IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 13, Issue 1 Ver. III (Jan. - Feb. 2017), PP 42-45
www.iosrjournals.org

Comparing the methods of Estimation of Three-Parameter Weibull distribution

Ramakrishnan.M[1] and Viswanathan. N[2]
[1]*Department of Mathematics, RKM Vivekananda College, Chennai, India*
[2]*Department of Statistics, Presidency College, Chennai, India*

Abstract: *Weibull distribution has many applications in engineering and plays an important role in reliability. Estimation of the location, scale and shape parameters of this distribution for both censored and non censored samples were considered by several authors. In this paper we compare Graphical oriented methods, "trial and error" approach, the approach of Jiang/Murthy and Maximum likelihood method developed by Bain & Engelhard for sample sets containing uncensored and censored sample. Importance of each method is discussed.*
Keywords: *Weibull distribution, Trial and error, Estimators and Maximum likelihood.*

I. Introduction

Weibull (1951) [1] publication marked the start of triumphant progress of Weibull distribution in statistical theory and as well as in applied statistics. Hundreds of authors around the globe have contributed to its development. Hundreds or even thousands of papers have been written on this distribution and the research is ongoing. Three parameter Weibull distributions is a physical model in reliability theory and survival analysis. The estimation of parameters of this distribution has been studied widely in the statistical literature. Lawless (1982) [2], David (1975) [3], Jiang/Murthy (1997) [4], Bartkute & Sakalauskas (2008) [5] and Bain and Engelhardt (1991) [6] have estimated the three parameters using graphical and Maximum Likelihood approaches. Our aim is to compare all the methods for the real data with and without censoring and identify the importance of censoring and effects of these estimators in Reliability estimation. Section 2 briefs three-parameter Weibull Distribution and its important functions used in Reliability. Section 3 contains the methods of estimating the location, scale and shape parameters of Weibull distribution. In Section 4, the methods are applied for the simulated data.

II. Three-Parameter Weibull Distribution

A random variable X has a three-parameter Weibull distribution with parameters a, b and c if its density function is given by

$$f_X(x|a,b,c) = \frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1} exp\left\{-\left(\frac{x-a}{b}\right)^{c}\right\}, x \geq a$$

The distribution of X is denoted as $X \sim We(a,b,c)$, Here 'a' is a location parameter or failure free time, 'b' is a scale parameter and 'c' is a shape parameter (Rinne) [3]. Changing 'a' when the other parameters are held constant will result in a parallel movement of the density curve over the abscissa. Enlarging (reducing) a causes a movement of the density to the right (to the left). Changing 'b' while 'a' and 'c' are held constant will alter the density at x in the direction of the ordinate. Enlarging 'b' will cause a compression or reduction of the density and reducing 'b' will magnify or stretch it while the scale on the abscissa goes into the opposite direction. The shape parameter is responsible for the appearance of a Weibull density.

Functions of three parameter Weibull distribution

1. The CDF of the three-parameter version is $F_X(x|a,b,c) = 1 - exp\left\{-\left(\frac{x-a}{b}\right)^{c}\right\}$
2. The reliability function is the complementary function to F:
3. $R_X(x|a,b,c) = exp\left\{-\left(\frac{x-a}{b}\right)^{c}\right\}$
4. Hazard rate is $h_X(x|a,b,c) = \frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1}$.

III. Estimation Procedures

3.1. Trial and Error Method (David, 1975)

Plotting $\hat{u}_{i:n} = ln\{-ln[1-\hat{F}(x_{i:n})]\}$ against $x^*_{i:n} = ln(x_{i:n})$ when the $x_{i:n}$ are sampled from $F(x) = 1 - exp\{-[(x-a)/bc$ will result into an approximately convex (concave) curve if a < 0 (a > 0). The curved plot can be converted into a fairly straight line when 'a' is known. The value of 'a' can be obtained by trial and error method.

$F(x) = 1 - exp\{-[(x-a)/b]^c\}$ is transformed to

Regression Model 1

$$\ln(x-a) = \frac{1}{c}\ln\{-ln[1-F(x)]\} + \ln b, \text{ Where } \hat{F}(x_{i:n}) = i/(n+1) \quad (1)$$

Regression Model 2

$$\ln\{-ln[1-F(x)]\} = c\ln(x-a) - c\ln b \quad (2)$$

First find the Residual Sum of Squares RSS $(\widehat{a})$ for the equation (2). From the minimum value of RSS $(\widehat{a})$, the value of $\hat{a}$ may be preferred. From the above two models $\hat{b}$ and $\hat{c}$ estimated.

3.2. Graphical Method

Jiang/Murthy (1997) has proposed a method, which is based on a Taylor's series expansion and which simultaneously estimates all three Weibull parameters. A plot of $u = \ln\{-ln[1-F(x)]\}$ versus $x^* = \ln(x)$ is a curve that intersects the x^* - axis at x_0^*. Re–transforming x_0^*. gives $x_0 = \exp(x_0^*.)$. When $\ln\{-ln[1-F(x)]\} = 0$ leads to $x_0 = a + b$

Reliability function may be rewritten as

$$R(x) = \exp\left[-\left(\frac{x - x_0 + b}{b}\right)^c\right]$$

$-\ln R(x) = \left(\frac{z}{b} + 1\right)^c$, where $z = x - x_0$

When $\frac{z}{b} < 1$ The Taylor's series expansion of the above equation is

$-lnR(x) \approx 1 + c\frac{z}{b} + \frac{c}{2}(c-1)\left(\frac{z}{b}\right)^2$ or

$y \approx \alpha + \beta z$, where $y = -\frac{\ln R(x)+1}{z}, \alpha = \frac{c}{b}, \beta = \frac{c(c-1)}{2b^2}$

By plotting y_i versus z_i and fitting a straight line, one can find estimates $\hat{\alpha}$ and $\hat{\beta}$ which may be re–transformed to estimates $\hat{a}$, $\hat{b}$ and $\hat{c}$ using above equations. We can get $\hat{c} = \frac{\hat{\alpha}^2}{\hat{\alpha}^2 - 2\beta}$, $\hat{b} = \frac{\hat{c}}{\hat{\alpha}}$, $\hat{a} = x_0 - \hat{b}$.

3.3. Maximum Likelihood Method:

In many applications the location parameter is assumed known, and thus may be taken to be zero, without loss of generality, by simply translating the data. Convert the data of Weibull Three parameter distribution into Two parameter Weibull distribution by the way of subtracting the values by location parameter a. Then $x \sim Wei(b, c)$

The Likelihood function for the first r ordered observations from a sample of size n is given by

$$L = f(x_{1:n}, \ldots, x_{r:n}) = \frac{n!}{(n-r)!}\left[\prod_{i=1}^{r} f_x(x_{i:n})\right][1 - F(x_{r:n})]^{n-r}$$

$$= \frac{n!}{(n-r)!}\left(\frac{c}{b}\right)^r \left(\prod_{i=1}^{r} \frac{x_{i:n}}{b}\right)^{c-1} \exp\left[-\left[\sum_{i=1}^{r}\left(\frac{x_{i:n}}{b}\right)^c + (n-r)\left[\frac{x_{r:n}}{b}\right]^c\right]\right]$$

On solving $\frac{\delta(logL)}{\delta b} = 0$ and $\frac{\delta(logL)}{\delta c} = 0$, it is seen that $\hat{c}$ is the solution of the equation

$$\frac{\sum_{i=1}^{r} x_{i:n}^{\hat{c}} \ln x_{i:n} + (n-r)x_{r:n}^{\hat{c}} \ln x_{r:n}}{\sum_{i=1}^{r} x_{i:n}^{\hat{c}} + (n-r)x_{r:n}^{\hat{c}}} - \frac{1}{\hat{c}} = \frac{1}{r}\sum_{i=1}^{r} \ln x_{i:n}$$

and $\hat{b} = \left[\frac{\sum_{i=1}^{r} x_{i:n}^{\hat{c}} + (n-r)x_{r:n}{}^{\hat{c}}}{r}\right]^{1/\hat{c}}$, To find $\hat{c}$, the Newton-Raphson method may be used.

IV. Implementation

4.1. Data

The dataset consisting of n=20 observations simulated from Weibull distribution having parameters a=15, b=30 and c=2.5 is taken and its ordered dataset is given in Table 1

Table 1: Ordered dataset from We (15, 30, 2.5)

i	1	2	3	4	5	6	7	8	9	10
$x_{i:20}$	22.8	26.3	28.8	30.9	32.4	34.4	35.6	37.3	38.5	39.9
i	11	12	13	14	15	16	17	18	19	20
$x_{i:20}$	41.6	43.5	44.7	46.2	48.4	49.5	52.0	53.8	57.3	66.4

To select the value of $\hat{a}$ by "trial and error approach", the Residual Sum of Squares (RSS) of the different values of 'a' is given below:

Table 2: Residual sum of squares for different values of 'a'

n (Total Samples)	20	20	20
m (Number of Samples taken)	20	15	10
a	**RSS**	**RSS**	**RSS**
5	0.172484	0.031631	0.003971
8	0.114322	0.016624	0.002800
9	0.096553	0.012500	**0.001900**
10	0.080339	0.009399	0.002330
11	0.066382	0.007428	0.003296
12	0.055656	**0.007170**	0.005100
13	**0.049544**	0.009341	0.008000
14	0.050041	0.014982	0.012686
15	0.060117	0.025658	0.019668
16	0.084336	0.043797	0.030137
20	0.588307	0.334334	0.173375

The least Residual Sum of Square values for complete data is 0.049544 and the data containing first 15 and 10 samples are 0.007170 and 0.001900 respectively. To select the approximate value of estimator 'a' graphical method also used. The following diagrams represent the shape of the curve for different values of "a" for complete and censored data.

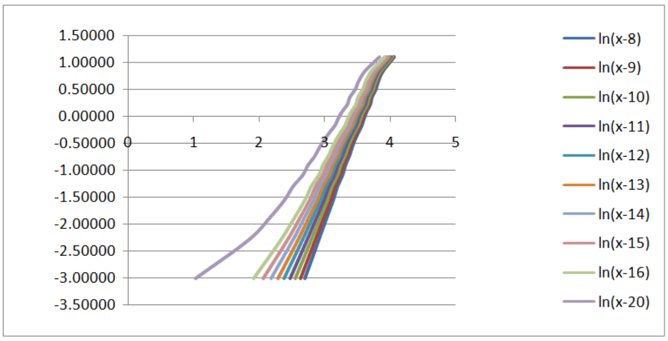

Figure 1: Graphical estimation of 'a' by trial and error for complete data.

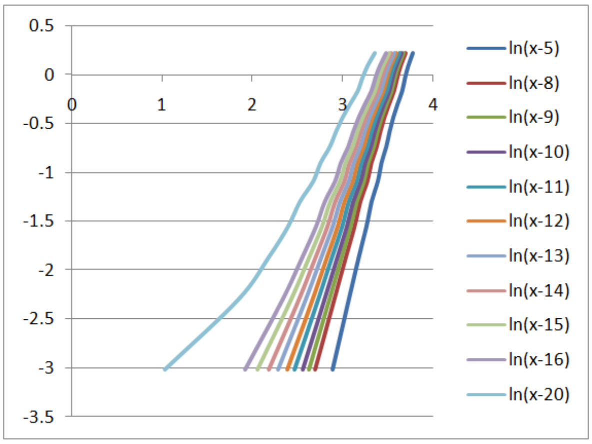

Figure 2: Graphical estimation of 'a' by trial and error for first 15 samples out of 20 samples

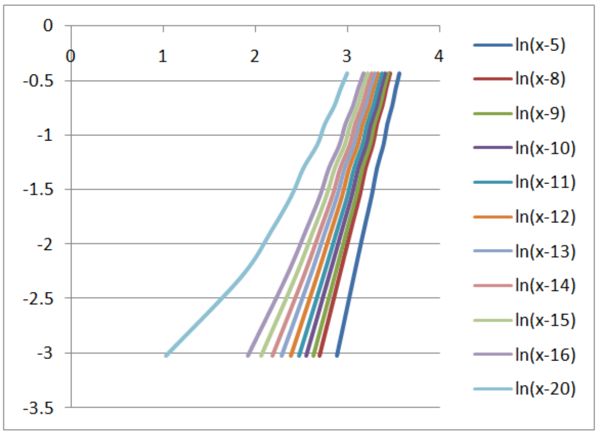

Figure 3: Graphical estimation of 'a' by trial and error for first 10 samples out of 20 samples

The curves display $\hat{u}_{i:n}$ as a function of $x^*_{i:n} = ln(x_{i:n} - \hat{a})$ and are obviously concave; i.e., 'a' must be greater than zero. Taking $\hat{a} = 5, 8,9,10,11,12,13,14,15,16$ and 20 moves the curves to the left, reduces the concavity and finally leads to convexity when $\hat{a} = 20$. To decide which of the Eleven curves is closest to linearity, select the one with residual sum of squares value nearer to zero using trial and error. From the table 2 and Fig.1, it is understood that the estimated value of 'a' is 13 for complete data and 12 and 9 for partial data containing first 15 and 10 samples respectively.

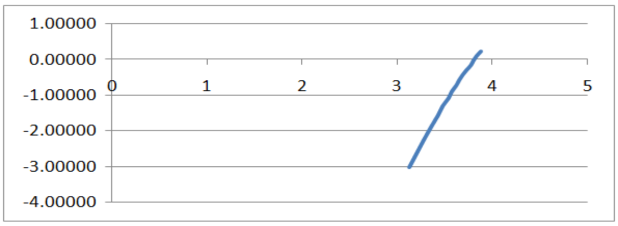

Figure 4: Graph of $\hat{u}_{i:n}$ versus $x^*_{i:n} = ln(x_{i:n})$

In Figure 2, the line cuts $x^*_{i:n}$ axis at 3.81. Therefore $x_0 = 45.1$

Table 3: Estimates of 'a', 'b' and 'c' for data set given in Table 1

	Model 1			Model 2			MLE		
m	20	15	10	20	15	10	20	15	10
a	13	12	9	13	12	9			
b	32.30473	33.15941	35.41811	32.32	33.16613	35.42264	30.96	30.69	29.17
c	2.5624	2.715554	3.208661	2.55	2.713966	3.207618	2.764	2.8446	3.1253

Jiang/Murthy	
alpha	0.081812
beta	0.001708
a	20.13124
b	24.96876
c	2.042743

From this, it is observed that the values of the estimators are very close in all the methods but it is different in Jiang/Murthy approach. When censoring involves the data it affects the parameter estimators also.

Table 4: Reliability function Estimates are listed Using MLE Method.

i	t	R(t),m=20	R(t),m=15	R(t),m=10
1	22.8	0.969571	0.97178	0.970963
2	26.3	0.924993	0.928481	0.919485
3	28.8	0.878073	0.881961	0.860974
4	30.9	0.82892	0.832598	0.797203
5	32.4	0.788718	0.791884	0.744034
6	34.4	0.729294	0.731293	0.664828
7	35.6	0.690938	0.691983	0.613792
8	37.3	0.633925	0.633345	0.538728
9	38.5	0.592326	0.590445	0.484957
10	**39.9**	**0.543002**	**0.539502**	**0.422738**
11	41.6	0.4829	0.477385	
12	43.5	0.416857	0.409179	
13	44.7	0.376454	0.367539	
14	46.2	0.328029	0.317779	
15	48.4	0.262356	0.250697	
16	49.5	0.232311		
17	52	0.171741		
18	53.8	0.135041		
19	57.3	0.079864		
20	66.4	0.013813		

From the above table, it is understood that when censoring is involved in the data it affects the reliability function estimates also. The probability of a subject surviving longer than time 39.9 is 0.54, 0.54 and 0.42 for number of samples 20, 15 and 10 among 20 samples respectively. Estimates are computed using R Software and with user defined functions and Microsoft Excel tools.

V. Conclusion

The performances of Graphical and Maximum Likelihood methods in the estimation of the parameters of the Weibull distribution were compared in this study. Many authors assumed that one of the parameters in Weibull three-parameter distributions is known. For fixing the values of Location parameter estimator, when using trial and error, it differs for complete and censored data. Jiang/Murthy method of estimation is used only for complete data. When tired for censored observations, it gives different results. When censoring is involved in the data structure, it affects the value of the parameter estimators. The Estimate of the Reliability function differ particular time because of censoring.

References

[1]. Weibull W, A statistical distribution functions with wide applicability, *J. Appl. Mech. 18, 1951, 292-297.*

[2]. Lawless, J.F. (1982): *Statistical Models and Methods for Lifetime Data*. Wiley, New York.

[3]. Rinne, H. (2009), *The Weibull Distribution The Hand book*. Chapman & Hall. New York.

[4]. Jiang, R. / Murthy, D.N.P., Comment on a general linear–regression analysis applied to the three–parameter Weibull distribution; *IEEE Transactions on Reliability 46, 1997,389–393.*

[5]. Bartkute. V and Sakalauskas. L, The method of three-parameter Weibull distribution estimation, *Acta ET Commentationes Universitatis TArtuensis De Mathematica,12, 2008, 65-78.*

[6]. Bain, L. J. and Engelhardt, M. E.(1991). Statistical *Analysis of Reliability and Life Testing Models*, 2nd Edition Marcel Dekker, INC New York.

[7]. Bain, L. J. and Antle, C. E. Estimation of parameters in the Weibull distribution. *Technometrics 9, 1967, 621-627.*

[8]. Davis, D. J. An analysis of some failure data. *J. Amer.Statist. Ass. 47, 1952, 113-150.*