Journal of Positive Sciences (JPS) eISSN: 2582-9351, https://positive-sciences.com JPS. 2021 Oct; 1(7): 6-10

Article info
Received on: 29.09.2021
Accepted on: 29.10.2021
Published on: 31.10.2021
doi: https://doi.org/10.52688/ASP57646

Research Article

Further methods of estimating the parameters of the weibull distribution

F. A. Al-Baidhani [1, *]

[1] Business Administration Department, Baghdad College of Economics Science University, Baghdad, Iraq

[*] fadhilajab@gmail.com

ABSTRACT

Probability-weighted moments, an order statistic technique, generalization of the total Q on test, maximum likelihood, generalized least squares, and ordinary least squares are all evaluated as methods for estimating the parameters of the Weibull distribution. The root mean squared error in the estimated 95 percent and 99 percent quantiles is used to compare (a) the effectiveness of parameter estimations and (b) the root mean squared error in the estimated 95 percent and 99 percent quantiles. Simulations are used to look into these requirements for sample sizes of 10 and 25, as well as a wide range of shape parameter values.

Keywords: Probability-weighted moments; order statistic method; generalization of the total Q on test; maximum likelihood generalized least squares

INTRODUCTION

The probability density function for the two parameter Weibull distribution is

$$f(x;\ \alpha,\beta) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} \exp\left\{-\left(\frac{x}{\alpha}\right)^{\beta}\right\}\ ,\ x > 0 \qquad (1)$$

Where the scale and form parameters are >0 β >0, respectively. The topic under consideration is that of estimating and, with quantile prediction as a secondary goal. We focus on samples that are relatively tiny in size, such as those seen in engineering reliability applications. We investigate three unknown methods (1, 2, 3) and compare them to three well-known methods using pseudo-random samples of size 10 and 25, for various values of. These are the ones. Methodology for Order Statistics (OSM) Probability-Weighted Moments are a type of probability-weighted moment (PWM) The Total Q on Test has been generalised (TQT) Maximum Likelihood Estimation (MLE) is a statistical technique for estimating the probability of (ML) Estimation with Generalized Least Squares (GLS) Estimation using Ordinary Least Squares (OLS)

Methods 4, 5 and 6 are described in detail in [1, 2], who contrasted the ML, GLS, and OLS methods with the [3, 4] method for small values of the shape parameter (5). We extend the range of values up to 30 here to cover a wide variety of practical values in materials engineering. Two sorts of comparisons are used to examine the utility of estimating methods: Estimation efficiency of parameters. In the predicted 95 percent and 99 percent confidence intervals, the root of the mean squared relative error.

METHODOLOGY

The following simple, unbiased estimators for the shape parameter (Beta) of the two Weibull distributions were proposed by [5] [6].

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

where s is the greatest integer that does not exceed (0.84 n), n is the sample size, and k_n is an unbiasing constant that varies with the sample size. To find k_n, utilize the table provided by [5, 6]. We get a simple estimator for the scale parameter (alpha) by equating _1', the first moment about the origin, to its predicted value. Generally speaking:

*Corresponding author
F. A. Al-Baidhani,
Business Administration Department, Baghdad College of Economics Science University, Baghdad, Iraq
e-mail: fadhilajab@gmail.com

$$\mu_1' = E\{x^r\} = \int_0^{\infty} x^r\, f(x;\, \beta, \alpha) dx = \alpha^r\, \Gamma\left(1 + \frac{r}{\beta}\right) \quad (3)$$

where Γ is the gamma function.

$$\text{Then} = \mu_1' = \alpha\Gamma\left(1 + \frac{1}{\beta}\right) \quad (4)$$

So, the scale estimator is:

$$\hat{\alpha} = \frac{\bar{x}}{\Gamma\left(1+\frac{1}{\beta}\right)} \quad (5)$$

where $\bar{x}$ is the sample mean and β is the shape estimate (2).

TQT

The Total Time on Test (TTT) idea was proposed by Epstein and Sobel (1953), and numerous expansions of this concept have been defined and researched (e.g. TTT-transform and TTT-plot) (1975). The scaled Total Q on Test transform (STQTT) and the empirical Q on Test ratio were proposed by Jewell (1977) as a generalization of the scaled TTT-transform and scaled TTT-plot (EQTR). He equated the cumulative hazard function, H(x), to the prototype failure function, Q(x), which contains all of the failure shape information, and is a constant. The EQTR tends to resemble the STQTT as the number of data points increases, so it can be used for parameter estimation and model identification. The EQTR for an ordered random sample x_i, i=1,2,…,n is calculated using the two parameter Weibull distribution. The cumulative distribution function is described as follows:

$$F(x) = 1 - exp\left\{-\left(\frac{x}{\alpha}\right)^{\beta}\right\}\ x \geq 0 \quad (6)$$

The survivor function is:

$$S(x) = exp\left\{-\left(\frac{x}{\alpha}\right)^{\beta}\right\}\ x \geq 0 \quad (7)$$

An initial estimate of the shape parameter can be obtained following [7]:

$$\beta_1 = \frac{2.989}{\ln\left(\frac{x_k}{x_h}\right)} \quad (8)$$

where x_k and x_h are the 97th and 17th sample percentiles respectively. It is convenient to reparametrize using:

$$\theta = \alpha^{-\beta} \quad Q(x) = x^{\beta} \quad and\ \ q(x) = \frac{dQ(x)}{dx} \quad (9)$$

Then the likelihood of the ordered random sample can be expressed as:

$$L\ (x_1, x_2, \ldots\ x_n; \theta) = \frac{n!}{(n-r)!}\ [S(x)]^{n-r}\ \prod_{i=1}^{r} f(x_i) \quad (10)$$

$$= \frac{n!}{(n-r)!} \prod_{i=1}^{r} q(x_i) \left\{\theta^r\ e^{-\theta[TQT(x_1,\ldots,x_r)]}\right\} \quad (11)$$

where the total Q on test statistic is:

$$TQT\ (x_1, x_2 \ldots . x_r) = \sum_{i=1}^{r} Q(x_i) + (n-r)\ Q(x_r) \quad (12)$$

$$EQTR = \frac{Q_i}{Q_n} = \frac{TQT\ (x_1,\ldots,x_i)}{TQT\ (x_1,\ldots,x_n)} \quad (13)$$

where Q_n is the actual finite sum:

$$Q_n = \int_0^{F_n^{-1}\left(\frac{i}{n}\right)} \bar{F}_n(U) dQ(U) = \frac{1}{n}\ TQT\ (x_1, \ldots, x_n) \quad (14)$$

Then we estimate β by the value that makes the plot of $\frac{Q_1}{Q_n}$ versus $\left(\frac{i}{n}\right), i = 1, \ldots n$ as near as possible to the straight line through the points (0,0) and (1,1). This can readily be done using the OPTIMISE facility in GENSTAT.

*Corresponding author
F. A. Al-Baidhani,
Business Administration Department, Baghdad College of Economics Science University, Baghdad, Iraq
e-mail: fadhilajab@gmail.com

This approach does not lead directly to an estimate of α. However an estimate can be obtained via the expression for the maximum likelihood estimate, but based on the TQT estimate of β.

$$\hat{\alpha} = \left[\frac{1}{n}\sum_{i=1}^{n} Q(x_i)\right]^{\frac{1}{\beta}} \tag{15}$$

The plot of EQTR (Q_i/Q_n) versus (i/n), I – 0, 1,..., n can be used for model identification since the EQTR should approximate to a straight line if the data is a random sample from the Weibull distribution.

PWM

Probability weighted moments, a generalization of the usual moments of a probability distribution, were introduced by [8]. The use of these moments to estimate the parameters of the generalized extreme-value (a close relation of the three-parameter Weibull distribution) was proposed by [9]. The probability weighted moment of order i, j, k for a random variable x with c.d.f. F(x) is defined as

$$M_{i,j,k} = E\,[x^i\,[F(x)]^j\,[1-F(x)]^k] \tag{16}$$

$$= \int_0^1 x(F)^i\,F^j\,(1-F)^k\,dF$$

Where i, j, k are non-negative integers and x(F) is the inverse c.d.f. PWM estimation proceeds by equating sample estimates of M100, M101 to the corresponding population expressions. For the two parameter Weibull, the population expressions are:

$$M_{1,0,k} = E\,[x(1-F)^k = \alpha(1-F)^{-\gamma}\,\Gamma(\gamma)] \quad k = 0, 1, 2, \ldots \tag{17}$$

where $\gamma = 1 + \frac{1}{\beta}$

The sample P.W.M.s obtained from an ordered random sample $\{x_i\}, i = 1, 2, \ldots., n$; from the two parameter Weibull distribution are

$a_0 = m_{100} = \bar{x}$

$a1 = m_{101} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{n-i}{n-1}\right) x_{(i)}$

β is the solution of the non-linear equation:

$m_{100} = 2^{\gamma} m_{101}$

$i.e. \hat{\beta} = \frac{1}{(\hat{\gamma}-1)}$ where $\hat{\gamma} = \frac{\log\left(\frac{m_{100}}{m_{101}}\right)}{\log(2)}$

and $\hat{\alpha} = \frac{\bar{x}}{\Gamma(\hat{\gamma})}$

EXPERIMENTAL MODEL

In order to illustrate the estimation methods in the previous sections, we use two data sets from [6][7][8] of sizes n = 15 and n = 26, respectively. The results of applying methods 1, 2 and 3, together with some intermediate stages are summarized in table 1.

Table 1: Results of TOT, PWM, OSM models

	β_1	$\hat{\beta}$	$\hat{\alpha}$	n		
TOT	2.084	1.8513	64.272	15		
	10.174	11.095	59567	26		
	a0	a1	$\hat{\beta}$	$\hat{\alpha}$	n	
PWM	58.200	20.200	1.8987	65.586	15	
	57573	27551	15.801	59526	26	
	$\hat{\beta}$	$\hat{\alpha}$	x	s	kn	n
OSM	2.0202	65.683	58.200	12.000	1.4004	15
	13.778	59785	57573	21.000	1.4479	26

*Corresponding author
F. A. Al-Baidhani,
Business Administration Department, Baghdad College of Economics Science University, Baghdad, Iraq
e-mail: fadhilajab@gmail.com

The results of all six methods are summarized in table 2.

Table 2: The results of all six methods proposed in this paper

	GLS	PWM	OLS	OSM	TQT	ML	n
$\hat{\beta}$	1.7860	1.8987	1.8608	2.0202	1.8513	2.1118	15
$\hat{\alpha}$	67.085	65.586	66.529	65.683	64.272	65.987	15
$\hat{\beta}$	11.638	15.801	16.027	13.778	11.095	11.690	26
$\hat{\alpha}$	59774	59526	59351	59785	59567	59707	26

SIMULATION

A simulation study was conducted in order to compare the six methods. For each of the shape parameter values $\beta = 0.5, 1, 2, 5, 10, 20\ and\ 30$, one thousand random samples of size 10 and 25 were generated from a Weibull distribution with $\alpha = 1$. The random number generator used was that in the GENSTAT package [5]. The methods are compared through (a) the relative efficiency (RE) of parameter estimates and (b) the relative bias (RB) and root mean squared relative error (RMSE) of the predicted quantities.

For the 95% and 99% quantities we define

$$MSE\ \hat{\beta} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\beta} - \beta\right)^2 \tag{18}$$

$$MSE\ \hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n}(\hat{\alpha} - \alpha)^2 \tag{19}$$

$$RE\ of\ \hat{\beta} = \frac{\left[\frac{0.608(\beta)^2}{n}\right]}{MSE\ (\hat{\beta})} \tag{20}$$

$$RE\ of\ \hat{\alpha} = \frac{\left[\frac{1.109\left(\frac{\alpha}{\beta}\right)^2}{n}\right]}{MSE\ (\hat{\alpha})} \tag{21}$$

$$RB = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\hat{T}_{pi}}{T_p} - 1\right) \text{ where p = 0.95 or 0.99} \tag{22}$$

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\hat{T}_{pi}}{T_p} - 1\right)^2\right]^{\frac{1}{2}} \tag{23}$$

To facilitate comparisons, the results of the simulation are presented graphically. Figures 1a, 1b, 2a and 2b show the relative efficiency (RE) of the estimates of parameters by each of methods 1-6, for different values of the shape parameter. As expected, the efficiency of each method for estimation for either parameter increases with sample size. Unexpectedly, in a few cases the efficiency exceeds 1. The graphs of the product of the root mean square relative error (RMSE) and β, against β is shown in Figures 3a, 3b, 4a and 4b. This product is used because it varies over a smaller range than does the RMSE.

For estimating the shape parameter, it is evident from Figures 1a and 1b that the GLS estimator is considerably more efficient than the other methods at both sample sizes, that efficiency increases with sample size. For most methods the efficiency is independent of the shape parameter, the exception being PWM. It is noteworthy that in both Figures the PWM estimator had low RE at $\beta = 0.5$, but high RE at $\beta = 1$ and at $\beta = 2$. Generally speaking, the second best method for estimating β is the OSM.

The RE of α is less clear cut, as seen in Figures 1a and 1b. For $2 < \beta < 30$ the GLS method is better than the other five methods, and the TQT method is worse than the other methods. For $\beta < 2$, and n = 10, OSM and TQT are the most efficient methods, but when $\beta < 2$ and n = 25, TQT is best, followed by ML.

For estimating the sample quantiles, we seek the method that has smallest RMSE. From Figures 1c and 1d it is evident that unless $\beta < 1$, ML is the best method, followed closely by TQT. For $\beta < 1$ Figures 3a and 3b show that PWM is the best method for the 95% quantile, with OSM as a close competitor, while Figures 1e and 1f show that in this case, OSM is the best method for the 99% quantile. In all of the Figures 1g and 1h the worst method for all values of β is OLS.

*Corresponding author
F. A. Al-Baidhani,
Business Administration Department, Baghdad College of Economics Science University, Baghdad, Iraq
e-mail: fadhilajab@gmail.com

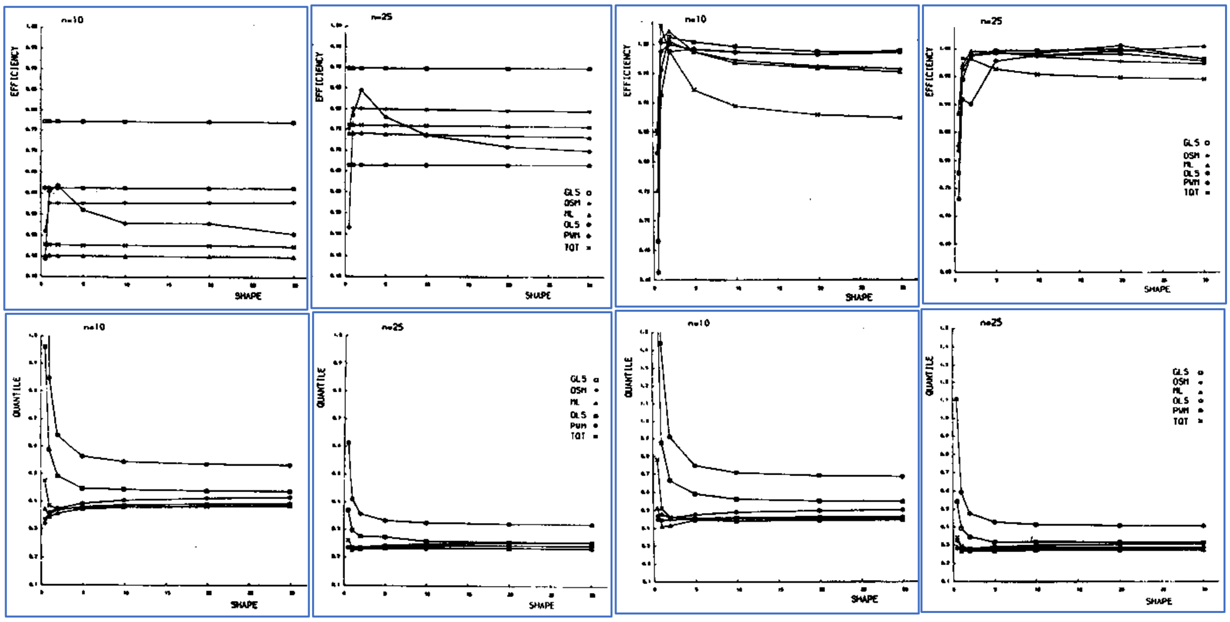

Figure 1: Empirical efficiency of estimates of shape parameter by various methods

CONCLUSION

PWM, OSM, TQT, ML, GLS, and OLS are compared using the relative efficiency of estimations of the shape and scale of alpha Weibull parameters, as well as the root mean square relative error of the projected 95 percent and 99 percent quantiles. The following conclusions are based on simulations with sample sizes of ten and twenty-five people. The optimal strategy to adopt clearly depends on the goal of the analysis as well as the value of. For estimating the shape, GLS is the best option. GLS is also the best method for determining the scale when the shape parameter is between 2 and 30. When the scale is equal to 2, TQT is an useful method for estimating it. When the shape parameter is greater than one, ML is the most accurate method for predicting the 95 percent and 99 percent confidence intervals.

REFERENCES

[1] Al-Baidhani, F.A. and Sinclair C.D. (1987). "Comparison of Methods of Estimation of Parameters of the Weibull Distribution". Communications in Statistics-Simulation and Computation, 16(2), 373-384.
[2] Barlow, R.E. and Campo, R. (1975). "Total Time on Test Processes and Applications to Failure Data Analysis". Reliability and Fault Tree Analysis, SIAM. Philadelphia 45-481.
[3] Dubey, S.D. (1967). "Some Percentile Estimators for Weibull Parameters". Technometrics, 9 (1), 119-129.
[4] Engelhardt, M. and Bain, L.J. (1977). "Simplified statistical Procedures for the Weibull or Extreme-Value Distribution". Technometrics,19 (3), 323-331.
[5] Epstein, L. and Sobel, M. (1953). "Life Testing". Journal of the American Statistical Association , 48, 486-502.
[6] Greenwood, J.A., Landwehr, J.M., Matalas, N.C. and Wallis, J.R. (1979). "Probability Weighted moments: Definition and relation to parameters of several distributions Expressible in Inverse Form". Water Resources Research, 15 (5), 1049-1054.
[7] Hosking, J.R.M., Wallis, J.R. and Wood, E.F. (1985). "Estimation of the Generalized Extreme-value Distribution by the Method of Probability-Weighted Moments". Technometrics , 27 (3), 251-261.
[8] Jewell, W.S. (1977). "Bayesian Life testing Using the Total Q on Test". In the Theory and Applications of Reliability , 1, 49-66.
[9] Shapiro, S.S. and Brain, C.W. (1987). "W-Test for the Weibull Distribution". Commun. Statist.-Simula., 16 (1) 209-219.

*Corresponding author
F. A. Al-Baidhani,
Business Administration Department, Baghdad College of Economics Science University, Baghdad, Iraq
e-mail: fadhilajab@gmail.com