9-2007

# Improved Maximum-Likelihood Estimation for the Common Shape Parameter of Several Weibull Populations

Zhenlin YANG
*Singapore Management University*, zlyang@smu.edu.sg

Dennis K. J. LIN
*Pennsylvania State University - Main Campus*

**DOI:** https://doi.org/10.1002/asmb.678

Appeared in: Applied Stochastic Models in Business and Industry, 2007, 23, 373-383.

# Improved Maximum Likelihood Estimation for the Common Shape Parameter of Several Weibull Populations

**Zhenlin Yang**
School of Economics and Social Sciences
Singapore Management University, email: zlyang@smu.edu.sg

**Dennis K. J. Lin**
Department of Supply Chain and Information Systems
The Pennsylvania State University, email: dkl5@psu.edu

**Summary**

The biasness problem of the maximum likelihood estimate (MLE) of the common shape parameter of several Weibull populations is examined in detail. A modified MLE (MMLE) approach is proposed. In the case of complete and Type II censored data, the bias of the MLE can be substantial. This is noticeable even when the sample size is large. Such a bias increases rapidly as the degree of censorship increases and as more populations are involved. The proposed MMLE, however, is nearly unbiased and much more efficient than the MLE, irrespective of the degree of censorship, the sample sizes, and the number of populations involved.

**Key Words:** Bias; Mean squared error; MLE; Modified MLE; Relative efficiency; Shape parameter; Type II censored data; Weibull distributions.

Appeared in: Applied Stochastic Models in Business and Industry, 2007, 23, 373-383.

# 1 Introduction

Equality of Weibull shape parameters across different groups of individuals is an important and common assumption in many applications. In regression problems with Weibull distributions, such an assumption is analogous to the constant variance assumption in normal regression models (Lawless, 1982, p.178). For example, lifetimes of manufactured items and breakdown voltages of electrical cable insulation are often assumed to follow Weibull distributions with a constant shape parameter where different manufacturing environments or different types of cable only alter the value of the scale parameter (Nelson, 1972; Stone and Lawless, 1979).

The most popular method of estimating the common shape parameter is the maximum likelihood method. Let $WB(\alpha, \beta)$ denote the Weibull population with cumulative distribution function $F(y, \alpha, \beta) = 1 - \exp\{-(y/\alpha)^\beta\}$, where $\alpha$ is the scale parameter and $\beta$ is the shape parameter. Let $t_{ij}(j = 1, \cdots, n_i)$ be the lifetimes and censoring times in the sample from the $i$th population $WB(\alpha_i, \beta)$ $(i = 1, \cdots, k)$, $r_i$ be the number of observed lifetimes in the $i$th sample, and $D_i$ be the set of individuals in the $i$th sample whose lifetimes are observed. Then the maximum likelihood estimator (MLE) $\hat{\beta}$ of the common shape parameter $\beta$ can be obtained by solving

$$\sum_{i=1}^{k} r_i \left( \frac{\sum_{j=1}^{n_i} t_{ij}^{\hat{\beta}} \log t_{ij}}{\sum_{j=1}^{n_i} t_{ij}^{\hat{\beta}}} \right) - \frac{\sum_{i=1}^{k} r_i}{\hat{\beta}} - \sum_{i=1}^{k} \sum_{j \in D_i} \log t_{ij} = 0. \tag{1}$$

See Lawless (1982, p183).

The MLE $\hat{\beta}$ is known to be biased (and sometimes significantly biased) when the sample sizes are small or when the data is heavily censored (Thoman *et al.* 1969). Such a biasness can mislead the subsequent inferences. In the case of a single random sample ($k = 1$), the biasness issue has been addressed by many authors (See, among the others, Bain and Engelhardt, 1991, p.221; Ross, 1994, 1996; Hirose, 1999; Yang and Xie, 2003; Ferrari *et al.*, 2007). However, the biasness issue in estimating the common shape parameter of several Weibull populations has not been addressed. First, it is not clear how biased the MLE of the common shape can be. Second, how can the MLE be corrected in a simple way to give a satisfactory estimator of the common shape?

In this paper, the biasness of the MLE is examined in detail and a simple modification on the profile likelihood equation (1) is introduced, based on the parameter orthorgonalization method of Cox and Reid (1987), to give a modified MLE (MMLE). It is found that for complete and Type II censored data the bias of the MLE can be substantial and remains noticeable even when sample sizes are fairly large. It increases rapidly as the degree of censorship increases and as the number of populations grows. The proposed MMLE, however, is nearly unbiased and much more efficient than the MLE, irrespective of the degree of censorship, the sample sizes, and the number of populations involved. For Type I censored data, the biasness problem of the MLE is less serious (as compared to the case of Type II censored data) and the improvement of the MMLE over the MLE is less significant. The computation for the MMLE is as simple as the computation for the MLE because the modification is simply to subtract a constant (depending on $k$) from the term $\sum r_i$ in (1). Our results generalize those of Yang and Xie (2003) for the special case of a single Weibull population, i.e., $k = 1$. Such a generalization is important as comparing several Weibull populations of the same shape is often of practical interest.

This paper is organized as follows. Section 2 derives the orthogonal parameters. The modified MLE is introduced in Section 3. Section 4 presents extensive simulation results for the properties of the MLE and MMLE. Two numerical examples are discussed in Section 5 for illustration. Concluding remarks and discussion are given in Section 6.

# 2 The Orthogonal Parameters

The biasness problem of the Weibull shape estimation is partly due to the fact that the estimators of the Weibull parameters are highly correlated. One way to alleviate the dependence of the parameter estimators is to reparameterize so that the parameters of interest and the nuisance parameters are orthogonal in the sense of Cox and Reid (1987). This way, inference on the parameters of interest is not affected (asymptotically) by the estimation of nuisance parameters. The impact of parameter orthgonalization is more significant when more nuisance parameters are involved.

Let $\alpha = \{\alpha_1, \cdots, \alpha_k\}$. Suppose that a reparameterization is made from $(\beta, \alpha)$ to $(\beta, \lambda)$. Denote by $(I_{\beta\beta}, I_{\beta\alpha'}, I_{\alpha\beta}, I_{\alpha\alpha'})$ and $(I_{\beta\beta}, I_{\beta\lambda'}, I_{\lambda\beta}, I_{\lambda\lambda'})$, respectively, the elements of the expected Fisher information matrix of $(\beta, \alpha)$ and $(\beta, \lambda)$. The $\beta$ and $\lambda$ are said to be orthogonal if $I_{\beta\lambda'} = 0$. It is often convenient to work with the original parameterization under which the orthogonality condition becomes,

$$I_{\alpha\alpha}\frac{\partial\alpha}{\partial\beta} + I_{\beta\alpha} = 0$$

where the $\alpha's$ are implicitly functions of $\beta$ and $\lambda$. To find the orthogonal parameters, it is necessary that both $I_{\alpha\alpha}$ and $I_{\beta\alpha}$ posses closed-form expressions so that partial differential equations can be set up and solved to give orthogonal parameters. This is clearly a difficult task when the likelihood involves censored data. The log likelihood based on complete samples has the form

$$\ell(\beta, \alpha) = m \log \beta + (\beta - 1)\sum_{i=1}^{k}\sum_{j=1}^{n_i} \log t_{ij} - \beta \sum_{i=1}^{k} n_i \log \alpha_i - \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\frac{t_{ij}}{\alpha_i}\right)^{\beta} \qquad (2)$$

where $m = \sum_{i=1}^{k} n_i$. It is easy to see that $I_{\alpha\alpha} = diag\{n_i\beta^2/\alpha_i^2, i = 1, \cdots, k\}$ and $I_{\beta\alpha} = \{-n_1(1-\gamma)/\alpha_1, \cdots, -n_k(1-\gamma)/\alpha_k\}'$ with $\gamma$ being the Euler's constant. Substituting these into the above condition leads to differential equations:

$$\left(\frac{\beta}{\alpha_i}\right)^2 \left(\frac{\partial\alpha_i}{\partial\beta}\right) - \frac{1-\gamma}{\alpha_i}, \qquad i = 1.\cdots, k.$$

with one set of solutions being:

$$\alpha_i(\beta, \lambda_i) = \lambda_i \exp\left(-\frac{1-\gamma}{\beta}\right), \qquad i = 1, \cdots, k.$$

These give the orthogonal parameters:

$$\lambda_i = \alpha_i \exp\left(\frac{1-\gamma}{\beta}\right), \qquad i = 1, \cdots, k.$$

Cox and Reid (1989) indicated that if $\lambda$ is orthogonal to $\beta$, so is any smooth function of $\lambda$. They suggested a basis for choosing the function so that the dependence between the parameter estimators is in the least informative fashion. Following their method, it is shown that the optimal orthogonal parameterization takes the log form, i.e.,

$$\lambda_i^o \propto \log \lambda_i = \log \alpha_i + (1-\gamma)/\beta, \; i = 1, \cdots, k. \tag{3}$$

The proportionality constant depends on $\beta$, a phenomenon similar to the Example 3 of Cox and Reid (1989). This implies that taking $\lambda_i^o = \log \lambda_i$ may not lead to the optimal function. Further improvement is possible by multiplying a $\beta$-dependent constant to $\log \lambda$. More on this issue is discussed next.

# 3 The Modified MLE

With the orthogonal parameters derived earlier, we are now ready to derive the modification to the likelihood equation (1). However, there is one difficult question: how is the orthogonal parameter setting for the censored data connected to that for the complete data? While the orthogonality condition depends on the expected Fisher information that is dependent on the type of data, a parameterization relates to only the intrinsic feature of the populations, hence should not be changed by the type of data. This leads us to the considerations of adopting the orthogonal parameters defined in (3) for general censored situation and making necessary adjustments based on numerical evidence. For the arbitrary censored data described in Section 1, the log likelihood is

$$\ell(\beta, \alpha) = m \log \beta + (\beta - 1) \sum_{i=1}^{k} \sum_{j \in D_i} \log t_{ij} - \beta \sum_{i=1}^{k} r_i \log \alpha_i - \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left( \frac{t_{ij}}{\alpha_i} \right)^{\beta} \tag{4}$$

where $m = \sum_{i=1}^{k} r_i$. For a given $\beta$, the restricted MLEs of $\alpha' s$ are

$$\hat{\alpha}_i(\beta) = \left( \frac{1}{r_i} \sum_{j=1}^{n_i} t_{ij}^{\beta} \right)^{1/\beta}, i = 1, \cdots, k.$$

Substituting these back into (4) gives the profile likelihood for $\beta$,

$$\begin{aligned}\ell_p(\beta) &= \ell(\beta, \hat{\alpha}_1(\beta), \cdots, \hat{\alpha}_k(\beta)) \\ &= m[\log\beta - 1] + (\beta - 1)\sum_{i=1}^{k}\sum_{j=1}^{n_i}\log t_{ij} - \beta\sum_{i=1}^{k} n_i \log\hat{\alpha}_i(\beta)\end{aligned}$$

and taking the derivative of $\ell_p(\beta)$ gives the profile likelihood equation (1). Under the orthogonal parameter setting $\lambda_i^o = \log\lambda_i$, the modified profile likelihood of Cox and Reid (1987) is defined as

$$\ell_m(\beta) = \ell_p(\beta) - \frac{1}{2}\log\det\left\{J_{\lambda^o\lambda^o}[\beta, \hat{\lambda}^o(\beta)]\right\} \tag{5}$$

where $J_{\lambda^o\lambda^o}[\beta, \hat{\lambda}^o(\beta)]$ is the element of the observed information matrix of the new parameterization $(\beta, \lambda^o)$ evaluated at the restricted MLE $\hat{\lambda}^o(\beta)$ for a given $\beta$. A simpler way for calculating this quantity is through the original parameterization:

$$J_{\lambda^o\lambda^o}[\beta, \hat{\lambda}^o(\beta)] = \left(\frac{\partial\alpha}{\partial\lambda^o}\right) J_{\alpha\alpha}(\beta, \alpha) \left(\frac{\partial\alpha}{\partial\lambda^o}\right)^T\Bigg|_{\alpha=\hat{\alpha}(\beta)}$$

where $J_{\alpha\alpha}(\beta, \alpha)$ is the element of the observed information matrix of $(\beta, \alpha)$, which is diagonal with the $i$th diagonal element being $\beta(\beta+1)\sum_{j=1}^{n_i} t_{ij}^{\beta}/\alpha_i^{\beta+2} - r_i\beta/\alpha_i^2$. This along with the expression (3) give the modification term:

$$-\frac{1}{2}\log\det\left\{J_{\lambda^o\lambda^o}[\beta, \hat{\lambda}^o(\beta)]\right\} \propto -k\log\beta.$$

Taking the derivative of $\ell_m(\beta)$ gives the modified likelihood equation:

$$\sum_{i=1}^{k} r_i \left(\frac{\sum_{j=1}^{n_i} t_{ij}^{\tilde{\beta}}\log t_{ij}}{\sum_{j=1}^{n_i} t_{ij}^{\tilde{\beta}}}\right) - \frac{\sum_{i=1}^{k} r_i - k}{\tilde{\beta}} - \sum_{i=1}^{k}\sum_{j\in D_i}\log t_{ij} = 0. \tag{6}$$

where $\tilde{\beta}$, the solution of (6), is the modified MLE (MMLE). Some explanations on the adjusting factor $k$ are as follows. In (6), $\sum_{i=1}^{k} r_i$ represents the total amount of information available and $k$ represents the total number of nuisance parameters to be estimated. When only the estimation of $\beta$ is concerned, its estimating equation should be penalized by the number of additional parameters estimated other than the parameters of interest. This explanation is consistent with the degrees of freedom reduction in the cases of $t$ and chi-squared tests.

There is a certain arbitrariness in choosing the constant in (3). It is possible that the modifier $k$ in (6) may not be the optimal choice. Also, the orthogonal parameters were derived from the complete samples and it is not clear how well it will work for censored data. Most importantly, the Cox-Reid method leads to the identification that it is necessary to modify $\sum r_i$ term in the likelihood equation by subtracting a constant from it. Empirical evidence provided in next section, however, reveals that in case of complete or Type II censored data, a modifier $k+1$ (instead of $k$) provides a dramatic improvement. The case of Type I censored data is more complicated than the case of Type II censored data, since $\sum r_i$ is no longer a fixed quantity. Monte Carlo simulation show that the modifier $k \sum r_i / \sum n_i$ works quite well.

In summary, the final modified likelihood equation takes the general form:

$$\sum_{i=1}^{k} r_i \left( \frac{\sum_{j=1}^{n_i} t_{ij}^{\tilde{\beta}} \log t_{ij}}{\sum_{j=1}^{n_i} t_{ij}^{\tilde{\beta}}} \right) - \frac{\sum_{i=1}^{k} r_i - c(k)}{\tilde{\beta}} - \sum_{i=1}^{k} \sum_{j \in D_i} \log t_{ij} = 0. \tag{7}$$

where $c(k) = k + 1$ for complete or Type II censored data and $c(k) = k \sum r_i / \sum n_i$ for Type I censored data.

# 4 Simulation Studies

A simulation study is carried out to assess the finite sample properties of the MLE and MMLE. Various values of $n$, $\beta$, $k$ and $p$ (proportion of non-censoring) are considered, allowing us to see the impact of sample size, population skewness, number of populations as well as the degree of censorship on the performance of the MLE and MMLE. For a given parameter setting, $k$ random samples are generated, one from each population, using IMSL subroutines RNWIB and SSCAL, and then censored. The MLE and MMLE of $\beta$ are computed and recorded. This procedure is repeated 10,000 times. The average and variance of the 10,000

Appeared in: Applied Stochastic Models in Business and Industry, 2007, 23, 373-383.

MLEs (or MMLEs) lead to Monte Carlo estimates of bias and mean-squared-error (MSE) of the two estimators. In Tables 1-2, $\mathrm{RB}_0$ and $\mathrm{RB}_1$ represent the relative bias (in percentage) of the MLE and MMLE, respectively, i.e., $\mathrm{RB}_0 = 100 \times (\hat{\beta} - \beta)/\beta$, and REF represents the relative efficiency of MMLE over MLE, that is, REF $= MSE(\hat{\beta})/MSE(\tilde{\beta})$.

**Complete and Type II censored data.** Comparing two Weibull populations of the same shape is of particular interest. From the simulation results reported in the upper portion of Table 1, we see that the MMLE is superior to the MLE. For all the cases simulated, the MMLE is nearly unbiased with the relative bias always less than 1%, but the relative bias of the MLE can be more than 30%; the MMLE can be as much as 151% (REF=2.51) more efficient than the MLE.

We next consider a more general case of estimating the common shape parameter for eight Weibull populations with different scale parameters. The simulation results are summarized in the lower portion of Table 1. The conclusions drawn from the two populations case still hold. Some further conclusions are as follows. First, with more populations involved ($n_i's$ unchanged), both estimators improved but the MMLE has a larger improvement over the MLE as reflected by the relative efficiency. For example, when $n_i = 20, i = 1, \cdots, k$, the REFs for the $k = 2$ case are smaller than the corresponding REFs for the $k = 8$ case. Second, when $n_i's$ decrease but $k$ increases (such that $\sum n_i$ increases), the performance of the MLE may get worse if the increase in $\sum n_i$ (the total number of observations) is not large enough to offset the increase in $k$ (the number of scale parameters). For example, from the case with $k = 2$ and $n_i = 20$ to the case with $k = 8$ and $n_i = 10$, the bias for the MLE at each censoring level becomes larger, despite of the fact that the total number of observations is increased by 40 (doubled) and the number of nuisance parameters is increased only by six.

*Table 1 here*

Furthermore, other parameter configurations have also been considered, including different values of $\alpha's$, more values for $k$, $n$ and $\beta$, different degrees of censorship for each sample, etc., and the results (not reported for brevity) are consistent with the patterns in Table 1. The simulation results further show that the performance of both estimators does not depend on the values of the scale parameters and depends very little on the true value of $\beta$. This is because both $(\hat{\beta} - \beta)/\beta$ and $(\tilde{\beta} - \beta)/\beta$ are invariant with respect to $\alpha$ and are asymptotically invariant with respect to $\beta$. The bias of MLE remains noticeable even when the sample sizes are large and it increases quickly as the degree of censorship increases. Moreover, the computation of the MMLE is as simple as that of the MLE.

**Type I censored data.** Type I censored data are more common in practice, but appear technically difficult for inferences (Nelson, 1982, p248). A similar phenomenon appears in the modification of the profile likelihood equation for the common Weibull shape. Monte Carlo simulation shows that using the modifier $c(k) = k \sum r_i / \sum n_i$ is far better than $k$. Table 2 presents some simulation results for $k = 2$ and 5 using the modifier $k \sum r_i / \sum n_i$. Unlike the cases of complete and Type II censored data where the MMLE offers a uniform and dramatic improvement over the MLE, the improvement is mild in the cases of Type I censored data. However, the biasness problem for the MLE based on Type I censored data is not as severe as the case of Type II censored data. In particular, the bias increases in a rather small magnitude with the increase of the degree of censorship, as compared with the case of Type II censored data. As a check on the stability of the simulation results with respect to the change in $\beta$ value, we report in Table 2 the results under three different values of $\beta$. Indeed, the results are quite stable.

*Table 2 here*

Other parameter configurations are simulated. All the results (not reported for brevity) are consistent with those reported in Table 2. It is generally concluded that the MMLE based on Type I censored data works well for light to moderate censored data in terms of bias reduction. But in terms of efficiency enhancement, it works well also for the heavy censoring case.

# 5 Numerical Examples

We now present two numerical examples to illustrate the use of the MLE and MMLE of the common shape parameter and their impact on the subsequent inferences such as estimating the scale parameters, the population means, the reliability, the percentile life, etc. In the discussions below, a quantity with a ˆ represents an estimator based on $\hat{\beta}$, and a quantity with a ˜ represents an estimator based on $\tilde{\beta}$.

**Example 1.** The data given below, analyzed by Lawless (1982, p189), are the failure voltages (in kilovolts per millimeter) of 40 specimens (20 each type) of electrical cable insulation: Type I Insulation: 32.0, 35.4, 36.2, 39.8, 41.2, 43.3, 45.5, 46.0, 46.2, 46.4, 46.5, 46.8, 47.3, 47.6, 49.2, 50.4, 50.9, 52.4, 56.3; Type II Insulation: 39.4, 45.3, 49.2, 49.4, 51.3, 52.0, 53.2, 53.2, 54.9, 55.5, 57.1, 57.2, 57.5, 59.2, 61.0, 62.4, 63.8, 64.3, 67.3, 67.7.

The MLEs and MMLEs based on individual sample assuming different shapes are: $\hat{\beta}_1 = 9.3833, \tilde{\beta}_1 = 8.8116, \hat{\beta}_2 = 9.1411, \tilde{\beta}_2 = 8.5783$, both supporting the assumption of equal shape. The MLE and MMLE for the assumed common shape parameter are $\hat{\beta} = 9.2611$ and $\tilde{\beta} = 8.8371$, with $\hat{\beta}$ being 4.8% larger than $\tilde{\beta}$, indicating that MLE might over-estimate the shape parameter. Calculations are also made by artificially censoring the data (e.g., taking the first 12 observations from each ordered sample), the results (available from the authors upon request) show a wider gap between MLE and MMLE, meaning that the bias of the MLE increases as the degree of censorship increases. This is consistent with the simulation results reported in the last section.

The MLEs and MMLEs of the two scale parameters are $\hat{\alpha} = (48.05,\ 59.54)$ and $\tilde{\alpha} = (47.79,\ 59.22)$, and of the two population means are $\hat{\mu} = (45.81,\ 56.78)$ and $\tilde{\mu} = (45.36,\ 56.20)$. Two methods give similar estimates of the scale parameters and the population means. However, as seen from Table 3, the MLEs and MMLEs of reliabilities and percentile lives are different, especially at the two tails of the distribution. In particular, the MLEs are larger than the MMLEs at left tail of the distribution, but smaller at the right tail.

*Table 3 here*

**Example 2.** The data given in McCool (1979) and displayed in Table 4 are the times of fatigue failure (in millions of cycles) of high-speed turbine engine bearings made out of five different compounds. The individual estimates of the shape parameters (see Table 4) do not show a wild difference among the estimated shape parameters of five compounds. Thus, one can assume a common shape parameter for different compounds, which is estimated to be 3.78 by MLE and 3.13 by MMLE. The gaps between MLE and MMLE are wider, as compared with the results in Example 1. The combined $\hat{\beta}$ is 20.8% larger than the combined $\tilde{\beta}$. Censoring the data artificially (i.e., taking first few observations in each ordered sample) further widens the gap. This may have a significant impact on the subsequent inferences.

The use of $\hat{\beta}$ leads to $\hat{\alpha}$ = (12.74, 7.80, 9.92, 14.38, 16.4) and $\hat{\mu}$ = (11.51, 7.05, 8.97, 12.99, 14.83), while the use of $\tilde{\beta}$ gives $\tilde{\alpha}$ = (12.40, 7.39, 9.69, 13.25, 16.08) and $\tilde{\mu}$ =(11.09, 6.61, 8.67, 11.86, 14.39). Thus, as a result of a larger gap between $\hat{\beta}$ and $\tilde{\beta}$ the gap between $\hat{\alpha}$ and $\tilde{\alpha}$ and the gap between $\hat{\mu}$ and $\tilde{\mu}$ become larger as compared with the corresponding results in Example 1.

Table 5 presents the estimates of reliabilities and percentile lives for the Type I bearings. Similar to Example 1, the two methods give quite different estimates on reliability and percentile life, especially at the two tails of the distribution.

*Tables 4 and 5 here*

# 6 Discussion

Accurate estimation of the Weibull shape parameter is a crucial engineering issue since the shape parameter determines the failure pattern. The usual MLE may over-estimate the shape by as much as 50%. This is especially true when sample size is small or the data is heavily (Type II) censored. Such a bias is clearly undesirable. In this paper, we proposed a new estimator of common Weibull shape, called the modified MLE or MMLE. It is seen that the MMLE is almost unbiased and is more efficient than the MLE. It is computationally as simple as the MLE, thus is highly recommended for the practical use. Our results generalize the results of Yang and Xie (2003) for the special case of a single Weibull population.

The large difference between the MLE and MMLE can give completely different conclusions about the failure mechanism. For example, if the true shape parameter is 0.9, meaning the failure rate is decreasing, its MLE could easily be 1.08 (20% over estimation), which indicates that failure rate is increasing. This is not likely to happen if the MMLE is used. Using the MLE or the MMLE of the common shape parameter can also give quite different estimates of reliabilities and percentile lives, especially at the two tails of the distribution as seen from the two examples given in the last section.

There are various related approaches available in the literature for the type of problems studied in this paper, including the modified profile likelihood approach (Barndorff-Nielson, 1983), the marginal or conditional likelihood approach (Fraser, 1968; Kalbfleisch and Sprott, 1970; Lawless and Mann, 1976). See also Barndorff-Nielsen (1994). It would be interesting, as a possible future work, to compare various available approaches when applied to the problem considered in this paper. In the special case of one population, Ferrari *et al.* (2007) show that the MMLE proposed by Yang and Xie (2003) outperforms the estimators based on the competing adjusted profile likelihoods.

Another commonly used method for estimating the Weibul shape parameter is the least squares estimation (LSE) method. The LSE is also biased and a bias-corrected LSE is proposed by Zhang *et al.* (2006). It would be interesting to first extend their bias-corrected LSE to the case of $k$ populations of the same shape, and then compare with our MMLE proposed in this paper.

## Acknowledgements

We would like to thank the Editor-in-Chief, Jef L. Teugels, an Associate Editor and a referee for the helpful comments.

# References

Barndorff-Nielsen, O. E. (1983). On a formula for the distribution of a maximum likelihood estimator. *Biometrika*, **70**, 343-365.

Barndorff-Nielsen, O. E. and Cox, D. R. (1994). *Inference and Asymptotics.* London: Chapman and Hall.

Bain, L. J. and Engelhardt, M. (1991). *Statistical Analysis of Reliability and Life-Testing Models.* New York: Marcel Dekker, Inc.

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

Cox, D. R. and Reid, N. (1989). On the stability of maximum-likelihood estimator of orthogonal parameters. *Canadian J. Statist.*, **17**, 229-233.

Ferrari, S. L. P., Ferreira da Silva, M. and Cribari-Neto, F. (2007). Adjusted profile likelihoods for the Weibull shape parameter. *J. Statist. Compu. & Simu.*, to appear.

Fraser, D. A. S. (1968). *The Structure of Inference.* New York: Wiley.

Hirose, H. (1999). Bias correction for the maximum likelihood estimates in two parameter Weibull distribution. *IEEE Trans. Dielect. Electr. Insul.*, **6**, 66-68.

Kalbfleisch, J. D. and Sprott, D. A. (1970). Applications of likelihood methods to models involving large numbers of parameters (with discussion). *J. Royal. Statist. Soc. B*, **32**, 175-208.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data.* New York: Wiley.

Lawless, J. F. and Mann, N. R. (1976). Tests for homogeneity of extreme value scale parameters. *Commum in Statistics - Theor Meth., A5*, 389-405.

McCool, J. I. (1979). Analysis of single classification experiments based on censored samples from the two-parameter Weibull distribution. *J. Stat. Plan. Infer.*, **3**, 39-68.

Nelson, W. (1982). *Applied Life Data Analysis.* New York: Wiley.

Nelson, W. B. (1972). Graphical analysis of accelerated life test data with the inverse power law model, *IEEE Trans. Reliab.*, **R21**, 2-11.

Ross, R. (1994). Formulas to describe the bias and standard deviation of the ML-estimated Weibull shape parameter. *IEEE Trans. Dielec. & Electr. Insul.*, **1**, 247-253.

Ross, R. (1996). Bias and standard deviation due to Weibull parameter estimation for small data sets. *IEEE Trans. Dielec. & Electr. Insul.*, **3**, 28-42.

Stone, G. C. and Lawless, J. F. (1979). Application of Weibull statistics to insulation aging tests. *IEEE Trans. Electr. Insul.*, **EI14**, 233-239.

Thoman, D. R., Bain, L. J. and Antle, C. E. (1969). Inferences on the parameters of the Weibull distribution. *Technometrics*, **11**, 446-460.

Yang, Z. and Xie, M. (2003). Efficient estimation of the Weibull shape parameter based on modified profile likelihood. *J. Statist. Compu. & Simu.*, **73**, 115-123.

Zhang, L. F., Xie, M. and Tang, L. C. (2006). Bias correction for the least squares estimator of Weibull shape parameter with complete and censored data. *Reliab. Eng. & Sys. Safety*, **91**, 930-939.

Table 1: *Simulation results based on Type II censored data*

| | $\beta = 0.5$ | | | $\beta = 1.0$ | | | $\beta = 2.0$ | | |
|---|---|---|---|---|---|---|---|---|---|
| $p$ | $RB_0$ | $RB_1$ | REF | $RB_0$ | $RB_1$ | REF | $RB_0$ | $RB_1$ | REF |
| | | | | $k = 2$ | | | | | |
| | $n_1 = n_2 = 20$ | | | $n_1 = 30, n_2 = 20$ | | | $n_1 = 50, n_2 = 30$ | | |
| 1.0 | 5.231 | 0.258 | 1.26 | 4.029 | 0.124 | 1.20 | 2.622 | 0.237 | 1.13 |
| 0.9 | 6.456 | 0.228 | 1.32 | 5.227 | 0.318 | 1.27 | 3.380 | 0.382 | 1.17 |
| 0.8 | 7.699 | 0.098 | 1.39 | 5.992 | 0.026 | 1.30 | 3.899 | 0.258 | 1.19 |
| 0.7 | 9.576 | 0.243 | 1.48 | 7.569 | 0.256 | 1.37 | 4.648 | 0.215 | 1.24 |
| 0.6 | 12.098 | 0.424 | 1.60 | 9.027 | -0.034 | 1.45 | 5.672 | 0.202 | 1.28 |
| 0.5 | 15.089 | 0.122 | 1.75 | 11.547 | -0.034 | 1.56 | 6.918 | 0.000 | 1.35 |
| 0.4 | 19.937 | -0.276 | 2.00 | 15.527 | -0.016 | 1.76 | 9.595 | 0.408 | 1.46 |
| 0.3 | 30.678 | 0.347 | 2.51 | 22.939 | 0.158 | 2.14 | 13.474 | 0.367 | 1.67 |
| | | | | $k = 8$ | | | | | |
| | $n_i = 10$ | | | $n_i = 20$ | | | $n_i = 50$ | | |
| 1.0 | 8.041 | 0.173 | 1.88 | 3.830 | 0.164 | 1.42 | 1.463 | 0.057 | 1.17 |
| 0.9 | 9.918 | 0.102 | 2.09 | 4.660 | 0.073 | 1.50 | 1.819 | 0.051 | 1.20 |
| 0.8 | 12.456 | 0.361 | 2.38 | 5.836 | 0.231 | 1.66 | 2.189 | 0.044 | 1.25 |
| 0.7 | 14.982 | 0.073 | 2.65 | 7.026 | 0.181 | 1.78 | 2.547 | -0.056 | 1.28 |
| 0.6 | 18.885 | 0.058 | 3.06 | 8.585 | 0.099 | 1.96 | 3.234 | 0.030 | 1.36 |
| 0.5 | 25.064 | 0.338 | 3.69 | 10.836 | 0.014 | 2.18 | 4.026 | -0.009 | 1.44 |
| 0.4 | 34.622 | 0.142 | 4.78 | 14.464 | -0.018 | 2.61 | 5.337 | 0.042 | 1.60 |
| 0.3 | 54.609 | 0.068 | 7.00 | 21.292 | 0.153 | 3.29 | 7.294 | -0.137 | 1.79 |
| 0.2 | – | – | – | 36.840 | -0.021 | 5.12 | 11.977 | -0.010 | 2.29 |

$RB_0$: Relative bias of MLE (in %)
$RB_1$: Relative bias of MMLE (in %)
REF: Relative efficiency of MMLE over MLE
$\alpha_i = i (i = 1, \cdots, k)$

Table 2: *Simulation results based on Type I censored data*

| | $\beta = 0.5$ | | | $\beta = 1.0$ | | | $\beta = 2.0$ | | |
|---|---|---|---|---|---|---|---|---|---|
| $p$ | $RB_0$ | $RB_1$ | REF | $RB_0$ | $RB_1$ | REF | $RB_0$ | $RB_1$ | REF |
| $k = 2$, $n_1 = n_2 = 10$, $\alpha_1 = 1, \alpha_2 = 2$ | | | | | | | | | |
| 0.9 | 9.537 | 1.942 | 1.32 | 9.457 | 1.870 | 1.318 | 9.505 | 1.911 | 1.317 |
| 0.8 | 8.926 | 0.913 | 1.29 | 9.290 | 1.247 | 1.301 | 8.991 | 0.974 | 1.297 |
| 0.7 | 9.860 | 1.353 | 1.30 | 9.825 | 1.331 | 1.304 | 9.621 | 1.134 | 1.299 |
| 0.6 | 11.955 | 2.892 | 1.33 | 10.596 | 1.635 | 1.303 | 10.650 | 1.686 | 1.302 |
| 0.5 | 12.499 | 3.020 | 1.31 | 12.550 | 3.073 | 1.304 | 12.564 | 3.076 | 1.318 |
| $k = 2$, $n_1 = 20, n_2 = 25, \alpha_1 = 1, \alpha_2 = 2$ | | | | | | | | | |
| 0.9 | 3.760 | 0.579 | 1.13 | 3.760 | 0.579 | 1.131 | 3.723 | 0.543 | 1.128 |
| 0.8 | 3.492 | 0.093 | 1.13 | 3.491 | 0.092 | 1.119 | 3.765 | 0.356 | 1.127 |
| 0.7 | 3.759 | 0.162 | 1.13 | 3.758 | 0.162 | 1.122 | 3.976 | 0.372 | 1.128 |
| 0.6 | 4.452 | 0.655 | 1.14 | 4.452 | 0.655 | 1.130 | 4.438 | 0.645 | 1.130 |
| 0.5 | 5.134 | 1.158 | 1.14 | 5.134 | 1.158 | 1.136 | 5.295 | 1.310 | 1.138 |
| 0.4 | 6.725 | 2.537 | 1.15 | 6.725 | 2.537 | 1.144 | 6.501 | 2.324 | 1.145 |
| 0.3 | 9.221 | 4.794 | 1.14 | 9.221 | 4.793 | 1.149 | 9.275 | 4.845 | 1.149 |
| $k = 5$, $n_i = 6$, $\alpha_i = 10, 20, 30, 40, 50$ | | | | | | | | | |
| 0.9 | 13.108 | -0.236 | 1.79 | 13.218 | -0.147 | 1.79 | 13.261 | -0.099 | 1.78 |
| 0.8 | 12.280 | -1.599 | 1.68 | 12.011 | -1.836 | 1.66 | 12.186 | -1.676 | 1.67 |
| 0.7 | 11.483 | -2.891 | 1.57 | 11.380 | -2.966 | 1.57 | 11.344 | -3.021 | 1.57 |
| 0.6 | 11.616 | -3.342 | 1.55 | 11.550 | -3.390 | 1.54 | 11.483 | -3.443 | 1.55 |
| 0.5 | 12.475 | -3.126 | 1.54 | 12.573 | -3.048 | 1.55 | 12.816 | -2.849 | 1.58 |
| 0.4 | 14.659 | -1.788 | 1.58 | 13.732 | -2.562 | 1.54 | 13.902 | -2.437 | 1.56 |
| $k = 5$, $n_i = 10$, $\alpha_i = 10, 20, 30, 40, 50$ | | | | | | | | | |
| 0.9 | 7.038 | -0.417 | 1.43 | 6.735 | -0.703 | 1.40 | 7.001 | -0.449 | 1.42 |
| 0.8 | 6.342 | -1.519 | 1.33 | 6.427 | -1.431 | 1.34 | 6.158 | -1.688 | 1.32 |
| 0.7 | 6.062 | -2.175 | 1.29 | 6.561 | -1.706 | 1.33 | 6.409 | -1.849 | 1.32 |
| 0.6 | 6.459 | -2.169 | 1.30 | 6.319 | -2.297 | 1.29 | 6.289 | -2.332 | 1.29 |
| 0.5 | 6.790 | -2.206 | 1.30 | 7.311 | -1.728 | 1.32 | 6.843 | -2.163 | 1.30 |
| 0.4 | 7.654 | -1.733 | 1.31 | 7.460 | -1.915 | 1.30 | 7.609 | -1.775 | 1.31 |
| 0.3 | 10.066 | 0.160 | 1.34 | 9.493 | -0.360 | 1.33 | 9.639 | -0.232 | 1.34 |

$RB_0$: Relative bias of MLE (in %)
$RB_1$: Relative bias of MMLE (in %)
REF: Relative efficiency of MMLE over MLE

Appeared in: Applied Stochastic Models in Business and Industry, 2007, 23, 373-383.

Table 3: *Estimated reliabilities and percentile lives of Type I insulation*

| Est. Reliab. at Time $t$ | | | Est. $p$th %tile Life | | |
|---|---|---|---|---|---|
| $t$ | MLE | MMLE | $p$ | MLE | MMLE |
| 25 | 0.9990 | 0.9978 | 0.005 | 29.10 | 27.28 |
| 30 | 0.9931 | 0.9878 | 0.01 | 31.08 | 29.36 |
| 35 | 0.9654 | 0.9486 | 0.05 | 36.27 | 34.89 |
| 38 | 0.9195 | 0.8916 | 0.10 | 38.83 | 37.66 |
| 42 | 0.7854 | 0.7444 | 0.25 | 42.70 | 41.89 |
| 45 | 0.6061 | 0.5677 | 0.50 | 46.41 | 45.97 |
| 48 | 0.3716 | 0.3530 | 0.75 | 49.56 | 49.48 |
| 51 | 0.1529 | 0.1579 | 0.90 | 51.99 | 52.21 |
| 54 | 0.0323 | 0.0422 | 0.95 | 53.31 | 53.68 |
| 56 | 0.0065 | 0.0115 | 0.99 | 55.52 | 56.19 |

Table 4: *Failure times of bearing specimens*

| Type | | | | | | $\hat{\beta}$ | $\tilde{\beta}$ |
|---|---|---|---|---|---|---|---|
| I | 3.03 | 5.53 | 5.60 | 9.30 | 9.92 | 2.59 | 2.22 |
| | 12.51 | 12.95 | 15.21 | 16.04 | 16.84 | | |
| II | 3.19 | 4.26 | 4.47 | 4.53 | 4.67 | 2.32 | 2.07 |
| | 4.69 | 5.78 | 6.79 | 9.37 | 12.75 | | |
| III | 3.46 | 5.22 | 5.69 | 6.54 | 9.16 | 3.13 | 2.70 |
| | 9.40 | 10.19 | 10.71 | 12.58 | 13.41 | | |
| IV | 5.88 | 6.74 | 6.90 | 6.98 | 7.21 | 1.94 | 1.75 |
| | 8.14 | 8.59 | 9.80 | 12.28 | 25.46 | | |
| V | 6.43 | 9.97 | 10.39 | 13.55 | 14.45 | 3.65 | 3.16 |
| | 14.72 | 16.81 | 18.39 | 20.84 | 21.51 | | |
| Combined | | | | | | 3.78 | 3.13 |

Appeared in: Applied Stochastic Models in Business and Industry, 2007, 23, 373-383.

Table 5: *Estimated reliabilities and percentile lives of Type I bearings*

| Est. Reliab. at Time $t$ | | | Est. $p$th %tile Life | | |
|---|---|---|---|---|---|
| $t$ | MLE | MMLE | $p$ | MLE | MMLE |
| 2 | 0.9991 | 0.9967 | 0.005 | 3.14 | 2.29 |
| 5 | 0.9713 | 0.9436 | 0.01 | 3.77 | 2.86 |
| 8 | 0.8418 | 0.7764 | 0.05 | 5.81 | 4.81 |
| 10 | 0.6700 | 0.6008 | 0.10 | 7.02 | 6.05 |
| 12 | 0.4503 | 0.4056 | 0.25 | 9.16 | 8.33 |
| 15 | 0.1565 | 0.1626 | 0.50 | 11.56 | 11.03 |
| 17 | 0.0509 | 0.0679 | 0.75 | 13.89 | 13.76 |
| 19 | 0.0107 | 0.0221 | 0.90 | 15.88 | 16.18 |
| 21 | 0.0013 | 0.0054 | 0.95 | 17.03 | 17.60 |
| 22 | 0.0004 | 0.0024 | 0.99 | 19.08 | 20.18 |