Pergamon

Appl. Math. Lett. Vol. 7, No. 5, pp. 53–56, 1994

0893-9659/94 $7.00 + 0.00

0893-9659(94)00070-0

Modified Weighted Least-Squares Estimators for the Three-Parameter Weibull Distribution

K. E. AHMAD
Mathematics Department, Faculty of Science
University of Assiut, Assiut, Egypt

(Received February 1993; revised and accepted May 1994)

Abstract—Maximum likelihood estimation for the parameters in the three-parameter Weibull distribution needs an initial value in the iteration scheme for likelihood equations. In certain cases, maximum likelihood estimation can break down for the three-parameter Weibull model, as no local maximum of the likelihood exists. To avoid these difficulties, we suggest a modification of the weighted least-squares estimators for the three-parameter Weibull distribution. The modification presented here is to find an estimate for a threshold parameter for this distribution.

Keywords—Modified weighted least-squares, Three parameter Weibull, Maximum likelihood, Survival function, Goodness of fit.

1. INTRODUCTION

The Weibull distribution has been extensively used in life testing and reliability problems. The distribution has been named after the Swedish scientist, Weibull, who proposed it for the first time in 1939 in connection with his studies on strength of material. Weibull [1] showed that the distribution is also useful in describing the "wear-out" or fatigue failures. Lieblein and Zelen [2] used it as a model for ball bearing failures. Mann [3] gives a variety of situations in which the distribution is used for other types of failure data.

In some situations, it may be believed that there is a threshold stress level below which the material is safe, but above which there is a nonnegligible probability of failure. There is, therefore, an interest in estimating either the threshold level or some other quantile of the probability distribution of failure strengths. The Weibull distribution is a very widely used parametric family, partly because it is a flexible three-parameter family which has been found in practice to be suitable for data on failure strengths and also failure times. It is, therefore, an important practical problem to be able to estimate the parameters of this distributions being of particular interest. Mann [4] has reviewed the Weibull distribution in her case concentrating on the three-parameter case. Cohen and Whitten [5] proposed modifications of both moment and maximum likelihood estimators for the three-parameter Weibull distribution. Smith and Naylor [6] developed and compared maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution.

2. MODIFIED WEIGHTED LEAST-SQUARES ESTIMATORS

White [7] proposed a method for estimating the shape and scale parameters of the Weibull distribution, which is based on a regression approach. The method proposed by White is applicable to censored samples as well as complete samples. The objective of this article is to find

The author would like to thank the referees who read the text very carefully and made several good suggestions.

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX a modification of weighted least-squares estimators for the three-parameter Weibull distribution. The statistical model considered in this article is that of a population containing N items, which are subjected to the life testing. Assume that the failure times T of all items in the population are independently identically distributed three-parameter Weibull random variables having survival function:

$$R(t) = \exp\left[-\left(\frac{t-\alpha}{\delta}\right)^{\gamma}\right], \qquad \text{for } t \in (0,\infty). \tag{1}$$

The first task is to find an estimate for a threshold parameter α.

An approximate value of α may be found as follows.

Since the population is finite, we can compute directly for each point t_i the survival function

$$R_i(t) = \frac{\text{number of failures up to time } t_i}{N},$$

for $i = 1, 2, \dots, N$. (If only a sample is available, then R_i can be the empirical survival function and N is the sample size.)

Plot the points $(t_i,\ Z_i = \log_e R_i(t))$, $i = 1, 2, \dots, N$ on ordinary cross-section paper and draw a smooth good fitting curve. Let $(t_i,\ Z_i = \log_e R_i(t))$, $i = 1, 2, \dots, M$ be arbitrary points on the curve so that

$$R_i(t) = \exp\left[-\left(\frac{t_i-\alpha}{\delta}\right)^{\gamma}\right], \qquad i = 1, 2, \dots, M.$$

Cheng and Fu [8] discussed choosing M, where they pointed out that the weighted least-squares estimators are relatively robust with respect to M. They also noted that, in practice, if we choose the number of intervals M between 10 and 20 (with equal lengths), then the weighted least-squares estimators will normally perform well.

Choose four points so that $Z_1/Z_2 = Z_3/Z_4$. Then, $(t_1-\alpha)/(t_2-\alpha) = (t_3-\alpha)/(t_4-\alpha)$ and hence,

$$\alpha = \frac{t_2t_3 - t_1t_4}{t_2 + t_3 - (t_1 + t_4)}.$$

If several such thresholds are chosen, then the average of all α's calculated from these thresholds will give a fairly precise results.

Consider $R_i(t) = \exp\left[-\left(\frac{t_i-\hat{\alpha}}{\delta}\right)^{\gamma}\right]$, $i = 1, 2, \dots, M$. Then,

$$\log_e \log_e R_i^{-1}(t) = -\gamma \log_e \delta + \gamma \log_e (t_i - \hat{\alpha}). \tag{2}$$

The above equation (2) can be written in the form

$$Y_i = A + BX_i + \epsilon_i, \qquad i = 1, 2, \dots, M, \tag{3}$$

where $X_i = \log_e (t_i - \hat{\alpha})$, $Y_i = \log_e \log_e R_i^{-1}(t)$, $A = -\gamma \log_e \delta$ and $B = \gamma$.

Now, for $i = 1, 2, \dots, M$, let

$$Y_i = \log_e \log_e \left(\frac{N}{N - \sum_{k=1}^{i} U_k}\right), \tag{4}$$

where U_i = observed number of failures during the time (t_{i-1}, t_i), $i = 1, 2, \dots, M$.

M = total number of time intervals.

N = the size of the sample.

If we consider the weighted sum of squares

$$Q = \sum_{i=1}^{M} U_i (Y_i - A - BX_i)^2,$$

with the weights $\{U_i\}$, $i = 1, 2, \ldots, M$, then equation (3) reduces to a simple linear regression problem and the weighted least-squares estimators for A and B are, therefore,

$$\hat{B} = \frac{\left[\sum_{i=1}^{M} U_i X_i Y_i - \frac{1}{U}\left(\sum_{i=1}^{M} U_i Y_i\right)\left(\sum_{i=1}^{M} U_i X_i\right)\right]}{\left[\sum_{i=1}^{M} U_i X_i^2 - \frac{1}{U}\left(\sum_{i=1}^{M} U_i X_i\right)^2\right]}$$

and

$$\hat{A} = \frac{1}{U}\left(\sum_{i=1}^{M} U_i Y_i - \hat{B}\sum_{i=1}^{M} U_i X_i\right),$$

where $U = \sum_{i=1}^{M} U_i$. Finally, $\hat{\gamma}$ and $\hat{\delta}$ can be estimated by $\hat{\gamma} = \hat{B}$ and $\hat{\delta} = \exp\left(\hat{A}/\hat{B}\right)$.

3. NUMERICAL EXAMPLE

Dumonceaux and Antle [9] cite data, obtained in a civil engineering context, of maximum flood levels (in millions of cubic feet per second) for the Susquehanna River at Harrisburg, Pennsylvania over 20 four-year periods as

.654	.613	.315	.449	.297	.402	.379	.423	.379	.3235
.269	.740	.418	.412	.494	.416	.338	.392	.484	.265

writing the above data in grouped data and estimating $R(t)$, using the above rule, we have

Interval	U	$R_i(t)$	$Z = \log_e R_i(t)$
$.250 \le t < .300$	3	0.85	−0.163
$.300 \le t < .350$	3	0.70	−0.357
$.350 \le t < .400$	3	0.55	−0.598
$.400 \le t < .450$	6	0.25	−1.386
$.450 \le t < .500$	2	0.15	−1.897
$.500 \le t < .550$	0	0.15	−1.897
$.550 \le t < .600$	0	0.10	−2.302
$.600 \le t < .650$	1	0.10	−2.302
$.650 \le t < .700$	1	0.05	−2.996
$.700 \le t < .750$	1		

After plotting the points (t_i, Z_i), $i = 1, 2, \ldots, M$ and drawing a smooth good fitting curve, choosing $Z_1 = -.163$, $Z_2 = -.357$, $Z_3 = -1.386$, $Z_4 = -3.036$, so that $Z_1/Z_2 = Z_3/Z_4$, hence, we have $t_1 = 0.3$, $t_2 = 0.35$, $t_3 = 0.45$ and $t_4 = 0.7$. So, we get $\hat{\alpha} = 0.2625$. Therewith, using the weighted least-squares estimation, we obtain $\hat{\gamma} = 1.28$ and $\hat{\delta} = 0.167$.

On the other hand, Cheng and Amin [10] give maximum likelihood estimates for the three-parameter Weibull model, using Dumonceaux and Antle [9] data, in the form $\hat{\alpha} = 0.261$, $\hat{\gamma} = 1.25$ and $\hat{\delta} = 0.173$. We give some idea of the goodness of fit by comparing the empirical *cdf* and *cdf*'s of the fitted model using MLE (maximum likelihood estimators) and MWLSE (modified weighted least-squares estimators), shown in Figure 1. The *cdf*'s of the fitted model using MLE and MWLSE are very similar as seen in Figure 1 and they provide a reasonable fit compared with the sample *cdf*.

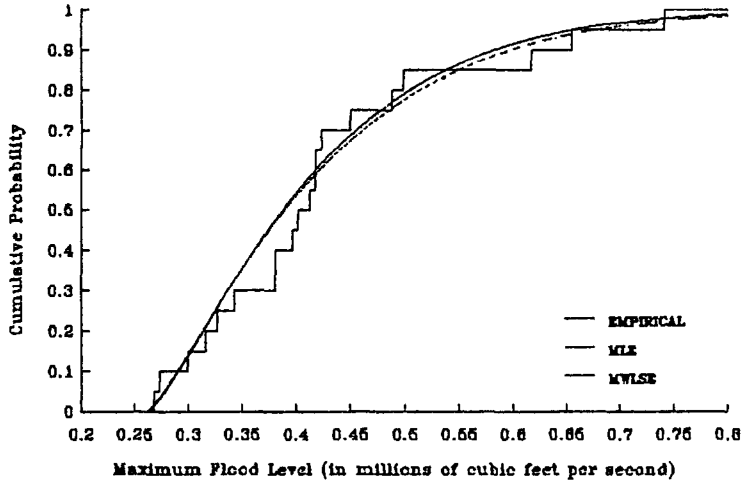

Figure 1. Empirical and fitted model using MLE and MWLSE.

4. CONCLUSIONS

In this paper, we suggest a modification of weighted least-squares estimators for the three-parameter Weibull distribution, the modification presented here is to find an estimate for a threshold parameter for this distribution. The advantage of this method is that it is suitable for the cases in which maximum likelihood estimation can break down for the three-parameter Weibull model, Cheng and Amin [10] gave some examples for certain cases in which maximum likelihood estimation breaks down for the Weibull model as no local maximum of the likelihood exists. Moreover, the suggested modification of weighted least-squares estimators for the three-parameter Weibull model can be used as good initial estimates for the maximum likelihood method.

REFERENCES

1. W. Weibull, A statistical distribution function of wide applicability, *J. Appl. Mech.* **11**, 293–297 (1951).
2. J. Lieblein and M. Zelen, Statistical investigation of the fatigue life of deep groove ball bearing, *J. of Res. Natl. Bur. Std.* **57**, 273–315 (1956).
3. N. Mann, Results on statistical estimation and hypotheses testing with application to the Weibull and extreme value distribution, Aerospace Research Laboratories, Wright-Patterson Air Force Base, OH, (1968).
4. N. Mann, Statistical estimation of parameters of the Weibull and Frechet distributions, In *Statistical Extremes and Applications* (Edited by J. Tiago de Oliveira), pp. 81–89, Reidel, Dordrecht, (1984).
5. A.C. Cohen and B.J. Whitten, Modified maximum likelihood and modified moment estimators for the three-parameter Weibull distribution, *Communication in Statistics Theory and Methods* **11** (23), 2631–2656 (1982).
6. R.L. Smith and J.C. Naylor, A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution, *App. Statist.* **36** (3), 358–369 (1987).
7. J.S. White, The moments of log-Weibull order statistics, *Technometrics* **11**, 373–386 (1969).
8. S.W. Cheng and J.C. Fu, Estimation of mixed Weibull parameters in life testing, *IEEE Trans. Rel.* **R-31** (4), 377–381 (1982).
9. R. Dumonceaux and C.E. Antle, Discrimination between the lognormal and Weibull distributions, *Technometrics* **15** (3), 923–926 (1973).
10. R.C.H. Cheng and N.A.K. Amin, Maximum likelihood estimation of parameters in the Inverse Gaussian distributions with unknown origin, *Technometrics* **23** (3), 257–263 (1981).