Estimation Accuracy of Weibull Distribution Parameters

Ahmad Mahir Razali, Ali A. Salih and Asaad A. Mahdi

School of Mathematical Sciences Faculty of Science and Technology Universiti Kebangsaan
Malaysia 43600 UKM Bangi, Selangor, Malaysia

Abstract: The Weibull variate is commonly used as a lifetime distribution in reliability applications. The two-parameter (shape and scale) Weibull distribution can represent a decreasing, constant or increasing failure rate, Hvans et al[3]. Further flexibility can be introduced into the Weibull distribution by adding the third parameter which is a location parameter. In this paper, we compared between three well known methods for estimating the parameters of Weibull distributions. These methods are; moments, maximum likelihood and least squares. We generated a set of data for the 2-parameter Weibull distribution, and another set for the 3-parameter Weibull distribution and we used these methods to estimate the parameters. We used the means square error, MSE and total deviation, TD as measurement for the comparison between these methods. We found that the moments method is the best method for estimating the parameters of the2-parameter and 3-parameter Weibull distributions because they gave the least value for the mean square error.

Key words: Weibull Distribution, parameter estimation, maximum likelihood, least square, moments method

INTRODUCTION

Weibull distribution is named after Walodi Weibull (1887 – 1979). It is very flexible and can through an appropriate choice of parameters, model many types of failure rate behaviors. This distribution can be found with two or three parameters; scale, shape and location parameters. The values of these parameters must be more than zero except for the location parameter which can be greater than and equal to zero. There are a number of methods for estimating the values of these parameters; some are graphical and others are analytical. In this article we will try to find the best analytical method to estimate the parameters of Weibull distribution. We will use the mean square error, MSE and total deviation, TD, as measurements to determine the accuracy.

Graphical methods include Weibull probability plotting, WPP, and hazard plot. These methods are not very accurate but they are relatively fast. Analytical methods include maximum likelihood method, MLM, least square method, LSM, and method of moments, MOM. These methods are considered more accurate and reliable compared to the graphical method. In this article we will compare between these methods using two and three parameter Weibull distributions to find the most accurate method.

Parameter Estimation: The probability density functions of the two and three parameter Weibull distributions are as follows, Ahmad and Ali[1];

$$f(x) = (\beta \frac{x^{\beta-1}}{\alpha^{\beta}})e^{-(\frac{x}{\alpha})^{\beta}} , x \geq 0 \qquad (1)$$

and

$$f(x) = [\frac{\beta(x-\gamma)^{\beta-1}}{\alpha^{\beta}}]e^{-[\frac{(x-\gamma)}{\alpha}]^{\beta}} , x \geq 0, \gamma \geq 0 \qquad (2)$$

The three parameters $\alpha > 0$, $\beta > 0$, and $\gamma \geq 0$ are the scale, shape and location parameters of the distribution respectively.

1. Graphical Methods
Weibull Probability Plotting, WPP: For the two parameter Weibull distribution, under the Weibull transformation, Murthy et al[7];

$$y = \beta x - \beta \ln(\alpha) \qquad (3)$$

This equation refers to a straight line with a slope β. The intercept with the y- axis is $-\beta \ln(\alpha)$ and with the x-axis is $\ln(\alpha)$.

Corresponding Author: Ahmad Mahir Razali, School of Mathematical Sciences Faculty of Science and Technology Universiti Kebangsaan Malaysia 43600 UKM Bangi, Selangor, MALAYSIA
E-mail: mahir@ukm.my ; wakilali@yahoo.com ; asaadmahdi@gmail.com

We will take the natural logarithm to the cumulative distribution function, cdf, to find the relation between cdf and the parameters.

$$F(x) = 1 - e^{-(\frac{x}{\alpha})^\beta} \qquad (4)$$

$$1 - F(x) = e^{-(\frac{x}{\alpha})^\beta}$$

$$\ln[1 - F(x)] = -(\frac{x}{\alpha})^\beta$$

$$\ln[-\ln(1 - F(x))] = -\beta \ln(\alpha) + \beta \ln(x) \qquad (5)$$

Also we need to rank the failure times from smallest to largest, and calculate the $F(x_i)$ by anyone of these approaches.

$$F(x_i) = \frac{i}{n}$$

$$F(x_i) = \frac{i}{n+1} \quad \text{called the mean rank approach}$$

$$F(x_i) = \frac{i - 0.5}{n} \quad \text{called the medium rank estimator}$$

$$F(x_i) = \frac{i - 0.3}{n + 0.4}$$

$$F(x_i) = \frac{i - 0.375}{n + 0.25}$$

Then we plot $F(x_i)$ versus x_i where n is the sample size and x_i is the failure times.

Hazard Plot: The cumulative hazard function H(x) is $(\frac{x}{\alpha})^\beta$, Al-Fawzan[2]

H(x) can be plotted versus failure times on hazard paper. The cumulative hazard function gets transformed into a linear relationship by taking logarithm.

$$H(x) = (\frac{x}{\alpha})^\beta \qquad (6)$$

$$\ln[H(x)] = \ln[(\frac{x}{\alpha})^\beta]$$

$$\ln[H(x)] = \beta[\ln x - \ln \alpha]$$

$$\ln(x) = \frac{1}{\beta} \ln[H(x)] + \ln \alpha \qquad (7)$$

Before plotting H(x) versus x_i we need to perform the following steps;

1. Rank the failure times from smallest to largest; $x_1 \leq x_2 \leq \ldots \leq x_n$
2. Find the hazard value for each failure;

$$\Delta H_i = \frac{1}{(n+1) - 1}$$

3. Find the H(x) by summing up the hazard values

$$H(x) = \Delta H_1 + \Delta H_2 + \ldots + \Delta H_i$$

4. Find ln(x) for $1 \leq i \leq n$
5. Plot ln[H(x)] versus $\ln(x_i)$ where $1 \leq i \leq n$
6. Fit a straight line

For the three parameter Weibull distribution we can also estimate the parameters graphically and analytically.
To use the WPP plot the Weibull transformation gets transformed into x, Murthy *et al*[7]

$$y = \beta \ln(e^x - \gamma) - \beta \ln(\alpha) \qquad (8)$$

where y is a non-linear function of x. Since WPP plot is a smooth curve, the asymptote of the WPP Plot are as follows;

For the 2-parameter Weibull distribution; as $x \to \infty$ asymptote is a straight line given by $y = \beta x - \beta \ln(\alpha)$.

As $x \to \ln(\gamma)$, the asymptote is a vertical line that intercepts the x-axis at $\ln(\gamma)$

There are a number of methods proposed to estimate the parameters from the plot. The simplest method to estimate γ is given by $\hat{\gamma} = x_1$ [a better estimate is $x_1 - \frac{1}{n}$]. Using this method the transformed data (given by $x_i - \hat{\gamma}$) is viewed as data generated from two-parameter Weibull distribution. The scale and shape parameters are obtained using the same procedure in the hazard plot. O'Connor suggests to estimate the location parameter by;

$$\hat{\gamma} = x_2 - \frac{(x_3 - x_2)(x_2 - x_1)}{(x_3 - x_2) - (x_2 - x_1)} \tag{9}$$

Also, Kececioglu[5] discussed two other methods for estimating the location parameter.

Li[6] proposed a method that depends on a two step iterative procedure. In the first step, he assumes that the location parameter is known, and the scale and shape parameters are estimated using the graphical method. In the second step, the shape parameter is assumed known, and the scale and location parameters are estimated by transforming the data using power law transformation so that the transformed data can be modeled by a two-parameter exponential distribution.

Jiang and Murthy[4] proposed another method which is a modification of the method proposed by Li[6].

2. Analytical Method: There are a number of analytical methods used to estimate the parameters of Weibull distribution. We will discuss some of these methods used for 2 and 3-parameter Weibull distribution;

Method of Moments, MOM: For the 2-parameter Weibull distribution, we can use the sample mean $\overline{x}$ and sample variance S^2, where

$$\vec{x} = \sum_{i=1}^{n} \frac{x_i}{n} \tag{10}$$

and $S^2 = \sum_{i=1}^{n} \frac{(x_i - x)^2}{n-1} \tag{11}$

We know that

$$\mu = E(x) = \alpha \Gamma(1 + \frac{1}{\beta}) \tag{12}$$

and

$$\sigma^2 = \alpha^2 \{ \Gamma(1 + \frac{2}{\beta}) - [\Gamma(1 + \frac{1}{\beta})]^2 \} \tag{13}$$

We can get $\hat{\beta}$ by dividing the variance on the square mean

$$\hat{\beta} = \frac{\Gamma(1 + \frac{2}{\hat{\beta}})}{\Gamma^2(1 + \frac{1}{\hat{\beta}})} - 1 \tag{14}$$

and

$$\hat{\alpha} = \frac{\overline{x}}{\Gamma(1 + \frac{1}{\hat{\beta}})} \tag{15}$$

For the 3-parameter Weibull distribution we will depend on the mean, variance and the third moment;

$$\mu = E(x) = \gamma + \alpha \Gamma(1 + \frac{1}{\beta}) \tag{16}$$

$$\sigma^2 = \alpha^2 \{ \Gamma(1 + \frac{2}{\beta}) - [\Gamma(1 + \frac{1}{\beta})]^2 \} \tag{17}$$

$$\mu_3 = \alpha^3 \{ \Gamma(1 + \frac{3}{\beta}) - 3\Gamma(1 + \frac{1}{\beta})\Gamma(1 + \frac{2}{\beta}) + 2[\Gamma(1 + \frac{1}{\beta})]^3 \}$$

By solving these three equations simultaneously, we obtain the parameter estimates. The coefficient of skewness, δ, can also be found;

$$\delta^2 = \frac{\{ \Gamma(1 + \frac{3}{\beta}) - 3\Gamma(1 + \frac{1}{\beta})\Gamma(1 + \frac{2}{\beta}) + 2[\Gamma(1 + \frac{1}{\beta})]^3 \}^2}{\{ \Gamma(1 + \frac{2}{\beta}) - [\Gamma(1 + \frac{1}{\beta})]^2 \}^3} \tag{18}$$

Maximum Likelihood Method, MLM: The likelihood function can be shown as follows;

$$L(\alpha, \beta) = \prod_{i=1}^{n} f(x_1, \alpha, \beta) \tag{19}$$

The maximum likelihood estimator, MLE of the parameter is the value of the parameter that maximizes L and MLM for the 2-parameter Weibull distribution can be obtained by solving the equations resulting from setting the two partial derivatives of L(α,β) to zero;

$$L(\alpha, \beta) = \prod_{i=1}^{n} (\frac{\beta x_i^{\beta-1}}{\alpha^\beta}) e^{-(\frac{x_i}{\alpha})^\beta} \tag{20}$$

$$\frac{\partial \ln L}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^{n} \ln x_i - \frac{1}{\alpha} \sum_{i=1}^{n} x_i^\beta \ln x_i = 0 \tag{21}$$

$$\frac{\partial \ln L}{\partial \alpha} = -\frac{n}{\alpha} + \frac{1}{\alpha^2} \sum_{i=1}^{n} x_i^{\beta} = 0 \qquad (22)$$

Then $\hat{\beta}$ is the solution of

$$\frac{\sum_{i}^{n} (x_i^{\hat{\beta}} \ln x_i)}{\sum_{i=1}^{n} x_i^{\hat{\beta}}} - \frac{1}{\hat{\beta}} - \frac{1}{n} \sum_{i=1}^{n} \ln x_i = 0 \qquad (23)$$

When the shape parameter is estimated then the scale parameter can also be estimated by;

$$\hat{\alpha} = \left(\frac{1}{n} \sum_{i=1}^{n} x_i^{\hat{\beta}}\right)^{\frac{1}{\hat{\beta}}} \qquad (24)$$

For the 3-parameter Weibull distribution, we know that the likelihood function is given by;

$$L(\alpha, \beta, \gamma) = \frac{\beta^n}{\alpha^{\beta n}} \left[\prod_{i=1}^{n} (x_i - \gamma)^{\beta-1}\right] e^{-\frac{1}{\alpha^{\beta}} \sum (x_i - \gamma)^{\beta}} \qquad (25)$$

By setting the partial derivative of equation (25) with respect to α, β and γ to zero, and solving the following set of equations simultaneously, we get the estimate of the parameters;

$$\hat{\alpha}^{\hat{\beta}} - \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\gamma})^{\hat{\beta}} = 0 \qquad (26)$$

$$\frac{\sum_{i=1}^{n} (x_i - \hat{\gamma})^{\hat{\beta}} \ln(x_i - \hat{\gamma})}{\sum_{i=1}^{n} (x_i - \hat{\gamma})^{\hat{\beta}}} - \frac{1}{\beta} - \frac{1}{n} \sum_{i=1}^{n} \ln(x_i - \hat{\gamma}) = 0 \qquad (27)$$

$$\left((\hat{\beta}-1) \sum_{i=1}^{n} (x_i - \hat{\gamma})^{-1} - \hat{\beta}\hat{\alpha}^{-\hat{\beta}} \sum_{i=1}^{n} (x_i - \hat{\gamma})^{\hat{\beta}-1}\right) = 0 \qquad (28)$$

Least Squares Method, LSM: The linear equation is as follows;

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

then;

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} \ln\left\{\ln\left[\frac{1}{(1 - \frac{i}{n+i})}\right]\right\}$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} \ln x_i$$

$$\hat{\beta} = \frac{\left\{n \sum_{i=1}^{n} (\ln x_i) \cdot \left(\ln\left[\ln\left[\frac{1}{1 - \frac{i}{n+1}}\right]\right]\right)\right\} - \left\{\sum_{i=1}^{n} \ln\left[\ln\left(\frac{1}{(1 - \frac{i}{n+1})}\right)\right]\right\} \sum_{i=1}^{n} \ln x_i}{\left\{n \cdot \sum_{i=1}^{n} (\ln x_i)^2\right\} - \left\{\sum_{i=1}^{n} (\ln x_i)\right\}^2} \qquad (29)$$

then,

$$\hat{\alpha} = e^{\left(\frac{\bar{y} - \bar{x}}{\hat{\beta}}\right)} \qquad (30)$$

For the 3-parameter Weibull distribution;
Let $X_1 \le X_2 \le \ldots \le X_n$ be the order statistics in random sample of Weibull distribution with p.d.f.

$$f(x, \alpha, \beta, \gamma) = \beta / \alpha \left[\frac{(x - \gamma)}{\alpha}\right]^{\beta-1} \exp\left\{-\left[\frac{(x - \gamma)}{\alpha}\right]^{\beta}\right\} \qquad (31)$$

where $x \ge \gamma$, $\alpha > 0$, $\beta > 0$

Now, new random variable will be defined, Yildirim[8];

$$U_i = 1 - \exp\left[-\left[\frac{(x_i - \gamma)}{\alpha}\right]^{\beta}\right] \qquad (32)$$

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

$$\log[-\log(1 - U_i)] = \beta[\log(x_i - \gamma)] - \beta \log \alpha \qquad (33)$$

U_i is an ith order statistic of a uniform distribution over the interval (0,1)

We can also obtain similar relations in equation (33) by using the hazard function

$$\log h(x) = \log[(\beta/\alpha)^\beta] + (\beta - 1)\log(x - \gamma)$$
(34)

or the cumulative hazard function

$$\log H(x) = \beta\log(x - \gamma) - \beta\log\alpha$$
(35)

or the survival function

$$\log[-\log S(x)] = \beta\log(x - \gamma) - \beta\log\alpha$$
(36)

and

$$a = -[-\beta\log\alpha], \quad b = \beta, \quad t_i = \log(x_i - \gamma)$$

Let e_i be the corresponding random error, then we get;

$$Y_i = a + bt_i + e_i, \qquad i=1,2,\dots,n$$
(37)

From this model, the least square estimation method can be used to estimate the parameters a and b;

$$E(Y) = E\{\log[-\log(1 - U_i)]\} = \log\{-\log[1 - E(U_i)]\}$$
(38)

By minimizing the sum square of errors we can get the other parameters;

$$Q_i(\alpha, \beta) = \sum\{\log[-\log(1 - U_i)] - \log[-\log[1 - E(U_i)]]\}^2$$

$$i=1,2,\dots,n$$
(39)

The least squares estimates of the α and β of Weibull distribution can be obtained by choosing the proper estimates of $E(\mu_i)$ and the survivor function $s(x_i)$. These estimates are;

$$\hat{\alpha}(\gamma) = \{\prod_{i=1}^{n}[x_i - \gamma]\}^{\frac{1}{n}}\{\prod_{i=1}^{n}[-\log[E(\mu_i)]]\}^{-\frac{1}{n}\beta(\gamma)}$$
(40)

and

$$\hat{\beta}(\gamma) = \frac{\sum_{i=1}^{n}(R_i - \bar{R}_i)[\log(x_i - \gamma) - (\log x)^t]}{[\sum_{i=1}^{n}[\log(x_i - \gamma) - (\log x)^t]^2]}$$
(41)

where i=1,2,...,n, $R_i = \log[-\log E(\mu_i)]$

$$\bar{R} = \frac{1}{n}(\sum_{i=1}^{n} R_i), (\log x)^t = (\frac{1}{n})\{\sum_{i=1}^{n} \log(x_i - \gamma)\}$$

Application: Table 1 below includes 20 generated data. These data were used to estimate the parameters of the 2-parameter Weibull distribution using the three methods illustrated above. The values of the scale parameter; $\alpha = 1$ and of the shape parameter $\beta = 3$. The results of the mean square error and total deviation are tabulated in tables 2 and 3.

Table 4 below includes 20 generated data. These data were used to estimate the parameters of the 3-parameter Weibull distribution using the three methods illustrated above. The values of the scale parameter; $\alpha = 1$, the shape parameter $\beta = 2$ and the location parameter $\gamma = 3$. The results of the mean square error and total deviation are tabulated in tables 5 and 6.

conclusion: Total deviation, TD, and mean square error MSE, are two measurements that give an indication of the accuracy of parameter estimation. For the 2 parameter Weibull distribution, Table 2 shows that the minimum value of the MSE is 0.0252 corresponding to the methods of moments MOM. Also, the minimum value of TD is 0.1485 corresponding to the maximum likelihood method, MLM. For the 3-parameter Weibull distribution, Table 5 shows that the minimum value for the MSE is 0.0239 corresponding to the MOM while Table 6 shows that the minimum value of TD is 0.9862 corresponding also to MOM. By repeating this experiment a number of times using different sample sizes and different values of parameters we can decide which method is the most suitable for estimation. Here, initially we can say that MOM is the best method used to estimate the parameter for the 2 and 3 parameter Weibull distributions taking into consideration the MSE as a measurement for comparison.

Table 1: Generated data for the 2-parameter Weibull distribution

0.2684	0.3842	0.4521	0.5590	0.5633	0.5864	0.6252	0.6762	0.6947	0.8330
0.8503	0.8802	1.0100	1.0150	1.0210	1.0300	1.1410	1.2210	1.4670	1.5530

Table 2: MSE for the 2-parameter Weibull distribution

Method	$\hat{\alpha}$	$\hat{\beta}$	MSE
MOM	0.9463	2.6160	0.02523
MLM	0.9481	2.7100	0.03013
LSM	1.2836	3.4816	1.30628

Table 3: Total deviation for the 2-parameter Weibull distribution

Method	$\hat{\alpha}$	$\hat{\beta}$	TD
MOM	0.9463	2.6160	0.1817
MLM	0.9481	2.7100	0.1485
LSM	1.2836	3.4816	0.4441

Table 4: Generated data for the 3-parameter Weibull distribution

3.369	3.376	3.423	3.426	3.469	3.500	3.520	3.544	3.603	3.743
3.769	3.781	3.847	3.866	3.891	3.935	4.148	4.363	4.410	5.086

Table 5: MSE for the 3-parameter Weibull distribution

Method	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	MSE
MOM	0.5300	1.1685	3.3012	0.0239
MLM	0.4087	0.8636	3.3684	0.0470
LSM	0.3176	1.4141	3.3324	0.4474

Table 6: Total deviation for the 3-parameter Weibull distribution

Method	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	TD
MOM	0.5300	1.1685	3.3012	0.9862
MLM	0.4087	0.8636	3.3684	1.2823
LSM	0.3176	1.4141	3.3324	1.0862

REFERENCES

1. Ahmad Mahir Razali and Ali A. Salih Al-Wakil, 2007. Evaluating the Bi-Weibull Distribution Function with a Number of Parameters. Proceedings of ICMS07, Universiti Kebangsaan Malaysia, pp: 805-810.
2. Al-Fawzan, M.A., 2000. Methods for Estimating the Parameters of the Weibull Distribution, King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia.
3. Hvans, M., N.A.J. Hasting, and J.B. Peacock, 2000. Statistical Distributions, 3rd Eds., Wiley Series in Probability and Statistic. New York.
4. Jiang, R. and D. Murthy, 1997. Comments on a General Linear-Regression Analysis Applied to the 3-parameter Weibull Distribution, IEEE Transaction on Reliability, 46: 389-393.
5. Kececioglu, D., 1991. Reliability Engineering Handbook, Prentice Hall, Eaglewood Cliffs, NJ.
6. Li, Y.M., 1994. A General Linear-Regression Analysis applied to the 3-Parameter Weibull Distribution, IEEE, Transaction on Reliability, 43: 255-263.
7. Murthy, D.N., M. Xie and R. Jiang, 2004. *Weibull Models*, Wiley Series in Probability and Statistics. New York.
8. Yildirim, F., 1990. Least Squares Estimation for the Weibull Parameters, Pakistan Journal of Statistics, 6(1)A: 93-104.