Appeared in: Journal of Statistical Computation and Simulation, 2003, 73, 115-123.

EFFICIENT ESTIMATION OF THE WEIBULL SHAPE PARAMETER BASED ON A MODIFIED PROFILE LIKELIHOOD

ZHENLIN YANG[a] and MIN XIE[b]

[a]*School of Economics and Social Sciences, Singapore Management University*
email: zlyang@smu.edu.sg
[b]*Department of Industrial and Systems Engineering, National University of Singapore*

The maximum likelihood estimator of the Weibull shape parameter can be very biased. An estimator based on the modified profile likelihood is proposed and its properties are studied. It is shown that the new estimator is almost unbiased with relative bias being less than 1% in most of situations, and it is much more efficient than the regular MLE. The smaller the sample or the heavier of the censoring, the more efficient is the new estimator relative to the regular MLE.

Keywords: Weibull distribution, Shape parameter, Bias, Mean squared error, Modified profile likelihood, Censored data.

1. INTRODUCTION

Because of its flexibility in modeling both increasing failure rate and decreasing failure rate, Weibull distribution is now widely used in reliability studies. As the failure rate trend for Weibull distribution is characterized by the value of its shape parameter, the estimation of the Weibull shape parameter is of particular interest. Several methods exist in the literature, such as the maximum likelihood estimation (MLE) method (Lawless, 1982), linear estimator (Lawless, 1982), method based on probability plot and a modified version of it (Drapella and Kosznik, 1999), shrunken estimator (Pandey and Singh), etc.

Among the above-mentioned methods, the MLE is a very popular one due to its simplicity and efficiency (Ross, 1994). The Weibull probability plot is usually used to get some rough estimates that might serve as starting values for numerical procedures in solving the likelihood equation. However, the MLE of the Weibull shape parameter is known to be biased, and can very biased in the case of small sample or heavy censoring (Mackisack and Stillman, 1996) (e.g, for a complete sample of size 10, the relative bias is about 21%). Ross (1994) provided a simple unbiasing formula for the case of complete sample, which was shown to work quite well, and added another formula (Ross, 1996) for the case of Type II censored data.

In this paper, we propose an estimator for the Weibull shape parameter based on modified profile likelihood of Cox and Reid (1987, 1989, 1992) under the notion of parameter orthogonalization. It turns out that the modification is extremely simple – it involves only a simple adjustment on the profile likelihood or the likelihood equation. The modified MLE performs surprisingly well as compared with the regular MLE. The relative bias can be reduced to less than 1% in most cases. Also, the modified MLE is much more efficient than the regular MLE. It is also more efficient than the unbiased MLE of Ross (1994, 1996). The modified MLE works the best for the complete or Type II censored data.

This paper is organized as follows. Section 2 discusses the use of parameter orthogonalization for the estimation of the Weibull parameters that provides the basic motivation for this study. Section 3 derives the formulas for the profile likelihood estimates for the Weibull shape parameter. Section 4 presents some simulation results regarding the performance of the new estimator relative to the regular likelihood estimator and its adjusted version.

2. PARAMETER ORTHOGONOLIZATION

A problem with the MLE for Weibull parameters is that the estimators are highly correlated. The basic idea behind the parameter orthogonalization is that if the two parameters are orthogonal, then the MLEs of the two parameters are asymptotically independent. Hence, making inference on one parameter is not affected (at least asymptotically) by whether the other parameter is estimated or given. Let Y be a Weibull random variable with the probability density function (pdf)

$$f(y;\, a,\, b) = \left(\frac{b}{a}\right)\left(\frac{y}{a}\right)^{b-1} \exp\left\{-\left(\frac{y}{a}\right)^{b}\right\}, \tag{1}$$

Appeared in: Journal of Statistical Computation and Simulation, 2003, 73, 115-123.

where b is the shape parameter and a is the scale parameter. We are now interested in estimating b in the presence of the nuisance parameter a.

Suppose that a reparameterization is made from (b, a) to (b, l) so that b and l are orthogonal in the sense that the element of the expected Fisher information matrix $I_{bl} = 0$ (Cox and Reid, 1987). The orthogonality condition can be reexpressed in terms of the original parameter setting as follows

$$i_{aa} \frac{\partial a}{\partial b} + i_{ba} = 0,$$

where $i_{aa} = E[-\partial^2 f(b,a)/\partial a^2] = (b/a)^2$, $i_{ba} = E[-\partial^2 f(b,a)/\partial b \partial a] = (g-1)/a$, and $g = 0.577215 \ldots$ is Euler's constant. Solving the resulted differential equation one gets $a = l \exp[(g-1)/b]$ and hence the orthogonal nuisance parameter:

$$l = a \exp[(1-g)/b].$$

The above result is given in Cox and Reid (1987), but there is a typographical error for the expression of i_{ba}. This orthogonal parameter setting was used to give a modified profile likelihood and hence an estimator of the Weibull shape parameter, but simulation results show that it is not quite satisfactory although it is able to reduce the bias by 50% or more, depending the true value of the shape parameter b.

It was noted in Cox and Reid (1989) that if l is orthogonal to b, so is any smooth function of l. This suggests that a further improvement is possible. Following the method of Cox and Reid (1989), we show that the optimal orthogonal parameterization takes the log form, i.e.

$$l^{\circ} \propto \log(l) = \log(a) + (1-g)/b.$$

However, similar to the example 3 considered in Cox and Reid (1989), the proportionality constant involves b, which is a phenomenon not being able to be explained by the original authors. Nevertheless, this leads us with a flexible choice of this constant.

Another important point is that the derivation of the orthogonal parameter is based on the complete sample. It would be of interest to investigate this orthogonal parameterization for the case of censored data. Naturally, this orthogonal parameterization should be reserved under censoring, but it is difficult to verify as in the case of censored data one cannot work out exact expressions for certain expectations. We will take primarily the log form, but allow some flexibility in the form of the b component with regard to different types of censoring.

3. MODIFIED PROFILE LIKELIHOOD

The purpose of modifications to the profile likelihood is to approximate more closely the likelihood function used in 'exact' marginal or conditional inference (Cox and Reid, 1992). In terms of the likelihood equation, the modified version should provide us with estimates that are much less biased than those corresponding to the profile likelihood. With these in mind, we investigate the usefulness of the method in the context of estimating the Weibull shape parameter.

Let $Y_1, Y_2, \ldots, Y_k$ be a random sample from WB(a , b) and $\ell(b, a)$ be the log likelihood function. The profile likelihood for b is defined as

$$\ell_p(b) = \ell[b, \hat{a}(b)],$$

where $\hat{a}(b)$ is the restricted MLE of a for a given b . The modified profile likelihood [1] is as follows

$$\ell_m(b) = \ell_p(b) - \frac{1}{2}\log\det\left[J_{ll}\left(b, \hat{l}(b)\right)\right], \tag{2}$$

where $J_{ll}\left(b, \hat{l}(b)\right)$ is the element of the observed information matrix for (b , l), evaluated for fixed b at the corresponding restricted MLE $\hat{l}(b)$.

3.1. The Case of Complete Data

Here, we give a detailed description of the method for the case of complete sample. The case of censored will be discussed later for different censoring assumptions. Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from WB(a , b). The log likelihood function is

$$\ell(b, a) = n\log\left(\frac{b}{a}\right) + (b-1)\sum_{i=1}^{n}\log\left(\frac{y_i}{a}\right) - \sum_{i=1}^{n}\left(\frac{y_i}{a}\right)^b \tag{3}$$

For a given b, the restricted MLE of a is

$$\hat{a}(b) = \left(\sum_{i=1}^{n} y_i^b / n\right)^{1/b}.$$

Substituting $\hat{a}(b)$ into (3) gives the profile log likelihood for b:

$$\ell_p(b) = n(\log n - 1) + n\log(b) - n\log\left(\sum_{i=1}^{n} y_i^b\right) + (b-1)\sum_{i=1}^{n}\log y_i .$$

The usual MLE of b is obtained by maximizing the profile log likelihood, or equivalently by solving the profile likelihood equation

$$S_p(b) = d\ell_p(b)/db = 0,$$

where the profile score function is

$$S_p(b) = \frac{n}{b} - n\frac{\sum_{i=1}^{n} y_i^b \log y_i}{\sum_{i=1}^{n} y_i^b} + \sum_{i=1}^{n}\log y_i . \tag{4}$$

By taking the parameterization

$$l^\circ = \frac{\log(l)}{b} = \frac{\log(a)}{b} + \frac{1-g}{b^2}, \tag{5}$$

and using the relationship,

$$J_{l^\circ l^\circ}[b, \hat{l}^\circ(b)] = J_{aa}(b, \hat{a}(b))\left(\partial a/\partial l^\circ\right)^2\Big|_{a=\hat{a}(b)},$$

one can easily see that $J_{l^\circ l^\circ}[b, \hat{l}^\circ(b)] \propto b^4$. Hence, from Equation (2), we have that

$$\ell_m(b) \propto \ell_p(b) - 2\log b$$

and the modified profile likelihood equation becomes,

$$S_m(b) = \frac{n-2}{b} - n\frac{\sum_{i=1}^{n} y_i^b \log y_i}{\sum_{i=1}^{n} y_i^b} + \sum_{i=1}^{n}\log y_i . \tag{6}$$

Appeared in: Journal of Statistical Computation and Simulation, 2003, 73, 115-123.

Solving the modified profile likelihood equation $S_m(b) = 0$ gives the modified MLE for b. This modification is simple, but works surprisingly well as shown by the Monte Carlo simulations next section.

3.2. The Case of Censored Data

For the case of censored data, it is reasonable to use the same orthogonal parameter setting. Only extra work is to derive, under different censorship mechanism, the observed information number $j_{ll}(b, \hat{l}(b))$, the key quantity in the modification term.

Let $Y_1, Y_2, \ldots, Y_r$ now denote the r smallest observations in a random sample of n from WB(b, a), i.e., the data are Type II censored. The observed information is found to be of the same form as in the case of complete sample if the same parameterization is used. This gives the modified the modified profile score function for b:

$$S_m(b) = \frac{r-2}{b} - \left(r \sum_{i=1}^{r}{}^{*} y_i^b \log y_i \right) \left(\sum_{i=1}^{r}{}^{*} y_i^b \right)^{-1} + \sum_{i=1}^{r} \log y_i \,, \tag{8}$$

compared with the usual profile score function

$$S_p(b) = \frac{r}{b} - \left(r \sum_{i=1}^{r}{}^{*} y_i^b \log y_i \right) \left(\sum_{i=1}^{r}{}^{*} y_i^b \right)^{-1} + \sum_{i=1}^{r} \log y_i \,, \tag{9}$$

where a notational convention, $\sum_{i=1}^{r}{}^{*} w_i = \sum_{i=1}^{r} w_i + (n-r)w_r$, is used.

In the situation of Type I censoring, define T_i, $i = 1, 2, \ldots, n$, as the actual Weibull lifetime, C_i the censoring time, and $Y_i = \min(T_i, C_i)$. Thus, the data ($Y_1, Y_2, \ldots, Y_n$) now represent the Type I censored data with r of them the real lifetimes and $n - r$ of them the censored lifetimes. We assume that $r \geq 1$. In this case, we employ the orthogonal parameterization

$$l^{\circ} = \log(l) = \log(a) + (1-g)/b \,.$$

The observed information is $j_{ll}[b, \hat{l}^{\circ}(b)] \propto b^2$ and we have from (2) that $\ell_m(b) = \ell_p(b) - \log b$. The profile and modified profile scores are

$$S_p(b) = \frac{r}{b} - \left(r \sum_{i=1}^{n} y_i^b \log y_i \right) \left(\sum_{i=1}^{n} y_i^b \right)^{-1} + \sum_{i \in D} \log y_i \tag{10}$$

and

$$S_m(b) = \frac{r-1}{b} - \left(r \sum_{i=1}^{n} y_i^b \log y_i \right) \left(\sum_{i=1}^{n} y_i^b \right)^{-1} + \sum_{i \in D} \log y_i \tag{11}$$

respectively, where D is the set of uncensored observations and r is the total number of them.

4. MONTE CARLO SIMULATION

In this section, some simulation results are presented to compare modified MLE with the traditional one. The bias and MSE of the two estimators are simulated. The reported results include: the relative bias, $E(\hat{b} - b)/b$, and the relative efficiency of $\hat{b}_M$ to $\hat{b}$,

$$\text{REF} = \text{MSE}(\hat{b})/\text{MSE}(\hat{b}_M).$$

We first consider the case of complete or Type II censored data. Several sample sizes (n) and degrees of censorship (r) are considered. Ten different b values are used for each combination of n and r. The a value is fixed at 100 as the estimation of b for both methods is independent of a. Simulation results are summarized in Table 1.

The results in Table 1 indicate that the new estimator denoted as MMLE is almost unbiased with relative bias being less than 1% in most of situations. It is also much more efficient than the regular MLE. The smaller the sample or the heavier of the censoring, the more efficient is the new estimator relative to the regular MLE.

Ross (1994) proposed a simple bias-reduction factor $(n-2)/(n-0.68)$ for the case of complete data, and added another formula $\{1+\sqrt{n/r}[1.37/(r-1.92)]\}^{-1}$ for Type II censored data (Ross, 1996). It was shown by simulation that these simple factors work well and can reduce the relative bias to less than 0.3%. However, these two formulas do not match when $n = r$, and there is no simple formula available for other type of censored data.

It is easy to see that in the cases of complete and Type II censored data, our modified MLE is more efficient than the bias-reduced MLE of Ross. For examples, for $n = r = 10$, the relative efficiency of the bias-reduced MLE to the regular MLE is 1.36, compared with 1.64 for the modified MLE, for $n = r = 20$, the numbers are, respectively, 1.15 and 1.30, and for $n = 30$ and $r = 10$, the numbers are 1.67 and 1.90, respectively.

Table 1. Relative Bias (%) and Relative Efficiency of MLE and Modified MLE Complete or Type II Censored Sample

b	Relative Bias MLE	Relative Bias MMLE	REF	Relative Bias MLE	Relative Bias MMLE	REF	Relative Bias MLE	Relative Bias MMLE	REF	Relative Bias MLE	Relative Bias MMLE	REF
	(n=10, r=10)			(n=20, r=10)			(n=20, r=15)			(n=20, r=20)		
0.5	17.11	1.82	1.64	21.57	0.52	1.82	12.30	0.70	1.48	7.75	0.94	1.30
0.8	17.16	1.83	1.64	21.52	0.48	1.83	11.90	0.34	1.47	7.61	0.80	1.29
1.0	17.33	2.01	1.64	21.66	0.59	1.84	12.40	0.79	1.48	7.57	0.76	1.30
1.5	16.90	1.62	1.64	21.43	0.38	1.84	12.44	0.83	1.48	7.52	0.72	1.29
2.0	16.80	1.52	1.63	22.64	1.40	1.86	11.95	0.39	1.47	7.76	0.93	1.31
3.0	16.90	1.62	1.62	21.50	0.45	1.83	12.02	0.45	1.47	7.62	0.81	1.30
4.0	16.94	1.64	1.64	21.47	0.41	1.85	12.36	0.74	1.48	7.50	0.71	1.29
5.0	16.92	1.66	1.62	22.21	1.03	1.84	12.28	0.67	1.48	7.56	0.76	1.29
8.0	16.25	1.05	1.61	21.77	0.68	1.83	11.79	0.24	1.46	7.83	1.01	1.30
10.0	16.40	1.22	1.62	21.79	0.70	1.85	12.00	0.42	1.47	7.38	0.59	1.29
	(n=30, r=10)			(n=30, r=15)			(n=30, r=20)			(n=30, r=25)		
0.5	23.04	0.53	1.90	13.08	0.06	1.52	9.11	0.33	1.36	6.63	0.36	1.26
0.8	22.26	-0.11	1.88	13.54	0.46	1.53	9.11	0.32	1.36	6.93	0.65	1.27
1.0	22.59	0.16	1.89	13.26	0.21	1.52	8.91	0.14	1.36	6.53	0.27	1.25
1.5	23.86	1.21	1.89	13.19	0.16	1.52	8.87	0.10	1.35	7.14	0.85	1.28
2.0	23.02	0.51	1.88	13.55	0.48	1.54	9.17	0.38	1.36	6.94	0.66	1.27
3.0	23.06	0.54	1.91	12.81	-0.17	1.51	9.16	0.38	1.36	6.65	0.39	1.26
4.0	22.77	0.31	1.91	13.76	0.66	1.55	8.88	0.13	1.36	6.81	0.54	1.27
5.0	22.63	0.19	1.89	13.47	0.40	1.53	8.73	-0.02	1.35	6.45	0.20	1.26
8.0	23.28	0.72	1.92	13.76	0.66	1.54	9.18	0.40	1.36	6.65	0.38	1.26
10.0	23.65	1.03	1.91	13.45	0.39	1.54	8.85	0.09	1.35	6.76	0.49	1.26

For the Type I censored case, we take the simple situation that the censoring time is constant across the observations. This is the case that all the testing units are put on test at the same time, and the test terminates at time C. The simulation is run at several degrees of censorship represented by p_0, the proportion of censoring. Fewer b values are considered relative to the complete or Type II censored case since the behavior of the estimators is quite stable with respect to the b value. The a value is again fixed at 100 since the estimation of b is independent of the a value. The simulation results are summarized in Table 2.

The results show that the modified MLE again performs much better than the regular MLE in terms of biasness and efficiency, although not as good as the cases of complete or Type II censored data. The bias of the MLE increases with p_0, the proportion of censoring, but the modified MLE does not. The relative efficiency of the modified MLE over the regular MLE increases with p_0 as well.

Table 2. Relative Bias (%) and Relative Efficiency of MLE and Modified MLE Type I Censored Data.

p_0	b	Relative Bias MLE	Relative Bias MMLE	REF	p_0	b	Relative Bias MLE	Relative Bias MMLE	REF
		(n = 20)					(n = 30)		
0.10	0.5	6.51	2.44	1.13	0.10	0.5	3.85	1.20	1.08
0.10	1.0	6.26	2.20	1.13	0.10	1.0	3.79	1.14	1.08
0.10	2.0	6.36	2.29	1.14	0.10	2.0	3.66	1.01	1.08
0.25	0.5	6.49	1.01	1.15	0.25	0.5	4.14	0.58	1.10
0.25	1.0	6.84	1.34	1.16	0.25	1.0	4.32	0.76	1.10
0.25	2.0	7.00	1.52	1.16	0.25	2.0	4.18	0.63	1.10
0.50	0.5	10.87	0.81	1.31	0.50	0.5	6.98	0.68	1.19
0.50	1.0	10.37	0.35	1.30	0.50	1.0	6.93	0.57	1.19
0.50	2.0	10.88	0.79	1.32	0.50	2.0	7.29	0.95	1.19
0.75	*	*	*	*	0.65	0.5	10.41	0.13	1.33
0.75	*	*	*	*	0.65	1.0	10.67	0.31	1.35
0.75	*	*	*	*	0.65	2.0	10.56	0.26	1.30
		(n = 100)					(n = 50)		
0.25	0.5	1.31	0.28	1.03	0.10	0.5	2.34	0.78	1.05
0.25	1.0	1.01	-0.03	1.03	0.10	1.0	2.36	0.79	1.05
0.25	2.0	1.28	0.24	1.03	0.10	2.0	2.38	0.81	1.05
0.50	0.5	2.09	0.32	1.06	0.25	0.5	2.41	0.31	1.06
0.50	1.0	1.99	0.21	1.05	0.25	1.0	2.43	0.34	1.06
0.50	2.0	1.85	0.09	1.05	0.25	2.0	2.76	0.66	1.07
0.75	0.5	4.13	0.11	1.12	0.50	0.5	3.87	0.23	1.11
0.75	1.0	4.12	0.09	1.12	0.50	1.0	3.60	-0.02	1.10
0.75	2.0	4.07	0.07	1.12	0.50	2.0	4.23	0.58	1.11
0.90	0.5	12.56	0.22	1.41	0.75	0.5	9.00	0.27	1.28
0.90	1.0	12.81	0.43	1.41	0.75	1.0	8.89	0.12	1.29
0.90	2.0	12.58	0.28	1.36	0.75	2.0	9.45	0.73	1.25

Note that all the simulations are carried out using Fortran 90, where an IMSL subroutine UVMID is used for solving the likelihood equations. For the purpose of real data analysis, it may be more convenient to use Mathematica or Maple to do the job. The Fortran code and the Mathematica code are available from the fist author upon request.

5. CONCLUSIONS

The traditional MLEs for the Weibull parameters are highly biased. This is especially so for the shape parameter, which is a very important parameter in reliability decision making and planning such as burn-in and replacement time determination. This problem is basically caused by the fact that the estimators are highly correlated. Through parameter orthogonalization, an inference procedure is developed in this paper.

Specifically, for the Weibull shape parameter, a modified profile likelihood under the notion of parameter orthogonalization is studied. The modification involves only a simple adjustment to the likelihood equation. Compared with the traditional MLE, the modified MLE performs surprisingly well. The modified MLE not only reduces the bias significantly, but is also more efficient than both the regular MLE and the unbiased MLE of Ross (1994, 1996).

Finally, it should be mentioned that a good estimator for the shape parameter is crucial in obtaining a good estimator for the scale parameter as the latter is often a function of the former and an estimate of the scale parameter can be computed when the shape parameter is estimated. Also, when the shape parameter is estimated, the distribution can be transformed to exponential with a simple power transformation and statistical inferences can be carried out easily (Xie, Yang and Gaudoin, 2000).

REFERENCES

D.R. Cox and N. Reid, "Parameter orthogonality and approximate conditional inference (with discussion)", *J. R. Statist. Soc.* B, vol 49, 1987, pp 1-30.

D.R. Cox and N. Reid, "On the stability of maximum-likelihood estimators of orthogonal parameters", *Canadian J. Statist.*, vol 17, 1989, pp 229-233.

D.R. Cox and N. Reid, "A note on the difference between profile and modified profile likelihood", *Biometrika,* vol 79, 1992, pp 408-411.

A. Drapella and S. Kosznik, "An alternative rule for placement of empirical points on Weibull probability paper", *Quality and Reliab. Eng. Int.*, vol 15, 1999, pp 57-59.

J.F. Lawless, *Statistical Models and Methods for Lifetime Data,* 1982; John Wiley & Sons, New York.

M.S. Mackisack and R.H. Stillman, "A cautionary tale about Weibull analysis", *IEEE Trans. Reliability,* vol 45, 1996, pp 244 -248.

M. Pandey and U.S. Singh, "Shrunken estimators of Weibull shape parameter from Type II censored samples", *IEEE Trans. Reliability*, vol 42, 1993, pp 81-86.

R. Ross, "Formulas to describe the bias and standard deviation of the ML-estimated Weibull shape parameter", *IEEE Trans. Dielectrics and Electrical Insulation,* vol 1, 1994, pp 247-253.

R. Ross, "Bias and standard deviation due to Weibull parameter estimation for small data sets", *IEEE Trans. Dielectrics and Electrical Insulation,* vol 3, 1996, pp 28-42.

M. Xie, Z. Yang and O. Gaudoin, "More on the mis-specification of the shape parameter with Weibull-to-exponential transformation", *Quality and Reliab. Eng. Int.*, vol. 16, 2000, pp 281-290.