 10(2), 244-251.

[8] Labban, J.A.; 2013, "Estimation of Weibull distribution Parameters with Complete Data Using T.O Method", Journal of Science of Computer and Mathematic, Vol 5(2), 123-129.

[9] Mark, A.N.; 2011, "Parameter Estimation for the Two-Parameter Weibull Distribution", Brigham Young University–Provo, Master Thesis.

[10] Paula, A.C.; Ricardo, C.S.; Carla, F.D.; Maria, E.V.; 2012, "Comparison of seven numerical methods for determining Weibull parameters for wind energy generation I the northeast region of Brazil", Appl. Energy 89:395-400.

[11] Silva, G.; Alexandre, P.; Daniel, F.; Everaldo, F.; 2004, "On the accuracy of the Weibull parameters estimators. European Wind Energy Conference and Exhibition (EWEC)", London UK.

[12] Weisser, D.; and Foxon, T.J.; 2003, "Implications of seasonal and diurnal variations of wind velocity for power output estimation of a turbine: A case study of grenada", Int. J. Energy Res. 2:1165-1179.