Comparing Different Estimators for Two Parameters Weibull Distribution under censored sample Type I

Dr. Shurooq A.K.AL-Sultany[1] and Sahar Ahmed Mohammed [2]
[1]Dept. of Mathematics/ College of Science/ Mustansiriyah University/ Iraq
E-mail:Shuroq80@uomustansiriyah.edu.iq
[2] Lecturer in Dept. of Mathematics/College of Science/ Mustansiriyah University/ Iraq.

Abstract

This paper discuss the failure censored sample taken from Weibull failure model with two parameters (p: shape and: scale parameter).We estimate the two parameters(P, θ) by maximum likelihood and white estimators , Assuming that we have a failure censored sample with the failure times of (r)items that failed. This type of censored is necessary in life testing application and in statistical quality control test.

Keyword: censored sample, Maximum likelihood, and white estimator.

الملخص

هذا البحث عينة مراقبة الفشل الماخوذة من نموذج فشل ويبل مع المعلمتين(الشكل والقياس). نقدر المعلمتين حسب مقدر الامكان ومقدر وايت بافتراض انه لدينا عينة مراقبة الفشل لازمنة فشل. هذا النوع من الرقابة ضروري في تطبيق اختبار الحياة وفي اختبار مراقبة الجودة الإحصائية.

1-Introduction

The estimation of parameters may done using complete sample available, where the failure times of all units (n) are recorded, but sometimes the life-testing experiments may be destructive, in that the items are destroyed at the end of experiment and then it cannot be used again. In this cases the experiments are executed on censored data, when we have $(r < n)$ units failure, and the sampling is done with high cost items, especially in medical experiments, we consider the failure-censored case the data here are life times of the r items failed $(X_{(1)} < X_{(2)} < \cdots < X_{(r)})$ and $(n-r)$ survived beyond $X_{(r)}$.

Here the times-censored data contains the life times of items that failure before the time t_0, ie $(X_{(1)} < X_{(2)} < \cdots < X_{(m)})$ assumed that m items are failed before (t_0) and the $(n-m)$ items have survived beyond t_0.

In 1996 (Fritz)[١] work or plotting weibull and estimating its parameters under censoring type

Waloddi weibull [٢] gives many applications for this distribution in fiber strength of Indian cotton and length of cyrtoideae.

In this type of censoring the time of censoring is fixed while (m) which represent the number of observation that failure before t_0, and it is random variable.

The censoring may be done in time or in the number of observation or in combination of two.

Here our research consider the first case of failure-censored sample, and we work on comparing two parameters (P, θ) of weibull distribution, using maximum likelihood method and white estimator, and also estimating reliability function.

2-Theoretical Aspects

The p.d.f of two parameters Weibull

X~ Weibull(P, θ) is $f(x\,;\,p,\theta) = \frac{P}{\theta} x^{p-1}\ e^{\frac{-x^p}{\theta}} \quad ;\ x > 0\ , \mathrm{P}, \theta > 0$ (1)

Where, p is shape parameter and θ is scale parameter

, The cumulative distribution function is $F(x;P,\theta) = (1 - e^{\frac{-x^p}{\theta}})$ (2)

And Reliability function is

R(x)= $e^{\frac{-x^p}{\theta}}$ (3)

While the formula for r^th moments about origin is [٣]

$$\mu' r = E(x^r) = \theta^{\frac{r}{p}} \Gamma \left(1 + \frac{r}{p}\right) \quad (4)$$

This lead to the Mean

$$\mu' = E(x) = \theta^{\frac{1}{p}} \Gamma \left(1 + \frac{1}{p}\right)$$

And

$$\sigma_x^2 = \theta^{\frac{2}{p}} [\Gamma \left(1 + \frac{2}{p}\right) - (\Gamma \left(1 + \frac{1}{p}\right))^2]$$

Also the hazard rate

$$\mu(t) = \frac{p}{\theta} t^{p-1} \quad (5)$$

When p >1, $\mu(t)$ is increasing function

And when p<1, $\mu(t)$ is decreasing

For p=1, $\mu(t)$ is$(\frac{1}{\theta})$ constant

This research aims to introduce the estimation of two parameters(p,θ) of failure censored sample, when the experiment of the test is terminated after (r) failures have been indicated earlier and the results of test for complete sample can be obtained from setting $r = n$, here the failure times observation$(X_{(1)} < X_{(2)} < \cdots < X_{(r)})$ are increasing order, then the units of sample are

$(r, n-r)$, the estimation is done using these set sol observation .the censored is necessary when the observations for experimentation are more valuable, and expensive, or the test must be terminated when we have r observations obtained.

So here we work on estimating (the two parameters p, θ) of Weibull distribution under censoring data. Weibull failure model is used in studies of strength of material, and time to failure module, also for describing the "wear out" (i.e. the fatigue failures).

3-Estimation of parameters

The estimation of two parameters (P, θ) is done through failure censored sample case; in this case the experiment of testing is terminated after r failures are obtained. To estimate the two parameters P, θ .

3.1-Maximum likelihood estimators

let us assume ,that we have censored sample with the failure times of r items

that have been failed i.e. let $X_{(1)} < X_{(2)} < \cdots < X_{(r)}$ be rth failure times which are arranged in increasing order of magnitude, here the likelihood function then

$$L^* = \text{k} \left(\frac{p}{\theta}\right)^r \prod_{i=1}^{r} x_i{}^{p-1}\, e^{-\frac{\sum_{i=1}^{r} x_i{}^p}{\theta}} \left[\int_{x_r}^{\infty} \frac{p}{\theta} x^{p-1} e^{-\left(\frac{x}{\theta}\right)} dx\right]^{n-r} \qquad (6)$$

Taking Logarithm of L^*[٤], Then

$$\mathbf{logl = logk + r[logp - log\theta] + (p-1)\sum_{i=1}^{r} logx_i - \frac{\sum_{i=1}^{r} x_i{}^p + (n-r)x_r{}^p}{\theta}}$$

$$\frac{\partial \boldsymbol{logl}}{\partial \boldsymbol{p}} = \frac{\boldsymbol{r}}{\boldsymbol{p}} + \sum_{i=1}^{r} \boldsymbol{logx_i} - \frac{\sum_{i=1}^{r} x_i{}^{(p)} logx_i + (n-r)x_r{}^p logx_r}{\theta} \qquad \textbf{(7)}$$

$$\frac{\partial \boldsymbol{logl}}{\partial \boldsymbol{\theta}} = \frac{-\boldsymbol{r}}{\boldsymbol{\theta}} + \frac{\sum_{i=1}^{r} x_i{}^p + (n-r)x_r{}^p}{\theta^2} \qquad \textbf{(8)}$$

From solving these equation (7) and (8)

We can obtain $p^{\wedge}{}_{MLE}$ and $\theta^{\wedge}{}_{MLE}$

Where

$$\boldsymbol{\theta}^{\wedge}{}_{\mathbf{MLE}} = \frac{\sum_{i=1}^{r} \mathbf{x_i}^{\mathbf{p}^{\wedge}} + (\mathbf{n}-\mathbf{r})\mathbf{x_r}^{\mathbf{p}^{\wedge}}}{\mathbf{r}} \qquad (9)$$

And from solving equation (7)

$$\frac{\mathbf{r}}{\mathbf{p}^{\wedge}} + \sum_{i=1}^{r} \mathbf{log x_i} - \frac{\sum_{i=1}^{r} \mathbf{x_i}^{\mathbf{p}^{\wedge}} \mathbf{log x_i} + (\mathbf{n}-\mathbf{r})\mathbf{x_r}^{\mathbf{p}^{\wedge}} \mathbf{log\, x_r}}{\boldsymbol{\theta}^{\wedge}}$$

This gives MLE for $p^{\wedge}$ (using newton-Raphson iteration method or any other suitable method).

Where the moment estimates are used as starting values for maximum likelihood yielding estimates p.

2)**white estimators**

This method depends on estimating (p,θ) on regression approach [٥]

This method is proposed by white (1969), it can be used for complete or censored data.

Since $\mathbf{F}(\boldsymbol{x_i}) = (\mathbf{1} - \boldsymbol{e}^{-\frac{-x_i^{p}}{\theta}})$

$$\boldsymbol{e}^{-\frac{x_i^{p}}{\theta}} = \mathbf{1} - \boldsymbol{F}(\boldsymbol{x_i})$$

$$\boldsymbol{e}^{-\frac{x_i^{p}}{\theta}} = \boldsymbol{u_i} \qquad \mathbf{0} \le \boldsymbol{u_i} \le \mathbf{1}$$

$$\mathbf{log}\, \boldsymbol{u_i} = -\frac{\boldsymbol{x_i}^{\boldsymbol{p}}}{\boldsymbol{\theta}}$$

$$\mathbf{log}[-\mathbf{log}\, \boldsymbol{u_i}] = \boldsymbol{plogx_i} - \boldsymbol{log\theta}$$

$$\mathbf{log}\, \boldsymbol{t_i} = \boldsymbol{plogx_i} - \boldsymbol{log\theta}$$

$$\boldsymbol{y_i} = \boldsymbol{plogx_i} + \mathbf{log}\left(\frac{\mathbf{1}}{\boldsymbol{\theta}}\right)$$

$$\boldsymbol{y_i} = \mathbf{log}\left(\frac{\mathbf{1}}{\boldsymbol{\theta}}\right) + \boldsymbol{plogx_i}$$

$$\boldsymbol{y_i = \propto + p \log x_i} \qquad (11)$$

$$\propto = \log\left(\frac{1}{\theta}\right) \Rightarrow \propto^{\wedge} = -\log(\theta^{\wedge})$$

$$(\theta^{\wedge} = e^{-\propto^{\wedge}}) \qquad (12)$$

First we obtain the values of

$$t_i = (-\log u_i) \; were \;\; 0 \le u_i \le 1$$

Then $y_i = (\log t_i)$

Regression model is

$$y_i = p\log x_i + \log\left(\frac{1}{\theta}\right)$$

$$y_i = p\log x_i - \log\theta$$

$$y_i = p\log x_i + \propto \qquad (13)$$

Use least square Regression Analysis we have

$$y_i = pz_i + \propto$$

$$p^{\wedge} = \frac{\sum y_i \; z_i - n\,\overline{y}\overline{z}}{\sum {z_i}^2 - n\,\overline{z}^2} \qquad (14)$$

$$\propto^{\wedge} = \bar{y} - \hat{p}\bar{z} \qquad (15)$$

$$y_i = \log(t_i), \qquad t_i = -\log u_i \qquad\qquad 0 \le u_i \le 1$$

$$x_i = \{-\theta \log u_i\}^{\frac{1}{p}}$$

From the data [6],

1.84, 0.05, 0.49, 0.07, 0.15, 2.60, 1.78, 0.57, 4.50, 0.05,

0.87, 0.72, 2.18, 2.38, 0.31, 1.01, 0.88, 1.56, 1.85, 2.11,

0.48, 1.71, 1.60, 0.27, 0.19, 1.03, 4.82, 0.62, 1.32, 0.38.

we had moment estimator for

$\mathbf{p}^{\wedge}_{\mathbf{mom}} = \mathbf{2.164}$

$\boldsymbol{\theta}^{\wedge}_{\mathbf{mom}} = \mathbf{3416}$

For complete sample

But for censored data at r=15 and for p=2

$$\boldsymbol{\theta}^{\wedge} = \frac{\sum_{i=1}^{r} x_i{}^{p} + (n-r)x_r{}^{p}}{r}$$

$\boldsymbol{\theta}^{\wedge}_{I} = \mathbf{1.9165}$ **with var ($\boldsymbol{\theta}^{\wedge}$)=0.0882**

And also for r=10 n=30 p=2

$\boldsymbol{\theta}^{\wedge}_{II} = \mathbf{1.948}$, $\boldsymbol{Var}(\boldsymbol{\theta}^{\wedge}_{II})$ ***=0.0887***

Since the estimators of $\theta^{\wedge}$ in case of two censured when (r=15) (r=10), have approximate variance

$\mathbf{Var}(\boldsymbol{\theta}^{\wedge}_{\mathbf{1}}) = \mathbf{0.0882}$ [**when r = 15**]

And $\mathbf{Var}(\boldsymbol{\theta}^{\wedge}_{\mathbf{1}}) = \mathbf{0.0887}$ [**where r = 10**]

We can depend on truncated (r=10) rather than (r=15), this is necessary for destructive test,

Work on finishing the unit.

While the results of second method(white method) According to the observation(x_i)

And for given values of ($\propto, \beta$)

We generate values of (y_i)

$\mathbf{y_i} = \propto + \boldsymbol{\beta}\mathbf{log}\mathbf{x_i}$

Let $\mathbf{z_i} = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{log}\mathbf{x_i}$

Then we estimate β and $\propto$ we can obtain different sets of estimators according to given (initial Sets of$\propto, \beta$)

The observation (x_i) here represented the time to failure, while the variable (y_i)

(dependent variable($y_i = logt_i$)

Where $t_i = (-logu_i)$ $(0 \leq u_i \leq 1)$

Then according to given values of (p, θ ,n) ,the values of (x_i) can also be generated

From $x_i = (-\theta \log u_i)^{\frac{1}{p}}$ (this is necessary when the data are not available)

So x_i is generated and $y_i = (\log t_i)$

$t_i = (-\log u_i)$ $0 \leq u_i \leq 1$

The fitted resulted model is

$$\boldsymbol{y^{\wedge}_i} = \propto^{\wedge} + \boldsymbol{\beta}^{\wedge} \boldsymbol{log x_i}$$

$$\propto^{\wedge} = \boldsymbol{log}\left(\frac{\mathbf{1}}{\boldsymbol{\theta}^{\wedge}}\right) \Rightarrow \propto^{\wedge} = -\boldsymbol{log\theta}^{\wedge}$$

$$-\propto^{\wedge} = \boldsymbol{log\theta}^{\wedge}$$

$$\boldsymbol{\theta}^{\wedge} = \boldsymbol{e}^{-\propto^{\wedge}} \ \ \textbf{\textit{while}}\ \boldsymbol{p}^{\wedge} = \boldsymbol{\beta}^{\wedge}$$

According to the data we give initial values and obtain estimators

Application

Table (1):Table represents the estimates of the parameters θ and P obtained through different methods, also mean square error and moment method for complete sample

True values	$\theta = 4$	MSE($\theta^\wedge$)	P=2	MSE($p^\wedge$)
Estimators				
MLE	**3.667**	**0.4553**	**1.875**	**0.3206**
MOM	**3.358**	**0.4032**	**2.0035**	**0.4511**
White	**3.949**	**0.2361**	**2.5461**	**0.5678**
	$\theta = 5.5$	MSE($\theta^\wedge$)	P=3	MSE($p^\wedge$)
MLE	**4.326**	**0.3532**	**2.885**	**0.3367**
MOM	**4.557**	**0.3367**	**2.066**	**0.2412**
White	**3.889**	**0.2461**	**2.705**	**0.4467**

The best estimator for shape parameter p is ($p^\wedge{}_{MLE}$) and for scale parameter $\theta^\wedge$ is White

While for θ the best one is $\theta^\wedge{}_{white}$ (which depend) on Regression method So according to these estimators of p and θ we can find the estimators of $R^\wedge(t_i)$

for three method

Table (2): Different Reliability Methods when θ=4 and P =2

t_i	$R_{real}(t_i)$	$\hat{R}_{MLE}(t_i)$	$\hat{R}_{MOM}(t_i)$	$\hat{R}_{White}(t_i)$
1	0.6065	0.5997	0.5507	0.5248
2	0.3679	0.3596	0.3032	0.2754
3	0.2231	0.2157	0.1670	0.1445
4	0.1353	0.1293	0.0919	0.0759
5	0.0821	0.0776	0.0506	0.0398

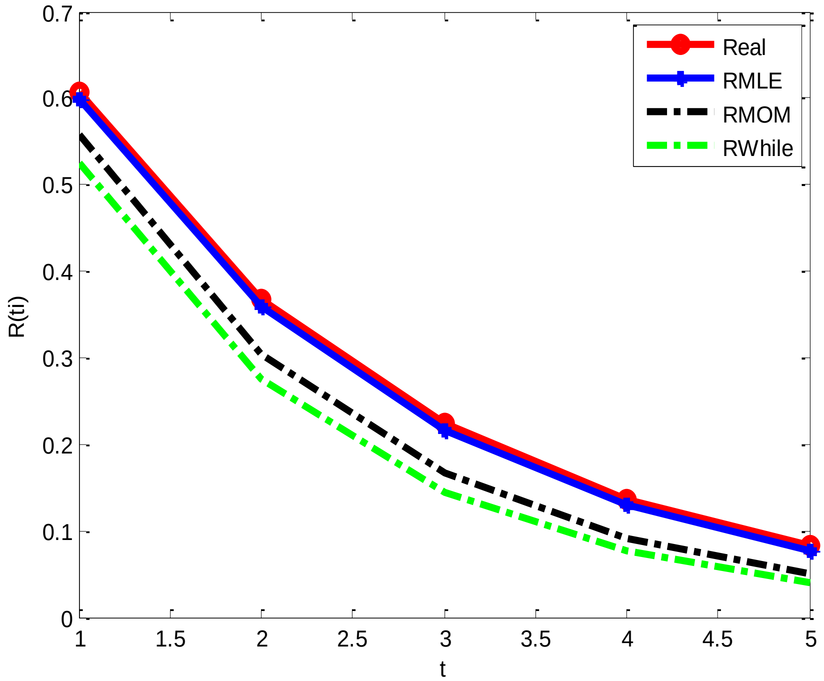

Figure 1: Plot of Different Reliability Methods

Table (3): Different Reliability Methods when θ=5.5 and P =3

t_i	$R_{real}(t_i)$	$\hat{R}_{MLE}(t_i)$	$\hat{R}_{MOM}(t_i)$	$\hat{R}_{White}(t_i)$
1	0.1599	0.5133	0.6355	0. 4988
2	0.0256	0.2635	0.4038	0. 2488
3	0.0041	0.1352	0.2566	0.1241
4	0.0001	0.0694	0.1631	0. 0619
5	0.0001	0.0356	0.1036	0.0309

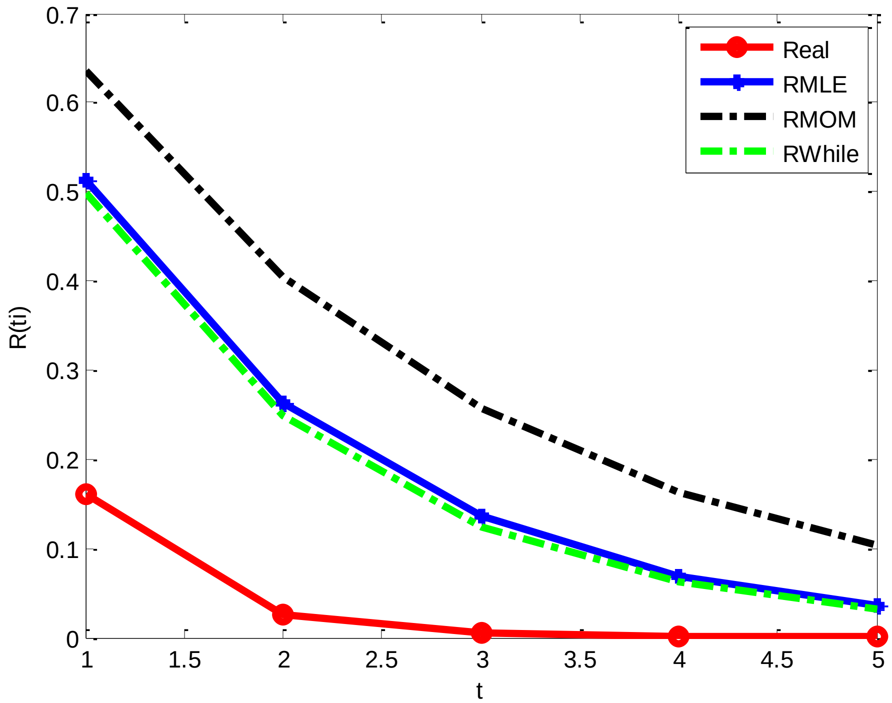

Figure 2: Plot of Different Reliability Methods

Conclusions

It's clear that $\hat{R}_{\text{MLE}}$ is greater than $\hat{R}_{\text{MOM}}$ and $\hat{R}_{\text{White}}$ for different set of estimated parameters and values of t_i so when θ=4 and P =2, $R^{\wedge}_{\text{MLE}}$ is better than $\hat{R}_{\text{MOM}}$ and $\hat{R}_{\text{White}}$.

The truncated can be done at $r = 10$ rather than $r = 15$ in case of destructive test, when the test work on finishing the units

Reference

[١] Fritz Scholz, 1996, "Plotting on Weibull Paper".
[٢] Weibull, W. (1951) A Statistical Distribution Function of Wide Applicability. Journal of Applied Mechanics, 18, 293-297.
[٣] Ivana Pobočíková and Zuzana Sedliačková, 2014, "Comparison of Four Methods for Estimating the Weibull Distribution Parameters "Applied Mathematical Sciences, Vol. 8, no. 83, 4137 – 4149.
[4] Cohen, A.C. (1965) Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples. Technometrics, 7, 579-88.
.[٥] K. P. SOMAN and K. B. MISRA ,(1992.),"A least square estimation of three parameters of a Weibull distribution", Microelectron. Reliab.,Vol. 32, No. 3, pp. 303-30.
[6] Albert K. Cohen, 1965, "THE SOCIOLOGY OF THE DEVIANT ACT: ANOMIE THEORY AND BEYOND", American Sociological Association, Vol. 30, No. 1.