International Journal of Advanced Engineering Research and Science (IJAERS)
Peer-Reviewed Journal
ISSN: 2349-6495(P) | 2456-1908(O)
Vol-8, Issue-9; Sep, 2021
Journal Home Page Available: https://ijaers.com/
Article DOI: https://dx.doi.org/10.22161/ijaers.89.21

Parameter estimation of the Weibull Distribution; Comparison of the Least-Squares Method and the Maximum Likelihood estimation

Edward Appau Nketiah, Li Chenlong*, Jing Yingchuan, Barbara Dwumah

Department of Statistics, Taiyuan University of Technology, China

Received: 12 Aug 2021,

Received in revised form: 16 Sep 2021,

Accepted: 21 Sep 2021,

Available online: 30 Sep 2021

***Keywords*— Asymptotic efficiency, Least-squares, Maximum likelihood, Parameter estimation, Weibull distribution.**

***Abstract*— *Weibull distribution is a very useful distribution in survival analysis, lifetime analysis, and reliability analysis. Several methods have been proposed to estimate the parameters of different distributions such as the method of moment, maximum likelihood, etc. In this paper, we analyze the 2-parameter Weibull distribution by simulating data on failure times of a product using the Monte Carlo approach and estimating the parameters of the distribution using the maximum likelihood estimation (MLE) and the least-squares method (LS). These methods were also investigated through applications in reliability analysis. The two approaches of estimating the parameters were compared, and the MLE obtained better performance than the least-squares method when the results for the parameters were assessed using the goodness of fit measures. Also, we obtained the asymptotic distribution of the MLE which was asymptotically efficient as the sample size increases. The inverse of the Fisher's information matrix which is the asymptotic variance-covariance matrix was also obtained.*

I. INTRODUCTION

The Weibull distribution has gained much weight in the real world and is mostly used in reliability and lifetime analysis or survival analysis. The 2-parameter Weibull distribution has the shape parameter (β) and the scale parameter (η or α) whose values affect the characteristic life of the distribution, the reliability function, and the failure rate. Most distributions such as normal distribution, gamma distribution, inverse gamma, and some other common distributions have two parameters which makes the 2-parameter Weibull distribution of immense interest. According to Sun(1997), the Weibull distribution does not belong to the 2-parameter exponential family neither are the two parameters orthogonal. In estimating the parameters of the Weibull distribution, the method of moment achieved the best result when it was compared with other estimators such as the MLE and the LS using the mean square error (MSE) and the total deviation (TD) for the 3-parameter and 2-parameter Weibull distribution (Razali *et.al*, 2009). Pobocikova *et.al* (2017) investigated the performance of six estimators using the Monte Carlo simulation for wind speed, the result of the study indicates that the maximum likelihood estimation method performs better in estimating the parameters of the 2-parameter Weibull distribution for wind speed when the bias and root mean square error (RMSE) were compared. However, according to Chu & Ke (2012), the least-squares method outwits the maximum likelihood approach when the sample size is small.

This paper aims to analyze the 2-parameter Weibull distribution by applying the Monte Carlo simulation to generate artificial data of failure times of a product through which we obtained the parameter estimates of the shape and scale parameters of the Weibull distribution by comparing the maximum likelihood estimation (MLE) method and that of the linear regression least-squares method. The two approaches of estimation were also applied to two different sets of data in reliability analysis.

The research also assesses the asymptotic efficiency of the maximum likelihood estimator of the shape and scale parameters of the 2-parameter Weibull distribution.

II. MAXIMUM LIKELIHOOD ESTIMATOR

The density function of the 2-parameter Weibull distribution is given as

$$f(x)=\frac{\beta x^{\beta-1}}{\eta^{\beta}}\exp\left\{-\left(\frac{x}{\eta}\right)^{\beta}\right\} \quad x>0\ ,\ \beta>0\ ,\ \eta>0 \tag{1}$$

The cumulative density function is given by

$$F(\mathrm{x})=1-\exp\left\{-\left(\frac{x}{\eta}\right)^{\beta}\right\} \tag{2}$$

where β is the shape parameter and η is the scale parameter. The reliability function is given by $R(x)=\exp\left\{-\left(\frac{x}{\eta}\right)^{\beta}\right\}$, the cumulative density function can be expressed in terms of the reliability function as $F(\mathrm{x})=1-R(x)$.

Maximum likelihood estimation is a method used to estimate the parameter of a probability distribution. This is obtained by maximizing the likelihood function of the distribution. The maximum likelihood estimator of the Weibull distribution is given as follows:

$$\begin{aligned}\prod_{i=1}^{n} f(x_i) &= \prod_{i=1}^{n}\frac{\beta x_i^{\beta-1}}{\eta^{\beta}}\exp\left\{-\left(\frac{x_i}{\eta}\right)^{\beta}\right\}\\ &=\left(\frac{\beta}{\eta^{\beta}}\right)^{n}\prod_{i=1}^{n} x_i^{\beta-1}\exp\left\{-\left(\frac{x_i}{\eta}\right)^{\beta}\right\}\end{aligned} \tag{3}$$

Taking the log of the likelihood in (3) we obtain the equation

$$L\left(f(\mathrm{x})\right)=n\log\beta-n\beta\log\eta+(\beta-1)\sum_{i=1}^{n}\log(x_i)-\sum_{i=1}^{n}\left(\frac{x_i}{\eta}\right)^{\beta} \tag{4}$$

The scale parameter can easily be obtained with an analytical method; the shape parameter β is very difficult to be obtained by the analytical method, however, it can be obtained by using the Newton-Raphson method or the least square method (Guure*et al*., 2012; Al-Fawzan, 2000). We obtain the scale parameter by differentiating (4) and solving the equation

$$\frac{\partial L\left(f(\mathrm{x})\right)}{\partial\eta}=0$$

$$\frac{\partial L\left(f(\mathrm{x})\right)}{\partial\eta}=-\frac{n\beta}{\eta}+\frac{\beta}{\eta}\sum_{i=1}^{n}\left(\frac{x_i}{\eta}\right)^{\beta}=0 \tag{5}$$

$$\eta=\left(\frac{\sum_{i=1}^{n}x_i^{\beta}}{n}\right)^{\frac{1}{\beta}}$$

Asymptotic efficiency

The asymptotic efficiency of the 2-parameter Weibull distribution can be obtained as follows, We find the partial derivatives of the parameters by differentiating the log of the Weibull distribution.

$$\omega_{11} = \frac{\partial \log f(x)}{\partial \eta^2} = \frac{\beta}{\eta^2} - \frac{\beta(\beta+1)}{\eta^2}\left(\frac{x}{\eta}\right)^{\beta}$$

$$\omega_{22} = \frac{\partial \log f(x)}{\partial \beta^2} = -\frac{1}{\beta^2} - \left(\frac{x}{\eta}\right)^{\beta}\left(\log\left(\frac{x}{\eta}\right)\right)^2$$

$$\omega_{12} = \frac{\partial \log f(x)}{\partial \eta\, \partial \beta} = -\frac{1}{\eta} + \frac{1}{\eta}\left(\frac{x}{\eta}\right)^{\beta} + \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta}\log\left(\frac{x}{\eta}\right)$$

$$\omega_{21} = \frac{\partial \log f(x)}{\partial \beta\, \partial \eta} = -\frac{1}{\eta} + \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta}\log\left(\frac{x}{\eta}\right) + \frac{1}{\eta}\left(\frac{x}{\eta}\right)^{\beta}$$

We obtain the determinant using the Hessian matrix of the Weibull distribution and ignoring all negative terms.

$$H(\theta) = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix}$$

where $\theta = (\beta, \eta)$

$\left|H(\theta)\right| = \frac{\beta+1}{\beta\, \eta^2}\left(\frac{x}{\eta}\right)^{\beta} + \frac{\beta(\beta+1)}{\eta^2}\left(\frac{x}{\eta}\right)^{2\beta}\left(\log\left(\frac{x}{\eta}\right)\right)^2$, considering $\log x < x+1$. We rewrite the determinant as

$$\left|H(\theta)\right| < \frac{\beta+1}{\beta\, \eta^2}\left(\frac{x}{\eta}\right)^{\beta} + \frac{\beta(\beta+1)}{\eta^2}\left(\frac{x}{\eta}\right)^{2\beta}\left(\frac{x}{\eta}+1\right)^2 \tag{6}$$

Considering $\frac{\eta_0}{2} < \eta < \frac{3\eta_0}{2}$ and $\frac{\beta_0}{2} < \beta < \frac{3\beta_0}{2}$ and define

$$\frac{4(\beta+1)}{\beta\eta_0^2}\left(\frac{2x}{\eta_0}\right)^{\beta} + \frac{4(\beta(\beta+1))}{\eta_0^2}\left(\frac{2x}{\eta_0}\right)^{2\beta}\left(\frac{2x}{\eta_0}+1\right)^2 = h_{\beta,\,\eta_0}(x) \tag{7}$$

We obtain the upper bound of $h_{\beta,\,\eta_0}(x)$ from (7), for any $x > 0$, $\left(\frac{2x}{\eta_0}\right)^{\beta} < \left(\frac{2x}{\eta_0}+1\right)^{2\beta_0}$

$$h_{\beta,\,\eta_0}(x) < \frac{4(2\beta_0+1)}{\frac{\beta_0}{2}\eta_0^2}\left(\frac{2x}{\eta_0}+1\right)^{2\beta_0} + \frac{4(2\beta_0(2\beta_0+1))}{\eta_0^2}\left(\frac{2x}{\eta_0}+1\right)^{2\beta_0}\left(\frac{2x}{\eta_0}+1\right)^2 \tag{8}$$

Then
$$E\left(h_{\beta,\,\eta_0}(x)\right) < E\left[\frac{4(2\beta_0+1)}{\frac{\beta_0}{2}\eta_0^2}\left(\frac{2x}{\eta_0}+1\right)^{2\beta_0} + \frac{4(2\beta_0(2\beta_0+1))}{\eta_0^2}\left(\frac{2x}{\eta_0}+1\right)^{2\beta_0}\left(\frac{2x}{\eta_0}+1\right)^2\right] < \infty$$

The kth moment is given as follows:

$$E\left(X^k\right) = \eta^k\Gamma\left(1+\frac{k}{\beta}\right) < \infty$$

Therefore, based on Bahadur (1964) we have

$$\sqrt{n}\left(\hat{\theta}_n - \theta_n\right) \xrightarrow{d} N_k\left(0, \left[I_n(\theta)\right]^{-1}\right)$$

Hence we obtain the Fisher's information matrix $I_n(\theta)$ and its inverse, the information matrix is given as

$$I(\theta) = -E\begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix}$$

For us to obtain the Fisher information in a simplified form, we make the same assumption made by Sun (1997) that for $i \geq 1$, we define $\gamma_i = \int_0^\infty [\log(y)]^i \exp(-y)\,dy$, and we let $Y = \left(\frac{x}{\eta}\right)^\beta$ be an exponential random variable with mean 1 whose density function is expressed as $f(y) = \exp(-y), \quad y \geq 0$

$$\beta E\left\{\left(\frac{x}{\eta}\right)^\beta \log\left(\frac{x}{\eta}\right)\right\} = \int_0^\infty y \log(y) f(y)\,dy$$

$$\int_0^\infty \left[1 + \log(y)\right] \exp(-y)\,dy = 1 + \gamma_1$$

$$\beta^2 E\left\{\left(\frac{x}{\eta}\right)^\beta \left[\log\left(\frac{x}{\eta}\right)\right]^2\right\} = \int_0^\infty y \log(y)^2 f(y)\,dy$$

$$\int_0^\infty \left[\log(y)^2 + 2\log y\right] \exp(-y)\,dy = \gamma_2 + 2\gamma_1$$

The Fisher's information matrix can be expressed as

$$I(\theta) = \begin{pmatrix} \dfrac{\beta^2}{\eta^2} & -\dfrac{1+\gamma_1}{\eta} \\ -\dfrac{1+\gamma_1}{\eta} & \dfrac{1+\gamma_2+2\gamma_1}{\beta^2} \end{pmatrix}$$

The determinant of the information matrix $\left|I(\theta)\right| = \dfrac{\gamma_2 - \gamma_1^2}{\eta^2}$, where $-\gamma_1$ is Euler's constant and $\gamma_2 - \gamma_1^2$ is the variance of $\log(\phi)$ where ϕ is an exponential random variable with a mean of 1.

$$\left[I(\theta)\right]^{-1} = \frac{\eta^2}{\gamma_2 - \gamma_1^2} \begin{pmatrix} \dfrac{1+\gamma_2+2\gamma_1}{\beta^2} & \dfrac{1+\gamma_1}{\eta} \\ \dfrac{1+\gamma_1}{\eta} & \dfrac{\beta^2}{\eta^2} \end{pmatrix}$$

III. LEAST-SQUARES ESTIMATION

The least-squares estimation of the 2-parameter Weibull distribution is also important and can be used to estimate the parameters of the Weibull distribution inasmuch as the maximum likelihood is used. In estimating the parameter of the Weibull distribution, we will consider the cumulative density function and obtain a regression equation from which we can deduce the equations for estimating the parameters of the 2-parameter Weibull distribution. The cumulative density function of the Weibull distribution in (2) becomes:

$$\exp\left\{-\left(\frac{x}{\eta}\right)^{\beta}\right\}=1-F(\mathrm{x})$$

Taking the log of the reliability function, we obtain

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

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

$$\ln(x)=\frac{1}{\beta}\ln\left\{-\ln\left[1-F(\mathrm{x})\right]\right\}+\ln(\eta) \tag{9}$$

From (9), let $Y_i=\ln(x_i)$, $X=\ln\left\{-\ln\left[1-F(\mathrm{x})\right]\right\}$, $a=\frac{1}{\beta}$ and $b=\ln(\eta)$, we write (9) as a linear function of the form $Y_i=aX+b$. By substituting Y_i and X into the least-squares formula, we will obtain the estimates for a and b through which we estimate the parameters β and η.

$$a=\frac{n\sum_{i=1}^{n}\ln\left\{-\ln\left[1-\hat{F}(\mathrm{x}_i)\right]\right\}\ln x_i-\sum_{i=1}^{n}\ln\left\{-\ln\left[1-\hat{F}(\mathrm{x}_i)\right]\right\}\sum_{i=1}^{n}\ln x_i}{n\sum_{i=1}^{n}(\ln x_i)^2-\left(\sum_{i=1}^{n}\ln x_i\right)^2} \tag{10}$$

$$b=\frac{\sum_{i=1}^{n}\ln(x_i)-a\sum_{i=1}^{n}\ln\left(-\ln\left[1-\hat{F}(\mathrm{x}_i)\right]\right)}{n} \tag{11}$$

hence we have $\beta=\frac{1}{a}$ and $\eta=\exp(b)$, where $F(\mathrm{x})$ can be estimated by the mean rank approach (Pobocikova,2012), $\hat{F}(\mathrm{x}_i)=\frac{i}{n+1}$.

IV. SIMULATION AND ANALYSIS

In this study, we first simulated data using the Monte Carlo simulation. Random datasets were generated from the uniform distribution for the study. The datasets generated from the uniform distribution were sampled uniformly from the range of 0 and 1 to represent the failure times(hours) for a particular product. The randomly selected samples from the uniform distribution were simulated with the model $X=\eta\left(-\ln(z)\right)^{1/\beta}$ to represent the failure times for the product as a random sample from the Weibull distribution, where z has the uniform distribution (0,1). We simulated 36,000 failure times and chose $n=50,100$, and 1000, we chose different scale and shape parameters to simulate the data for the Weibull distribution, and the sample sizes n were chosen after each simulation, this was replicated 1000 times in each simulation. The shape parameter values β were chosen so that $\beta=0.5$ will represent the scenario when $\beta<1$ and the other values (2,3) of the shape parameter when $\beta>1$, and $\beta=1$ when we have constant failure times. The two parameters for the Weibull distribution were estimated through the use of the maximum likelihood estimation and the least-squares method. The performance of these two estimation techniques was compared using the goodness of fit measures. The maximum likelihood approach estimates the parameters of the 2-parameter Weibull distribution better than the least-squares method. The results for the parameter estimation using the maximum likelihood and the least-squares method are tabulated in Table 1. By comparing the goodness of fit for the MLE and the LS, the MLE performs better in estimating the model parameters than the least-squares method.

The probability density function and the cumulative density of the different parameter values used to simulate the Weibull distribution for this study are shown in Fig.1. The probability plot and the cumulative density function (CDF) are shown in Fig. 2.

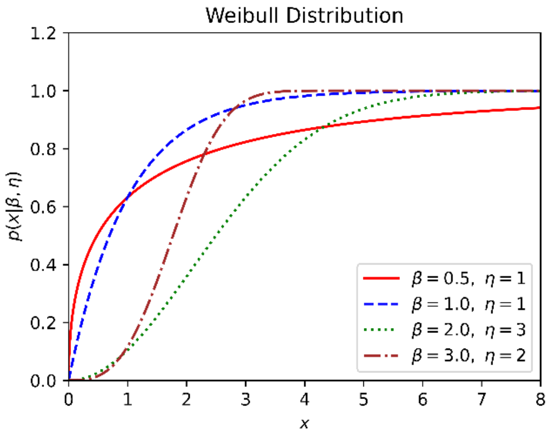
Table 1. Parameter estimation of the Weibull distribution

Maximum Likelihood Estimation

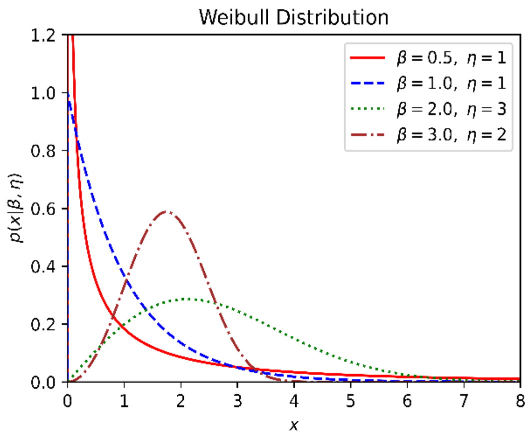
	Parameter Values		Estimates		Confidence Interval		Goodness of fit			
size	η (α)	β	$\hat{\eta}$ ($\hat{\alpha}$)(SE)	$\hat{\beta}$ (SE)	CI($\hat{\eta}$ ($\hat{\alpha}$))	CI($\hat{\beta}$)	Log-likelihood	AICc	BIC	AD
50	2	0.5	1.8310(0.5454)	0.4949(0.0544)	(1.0213,3.2829)	(0.3990,0.6138)	-87.4434	179.142	182.711	0.4928
	1	1	1.2068(0.1495)	1.2060(0.1315)	(0.9467,1.5384)	(0.9739,1.4934)	-54.8121	113.880	117.448	0.4579
	3	2	3.2545(0.2276)	2.1122(0.2389)	(2.8377,3.7325)	(1.6923,2.6363)	-88.0339	180.323	183.892	1.0994
	2	3	2.0121(0.0952)	3.1467(0.3563)	(1.8340,2.2075)	(2.5205,3.9285)	-47.8072	99.8696	103.438	0.4987
100	2	0.5	1.9099(0.4687)	0.4288(0.0344)	(1.1807,3.0896)	(0.3664,0.5018)	-170.569	345.261	350.348	0.5962
	1	1	1.2241(0.1171)	1.1052(0.0853)	(1.0148,1.4766)	(0.9501,1.2856)	-115.496	235.115	240.202	0.7073
	3	2	2.8691(0.1373)	2.1891(0.1756)	(2.6122,3.1512)	(1.8706,2.5617)	-160.233	324.589	329.676	0.8453
	2	3	1.9377(0.0720)	2.8296(0.2240)	(1.8017,2.0841)	(2.4229,3.3045)	-100.351	204.826	209.912	0.3510
1000	2	0.5	1.9012(0.1252)	0.5065(0.0125)	(1.6710,2.1631)	(0.4826 ,0.5315)	-1767.24	3538.5	3548.3	0.7730
	1	1	1.0413(0.0345)	1.0042(0.0247)	(0.9757,1.1112)	(0.9569,1.0538)	-1038.67	2081.36	2091.16	0.2608
	3	2	3.0284(0.0495)	2.0362(0.0505)	(2.9329,3.1270)	(1.9395,2.1376)	-1692.99	3389.98	3399.79	0.2070
	2	3	2.0028(0.0212)	3.1398(0.0776)	(1.9616,2.0449)	(2.9913,3.2956)	-944.250	1892.51	1902.32	0.3783

Least-squares Estimation

size	η (α)	β	$\hat{\eta}$ ($\hat{\alpha}$) (SE)	$\hat{\beta}$ (SE)	CI($\hat{\eta}$ ($\hat{\alpha}$))	CI($\hat{\beta}$)	Log-likelihood	AICc	BIC	AD
50	2	0.5	1.8310(0.5454)	0.4949(0.0544)	(1.0213,3.2829)	(0.3990,0.6138)	-87.4434	179.142	182.711	0.4928
	1	1	1.2010(0.1463)	1.2223(0.1329)	(0.9459,1.5248)	(0.9877,1.5125)	-54.8233	113.902	117.471	0.4810
	3	2	3.3544(0.3018)	1.7234(0.2069)	(2.8121,4.0014)	(1.3620,2.1806)	-89.7926	183.84	187.409	1.5810
	2	3	2.0162(0.1016)	2.9541(0.3382)	(1.8266,2.2254)	(2.3603,3.6972)	-47.9805	100.216	103.785	0.4353
100	2	0.5	1.9055(0.4938)	0.4123(0.0333)	(1.1467,3.1666)	(0.3519,0.4831)	-170.697	345.518	350.604	0.4999
	1	1	1.2004(0.1067)	1.1784(0.0897)	(1.0085,1.4289)	(1.0150,1.3680)	-116.000	236.124	241.21	1.0681
	3	2	2.9312(0.1672)	1.8901(0.1571)	(2.6211,3.2779)	(1.6059,2.2245)	-162.173	328.469	333.556	1.1726
	2	3	1.9515(0.0784)	2.6302(0.2117)	(1.8037,2.1114)	(2.2463,3.0798)	-100.866	205.855	210.941	0.3884
1000	2	0.5	1.8555(0.1157)	0.5229(0.0127)	(1.6420,2.0967)	(0.4985 ,0.5485)	-1768.47	3540.94	3550.75	1.4114
	1	1	1.0383(0.0341)	1.0122(0.0249)	(0.9736,1.1076)	(0.9646,1.0621)	-1038.74	2081.5	2091.31	0.3034
	3	2	3.0364(0.0507)	1.9972(0.0498)	(2.9387,3.1374)	(1.9019,2.0972)	-1693.37	3390.76	3400.56	0.2147
	2	3	2.0024(0.0213)	3.1323(0.0775)	(1.9611,2.0445)	(2.9841,3.2879)	-944.255	1892.52	1902.32	0.3718



1A: Probability density plots for the different parameter values



1B: Cumulative density plots for the different parameter values

Fig.1: Plots of the probability density and cumulative density of the Weibull distribution with different values of the shape (β) and scale parameter (η (α)).

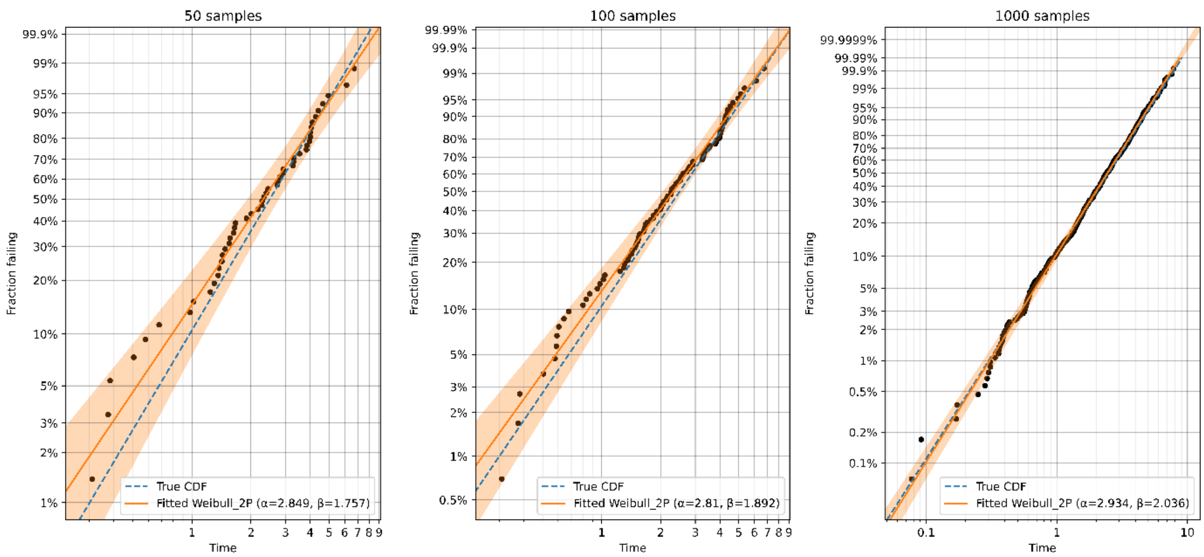


Fig.2: Plot of the cumulative density function and the probability plot of the data values of three different samples randomly selected from the Weibull distribution and their estimated values of the shape (β) and scale parameter (η (α)).

4.1 Applications in Reliability Analysis

Health

Considering the application of Weibull distribution in reliability analysis, we use the data on 128 randomly selected remission times (months) of bladder cancer patients. The data have been used by many researchers for different research activities, Zea *et al.* (2012) used the data comparing sub-model including beta-Pareto distribution and others, and the beta exponential Pareto distribution. Almheidat *et al*. (2015) also investigated the Cauchy-Weibull distribution with this dataset. We apply the Weibull distribution to fit the same data, however, we removed all outliers that are 2-standard deviation away from the mean from the dataset before fitting the Weibull distribution to the data. In order to ascertain that the correct distribution is fitted, we fitted many other distributions such as the Lognormal, Gamma, Exponential, Loglogistic, etc. to the data after removing all outliers from the data and then select the best distribution that fits the data. This reduces the sample size from 128 to 122 remission times of bladder cancer patients. Fig.3 indicates the fit of distributions to the data. The data for the 128 remission times (months) of bladder cancer patients are shown below in Table 2.

Table 2: Remission times (months) of bladder cancer patients

0.08	1.46	2.69	3.57	4.51	5.62	7.32	8.66
11.98	15.96	25.82	0.2	1.76	2.69	3.64	4.87
5.71	7.39	9.02	12.02	16.62	26.31	0.4	2.02
2.75	3.7	4.98	5.85	7.59	9.22	12.03	17.12
32.15	0.5	2.02	2.83	3.82	5.06	6.25	7.62
9.47	12.07	17.14	34.26	0.51	2.07	2.87	3.88
5.09	6.54	7.63	9.74	12.63	17.36	36.66	0.81
2.09	3.02	4.18	5.17	6.76	7.66	10.06	13.11
18.1	43.01	0.9	2.23	3.25	4.23	5.32	6.93
7.87	10.34	13.29	19.13	46.12	1.05	2.26	3.31
4.26	5.32	6.94	7.93	10.66	13.8	20.28	79.05
1.19	2.46	3.36	4.33	5.34	6.97	8.26	10.75
14.24	21.73	1.26	2.54	3.36	4.34	5.41	7.09
8.37	11.25	14.76	22.69	1.35	2.62	3.48	4.4
5.41	7.26	8.53	11.64	14.77	23.63	1.4	2.64
3.52	4.5	5.49	7.28	8.65	11.79	14.83	25.74

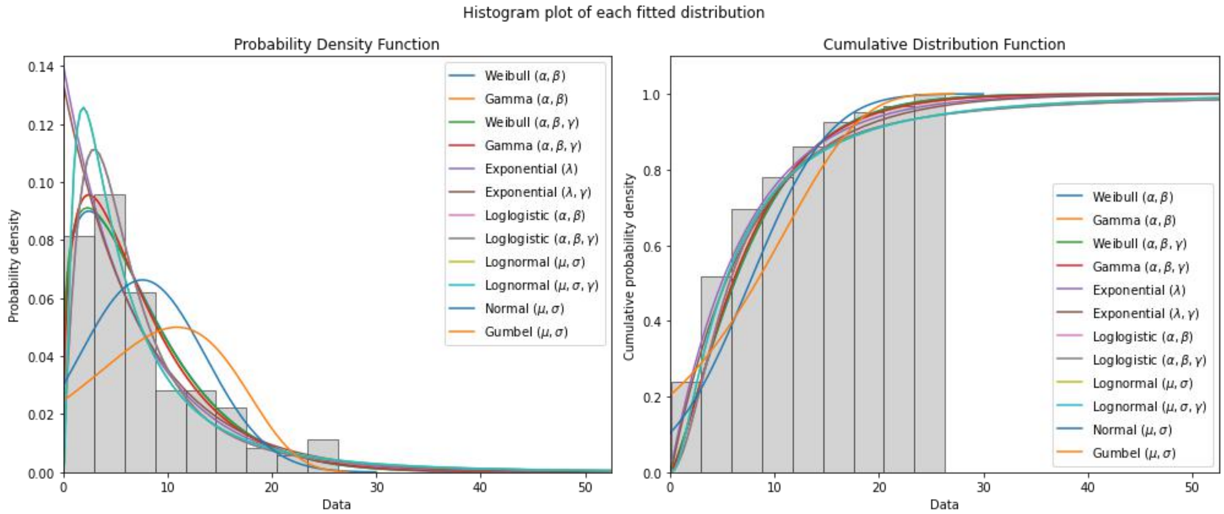


Fig.3: Histogram with each fitted distribution with the plots of its PDF and CDF.

Table 3: Comparison of the various distributions for the remission times data

Distribution	Alpha	Beta	Gamma	Mu	Sigma	Lambda	Log-likelihood	AICc	BIC	AD
Weibull_2P	8.204	1.276					-364.19	732.481	737.988	0.322
Gamma_2P	5.130	1.482					-364.34	732.787	738.294	0.262
Weibull_3P	8.128	1.258	0.0413				-364.11	734.421	742.63	0.315
Gamma_3P	5.130	1.482					-364.34	734.889	743.098	0.262
Exponential_1P						0.13153	-369.77	741.58	744.351	3.479
Exponential_2P			0.0799			0.13293	-368.19	740.483	745.99	2.16
Loglogistic_2P		1.864					-370.61	745.312	750.819	0.566
Loglogistic_3P		1.864	0				-370.61	747.414	755.623	0.566
Lognormal_2P				1.655	0.998		-374.72	753.539	759.046	1.365
Lognormal_3P				1.655	0.998		-374.72	755.642	763.85	1.365
Normal_2P				7.603	6.021		-392.13	788.35	793.857	4.218
Gumbel_2P				10.922	7.362		-420.56	845.226	850.734	8.34

Table 3 compares the various distributions using the AICc, and BIC and it was observed that the 2-parameter Weibull distribution is the best fit for the data. Fig. 4 shows the probability plot of each of the distributions compared, from Fig.4 there is an indication that the 2-parameter Weibull distribution fits better than the other distributions.

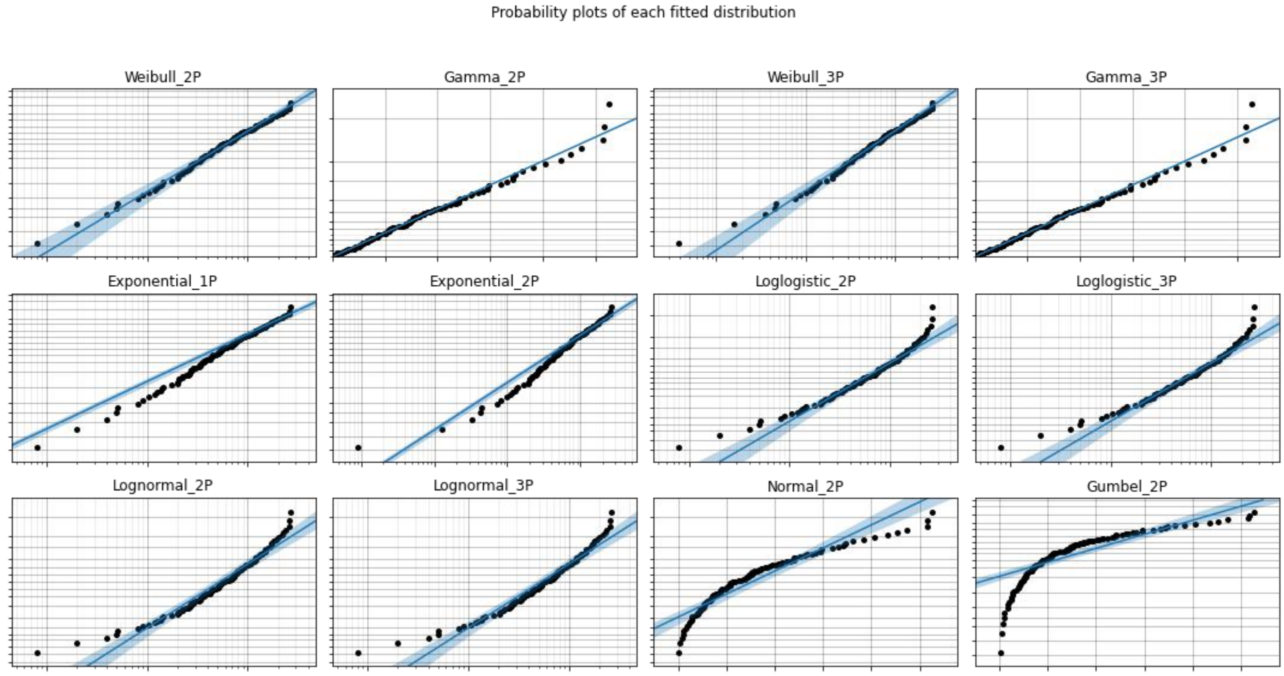


Fig.4: Probability plot of the remission times (months) of the bladder cancer patients for the various distributions.

The 2-parameter Weibull distribution is considered as the best distribution for the data, we, therefore, study the reliability of the remission times of the bladder cancer patients using the distribution chosen. The maximum likelihood estimation method and the least-squares method are considered in fitting the parameters of the Weibull distribution, the results of the two methods are compared using the AICc and the BIC. The parameter estimates are shown in Table 4. The MLE approach performs better in estimating the parameters of the remission times of the bladder cancer patients compared with the least-squares method.

Table 4: Parameter estimate of the Weibull distribution for the remission times

	Parameter	Point Estimate	Standard Error	Lower CI	Upper CI	Log-likelihood	AICc	BIC	AD
MLE	Alpha	8.20364	0.613436	7.08528	9.49852	-364.19	732.481	737.988	0.322
	Beta	1.27557	0.0898488	1.11108	1.4644				
LS	Alpha	8.24435	0.63076	7.0963	9.57813	-364.24	732.576	738.084	0.317
	Beta	1.25165	0.08858	1.08953	1.43789				

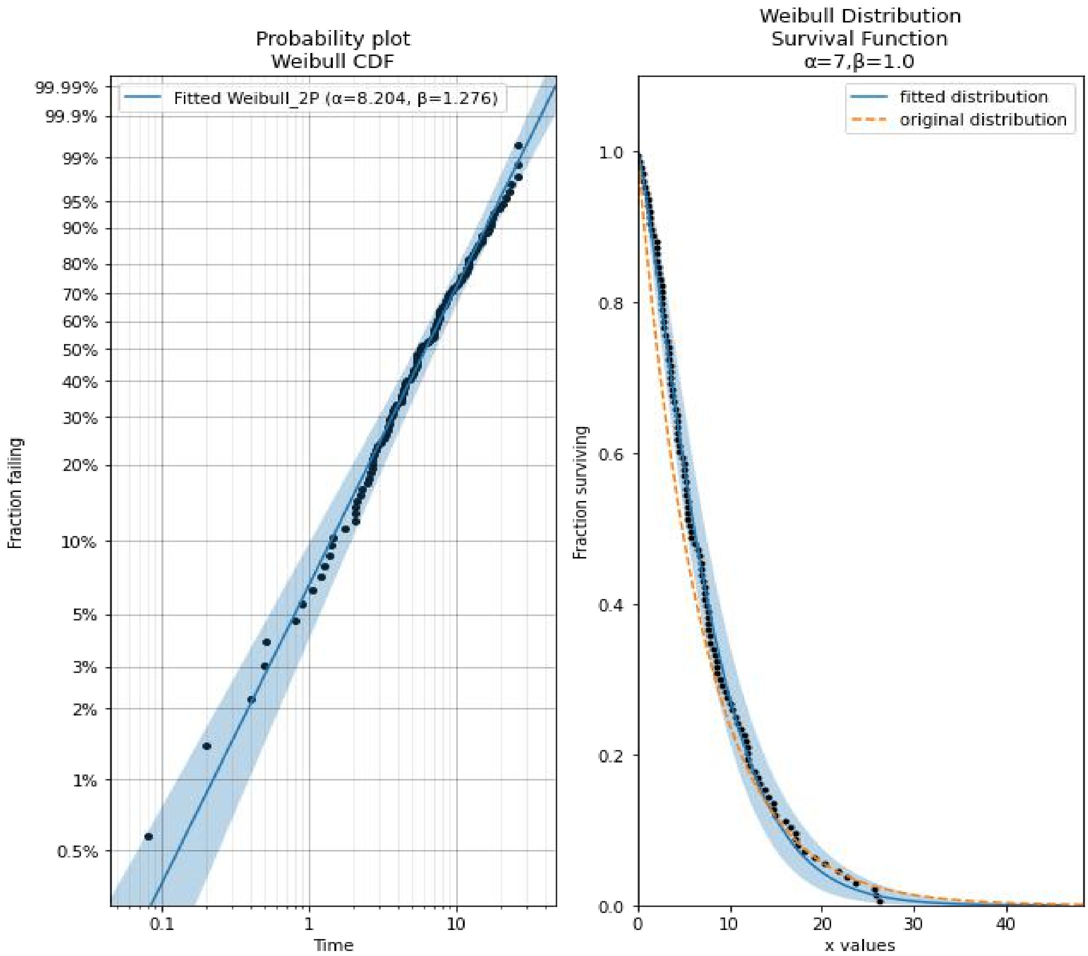


Fig.5: Probability plot and the plot of the reliability function for the remission times.

The plot in Fig.5 gives the probability plot for the Weibull distribution and the reliability function for the remission times (month) for the bladder cancer patients. The estimated shape parameter (1.276) using the MLE indicates that the remission rate of the bladder cancer patient increase as the remission time increases. The B10 life of the data was 1.4 months which indicates that we expect 10% of the bladder cancer patients to be remitted after 1.4 months.

Industrial

We again consider breaking stress of carbon fibers of 50mm length (GPa), the data was first used by Nicholas and Padgett (2006), Lemonte (2011) using β -Birnbaum–Saunders distribution, Al-Aqtash *et.al* (2014) use the data to study the application of Gumbel-Weibull distribution. For our study, we remove all outliers that are 2- standard deviation away from the mean, this reduces the dataset from 66 to 59. The dataset was fitted with other distributions and the best distribution was the 2-parameter Weibull distribution. Fig.6 illustrates the fit of the distributions to the breaking stress of the carbon fibers data. The carbon fibers data is shown in Table 5.

Table 5: Breaking stress of Carbon fibers

0.39	0.85	1.08	1.25	1.47	1.57	1.61	1.61	1.69	1.8	1.84
1.87	1.89	2.03	2.03	2.05	2.12	2.35	2.41	2.43	2.48	2.5
2.53	2.55	2.55	2.56	2.59	2.67	2.73	2.74	2.79	2.81	2.82
2.85	2.87	2.88	2.93	2.95	2.96	2.97	3.09	3.11	3.11	3.15
3.15	3.19	3.22	3.22	3.27	3.28	3.31	3.31	3.33	3.39	3.39
3.56	3.6	3.65	3.68	3.7	3.75	4.2	4.38	4.42	4.7	4.9

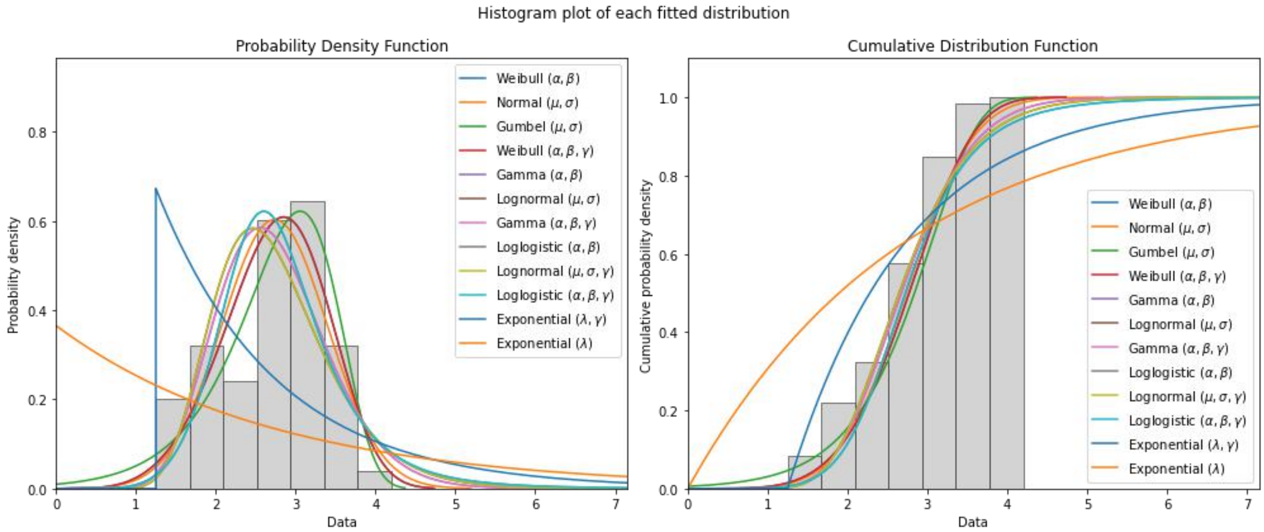


Fig.6: Histogram with each fitted distribution with the plots of its PDF and CDF.

Table 6: Comparison of the various distributions for the breaking stress of carbon fibers data

Distribution	Alpha	Beta	Gamma	Mu	Sigma	Lambda	Log-likelihood	AICc	BIC	AD
Weibull_2P	2.990	4.836					-58.2972	120.809	124.749	0.578
Normal_2P				2.736	0.661		-59.3269	122.868	126.809	0.771
Gumbel_2P				3.056	0.591		-59.9252	124.065	128.005	0.519
Weibull_3P	2.990	4.836	0.000				-58.2972	123.031	128.827	0.578
Gamma_2P	0.180	15.203					-61.4958	127.206	131.147	1.265
Lognormal_2P				0.973	0.268		-63.3661	130.946	134.887	1.595
Gamma_3P	0.180	15.203	0.000				-61.4958	129.428	135.224	1.265
Loglogistic_2P	2.723	6.619					-63.6408	131.496	135.437	1.192
Lognormal_3P			0.000	0.973	0.268		-63.3661	133.169	138.965	1.595
Loglogistic_3P	2.723	6.619	0.000				-63.6408	133.718	139.514	1.192
Exponential_2P			1.250			0.673	-82.3637	168.942	172.883	7.952
Exponential_1P						0.366	-118.378	238.827	240.834	15.551

Table 6 compares the distributions and the best distribution that best fit the breaking stress of the carbon fibers data after removing all observations that were 2-standard deviation away from the mean was the 2-parameter Weibull distribution. The comparison was done by using the AICc and BIC. From Fig.7, it can be observed that the probability plot of the 2-parameter Weibull distribution achieves a better result than the other distributions.

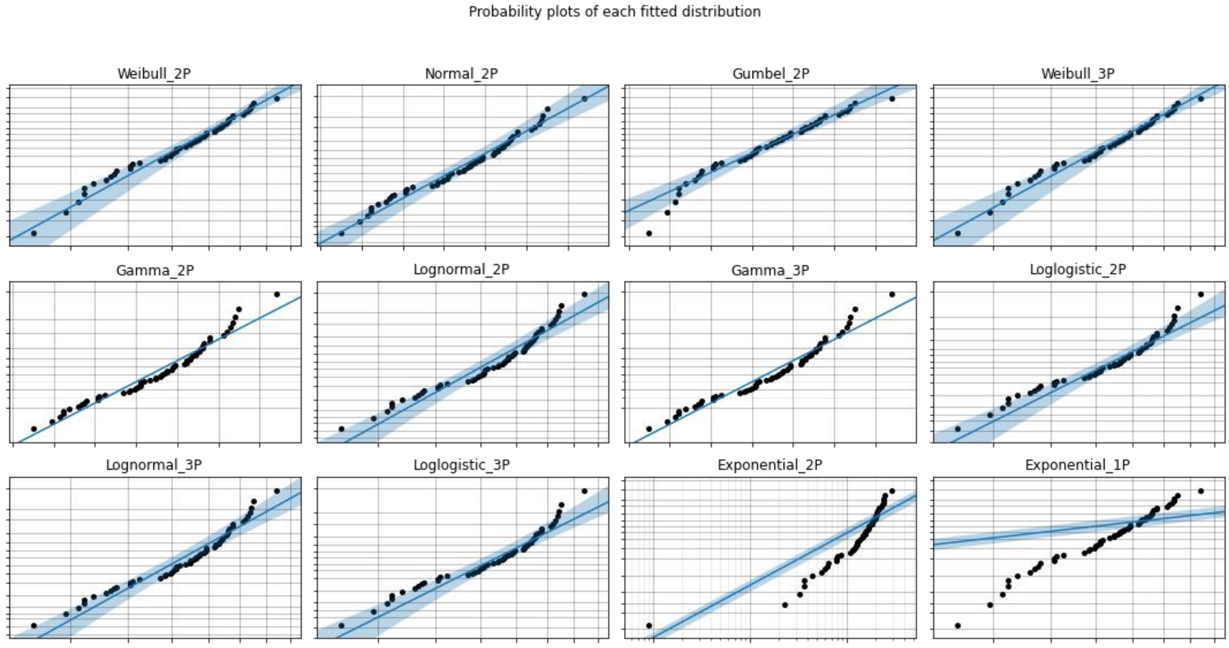


Fig.7: Probability plot of breaking stress of carbon fibers data for the various distributions.

We study the breaking stress of carbon fibers using the Weibull distribution to assess its reliability. We use the maximum likelihood estimation and the least-squares method to estimate the scale and shape parameter of the carbon fibers data, we compare these two methods using the AICc and the BIC. Table 7 presents the parameter estimate of the breaking stress of the carbon fibers using the 2-parameter Weibull distribution.

Table 7: Parameter estimate of the Weibull distribution for the breaking stress of carbon fibers

	Parameter	Point Estimate	Standard Error	Lower CI	Upper CI	Log-likelihood	AICc	BIC	AD
MLE	Alpha	2.99031	0.08466	2.8289	3.16093	-58.297	120.809	124.749	0.578
	Beta	4.83619	0.50136	3.94694	5.92579				
LS	Alpha	2.99167	0.08904	2.82214	3.17138	-58.436	121.086	125.027	0.547
	Beta	4.83619	0.50136	3.94694	5.92579				

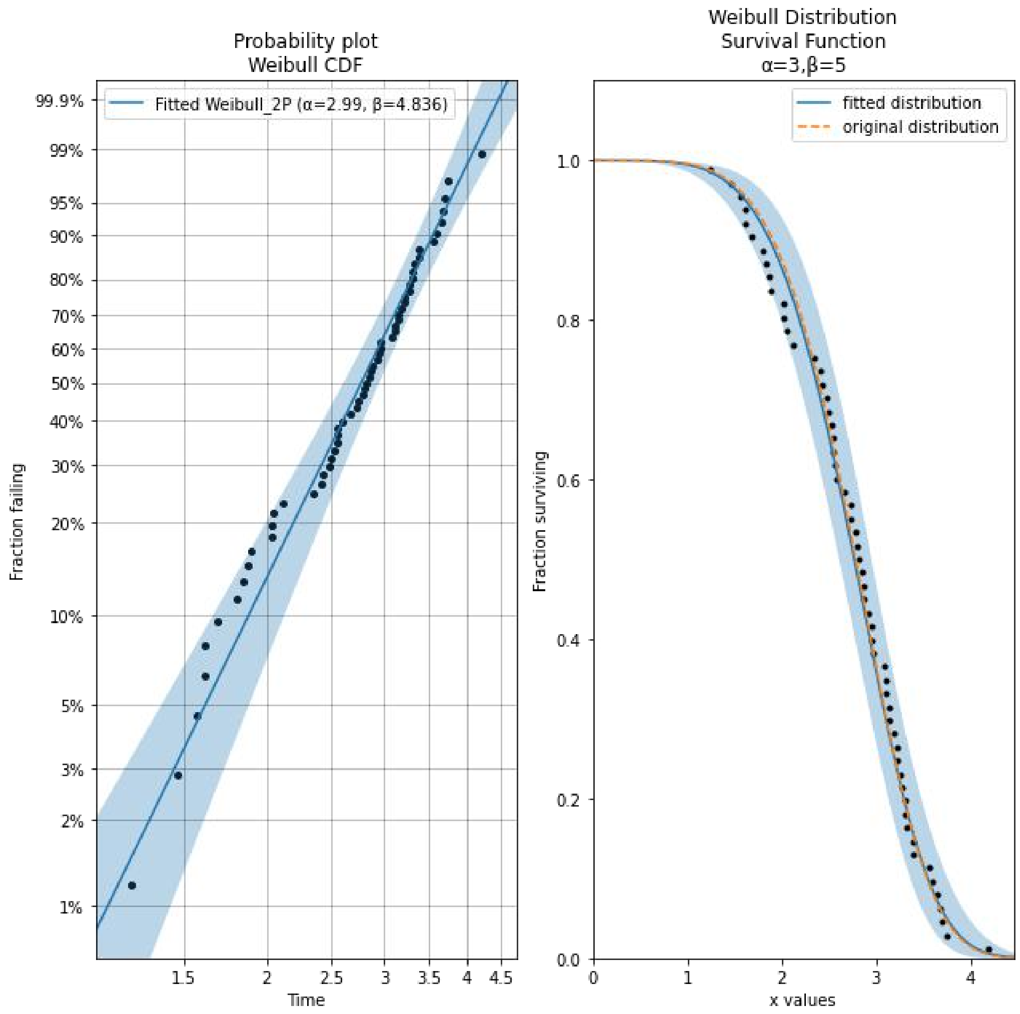


Fig.8: Probability plot and the plot of the reliability function for the breaking stress of carbon fibers.

The plot in Fig.8 gives the probability plot for the Weibull distribution and the reliability function for the breaking stress of the carbon fibers data. The maximum likelihood estimator gives the best estimate with an estimated shape parameter (4.836) which indicates that the rate of breaking stress increase as the stress time increases. The B10 life of the data was 2.77206 GPa which indicates that we expect 10% of the carbon fibers to break after obtaining breaking stress of 2.77206 GPa.

V. CONCLUSION

The paper examines two approaches for estimating the parameters of the 2- parameter Weibull distribution. In this research, we compared two estimation methods using artificial data generated with Monte Carlo simulation and two real datasets. The research findings indicate that the maximum likelihood estimation method performs better than the least-squares approach.

In general, the Weibull distribution is very useful in many areas such as survival analysis, reliability analysis, etc. The estimation of the 2-parameter or 3-parameter Weibull distribution with or without censored data has gained much attention in these areas of study. In this paper, we compared the maximum likelihood estimation and the least-squares method, and by assessing the properties of the distribution and the estimates of the 2-parameter Weibull distribution using the goodness of fit measures. We observed that the maximum likelihood estimation performs better than the least-squares method and the ML estimator is asymptotically efficient as it converges in distribution as the sample size increases.

In order to investigate the application of the Weibull distribution in reliability analysis. We estimated the parameter of the Weibull distribution considering two different datasets using the MLE and LS and examined the failure rate of the datasets. The findings also indicate that the MLE performs better than the least-squares method.

REFERENCES

[1] Al-Fawzan, M. A. (2000). Methods for estimating the parameters of the Weibull distribution. *King Abdulaziz City for Science and Technology, Saudi Arabia*.

[2] Al-Aqtash, R., Lee, C., & Famoye, F. (2014). Gumbel-Weibull distribution: Properties and applications. *Journal of Modern applied statistical methods*, *13*(2), 11.

[3] Almheidat, M., Famoye, F., & Lee, C. (2015). Some generalized families of Weibull distribution: Properties and applications. *International Journal of Statistics and Probability*, *4*(3), 18.

[4] Bahadur, R. R. (1964). On Fisher's bound for asymptotic variances. *The Annals of Mathematical Statistics*, *35*(4), 1545-1552.

[5] Cordeiro, G. M., & Lemonte, A. J. (2011). The β-Birnbaum–Saunders distribution: An improved distribution for fatigue life modeling. *Computational Statistics & Data Analysis*, *55*(3), 1445-1461.

[6] Chu, Y. K., & Ke, J. C. (2012). Computation approaches for parameter estimation of Weibull distribution. *Mathematical and computational applications*, *17*(1), 39-47.

[7] Guure, C. B., Ibrahim, N. A., & Ahmed, A. O. M. (2012). Bayesian estimation of 2-parameter weibull distribution using extension of Jeffreys' prior information with three loss functions. *Mathematical Problems in Engineering*, *2012*.

[8] Nicholas, M. D. & Padgett, W. J. (2006). A bootstrap control for Weibull percentiles. *Quality and Reliability Engineering International*, *22*, 141-151.

[9] Pobocikova, I., & Sedliackova, Z. (2012). The least square and the weighted least square methods for estimating the Weibull distribution parameters–a comparative study. *Communications-Scientific letters of the University of Zilina*, *14*(4), 88-93.

[10] Pobocikova, I., Sedliackova, Z., Michalkova, M., & George, F. (2017). Monte Carlo Comparison of the Methods for Estimating the Weibull Distribution Parameters–Wind Speed Application. *Communications-Scientific letters of the University of Zilina*, *19*(2A), 79-86.

[11] Razali, A. M., Salih, A. A., & Mahdi, A. A. (2009). Estimation accuracy of Weibull distribution parameters. *Journal of Applied Sciences Research*, *5*(7), 790-795.

[12] Sun, D. (1997). A note on noninformative priors for Weibull distributions. *Journal of Statistical Planning and Inference*, *61*(2), 319-338.

[13] Zea, L. M., Silva, R. B., Bourguignon, M., Santos, A. M., & Cordeiro, G. M. (2012). The beta exponentiated Pareto distribution with application to bladder cancer susceptibility. *International Journal of Statistics and Probability*, *1*(2), 8.