



Maximum likelihood estimation of the Weibull distribution with reduced bias

Enes Makalic¹ · Daniel F. Schmidt²

Received: 18 October 2022 / Accepted: 15 March 2023 / Published online: 17 April 2023
© The Author(s) 2023

Abstract

In this short note, we derive a new bias adjusted maximum likelihood estimate for the shape parameter of the Weibull distribution with complete data and type I censored data. The proposed estimate of the shape parameter is significantly less biased and more efficient than the corresponding maximum likelihood estimate, while being simple to compute using existing maximum likelihood software procedures.

Keywords Survival analysis · Kullback–Leibler divergence · Bias · Estimation

1 Introduction

The Weibull distribution, with probability density function

$$p(y|\theta) = \left(\frac{k}{\lambda^k}\right) y^{k-1} \exp\left(-\left(\frac{y}{\lambda}\right)^k\right), \quad (1)$$

where $\theta = (k, \lambda)^\top$ and $k > 0$ is the shape parameter and $\lambda > 0$ is the scale parameter, is a popular distribution in analysis of survival data. Given data $\mathbf{y} = (y_1, \dots, y_n)^\top$, a common approach to estimating the parameters of a Weibull distribution, θ , is via the method of maximum likelihood (ML), in which the parameters are set to values that maximise the log-likelihood of the data

$$\ell(\theta) = -n \log\left(\frac{\lambda^k}{k}\right) + (k-1) \left(\sum_{i=1}^n \log y_i\right) - \sum_{i=1}^n \left(\frac{y_i}{\lambda}\right)^k. \quad (2)$$

The ML estimate of λ is

$$\hat{\lambda}(\mathbf{y}) = \left(\frac{1}{n} \sum_{i=1}^n y_i^k\right)^{\frac{1}{k}}, \quad (3)$$

and the ML estimate of k , $\hat{k}(\mathbf{y})$, is defined implicitly by the estimating equation

$$\frac{n}{k} + \sum_{i=1}^n \log y_i - \frac{n \sum_{i=1}^n y_i^k \log y_i}{\sum_{i=1}^n y_i^k} = 0, \quad (4)$$

and must be obtained by numerical optimisation.

The ML estimate of the Weibull distribution scale parameter λ has negligible bias, even for relatively small sample sizes. In contrast, the ML estimate of the shape parameter k is known to be strongly biased for small sample sizes. Ross (1994) derived a simple bias-reduction adjustment formula for the ML estimate of k

$$\hat{k}_R(\mathbf{y}) = \left(\frac{n-2}{n-0.68}\right) \hat{k}_{ML}(\mathbf{y}), \quad (5)$$

and later extended his approach to censored data (Ross 1996). Hirose (1999) proposed an alternative bias correction method for data with no censoring that was derived by fitting a non-linear function to simulation results. Teimouri and Nadarajah (2013) develop improved ML estimates for the Weibull distribution based on record statistics. In contrast, Yang and Xie (2003) use the modified profile likelihood proposed by Cox

✉ Enes Makalic
emakalic@unimelb.edu.au

Daniel F. Schmidt
daniel.schmidt@monash.edu

¹ Melbourne School of Population and Global Health,
University of Melbourne, Parkville, VIC 3010, Australia

² Faculty of Information Technology, Monash University,
Clayton, VIC 3800, Australia

and Reid (1987, 1992) to derive an alternative ML estimate of k (MLC) from the estimating equation

$$\frac{n-2}{k} + \sum_{i=1}^n \log y_i - \frac{n \sum_i y_i^k \log y_i}{\sum_i y_i^k} = 0. \quad (6)$$

Using simulations, Yang and Xie showed that their estimate of k is less biased than the ML estimate and more efficient than the estimate (5) proposed by Ross. In a follow-up paper, Shen and Yang (2015) developed a profile ML estimate of k in the case of complete and censored samples, and showed that it outperformed MLC in simulations with complete data.

In this paper, we introduce new bias adjusted maximum likelihood estimates for the Weibull distribution for both complete and type I censored data. In addition, we derive a novel formula for the Kullback and Leibler (1951) divergence between two Weibull distributions under type I censoring, a result that does not appear to be widely known.

1.1 Type I censored data

In survival analysis, one typically does not observe complete data and instead has joint realisations of the random variables $(Y = y, \Delta = \delta)$ where $Y = \min(T, c)$ and

$$\Delta = I(T \leq c) = \begin{cases} 1, & \text{if } T \leq c \text{ (observed survival)} \\ 0, & \text{if } T > c \text{ (observed censoring)} \end{cases},$$

where the random variable T denotes the survival time and $c > 0$ is the fixed censoring time. The data comprises the survival time $T = t$ of an item if this is less than the corresponding censoring time c (i.e., $T \leq c$); otherwise, we only know that the item survived beyond time c (i.e., $T > c$).

The log-likelihood of data $D = \{(y_1, \delta_1), \dots, (y_n, \delta_n)\}$ under type I censoring is

$$\ell(\theta) = d \log \left(\frac{k}{\lambda^k} \right) - \frac{1}{\lambda^k} \sum_{i=1}^n y_i^k + \sum_{i=1}^n \log y_i^{\delta_i (k-1)}, \quad (7)$$

where $d = \sum_{i=1}^n \delta_i$ is the number of uncensored observations. The maximum likelihood (ML) estimate of λ is then

$$\hat{\lambda}^k(\mathbf{y}) = \frac{1}{d} \sum_{i=1}^n y_i^k, \quad (8)$$

and $\hat{k}(\mathbf{y})$ is obtained from the estimating equation

$$\frac{d}{k} + \sum_{i=1}^n \delta_i \log y_i - \frac{d \sum_i y_i^k \log y_i}{\sum_i y_i^k} = 0. \quad (9)$$

As in the case of complete data, the ML estimate of k for type I censored data has large bias for small sample sizes,

and for large amounts of censoring. Based on the modified profile likelihood approach, Yang and Xie (2003) propose an alternative estimate of k

$$\frac{d-1}{k} + \sum_{i=1}^n \delta_i \log y_i - \frac{d \sum_i y_i^k \log y_i}{\sum_i y_i^k} = 0. \quad (10)$$

We note that the above score function requires that $d > 1$ to yield a positive estimate for k . Yang and Xie demonstrated that their proposed estimate of k is less biased and more efficient than the regular ML estimate.

Shen and Yang (2015) derived a new second- and third-order bias correction formula for the shape parameter of the Weibull distribution without censoring and with general right-censoring models. Although the new estimate is shown to be effective in correcting bias, it must be computed through bootstrap simulation. The same procedure was later extended to include Weibull regression with complete and general right censoring (Shen and Yang 2017).

More recently, Choi et al. (2020) examine a different problem of Weibull parameter overestimation caused by mass occurrences of (censored) events in the early time period and develop an expectation maximization (EM) algorithm to reduce bias.

Maximum likelihood estimation of the Weibull distribution under more sophisticated censoring schemes has also been studied. Progressive hybrid censoring and generalized progressively hybrid censored data was examined in Lin et al. (2012) and Zhu (2020), respectively. Ng and Wang (2009) and Teimouri (2022) study ML estimation of the Weibull distribution with progressively type I interval censored data. An R package for both progressively type I and type II censored data was developed in Teimouri (2021). Additionally, ML estimation of the Weibull distribution with generalized type I censored data and block censoring was examined by Starling et al. (2021) and Zhu (2020), respectively.

2 A simple adjustment to maximum likelihood estimates to reduce estimation bias

In a landmark paper, Cox and Snell (1968) derived an approximation to the finite sample bias of ML estimates for independent, but not necessarily identically distributed, data (see Appendix A for details). The ML estimate with reduced bias, $\tilde{\theta}_{ML}$, is given by

$$\hat{\theta}_{MML} = \hat{\theta}_{ML} - \text{Bias}(\hat{\theta}_{ML}), \quad (11)$$

where the Cox and Snell formula for $\text{Bias}(\theta)$ is given in Appendix A, and is evaluated at the usual ML estimate $\hat{\theta}_{ML}$. A benefit of this bias approximation formula is that it can be computed even if the ML estimate is not available in closed

form. A similar approach to the above was used to derive bias adjusted ML estimates for the unit Weibull distribution (Mazucheli et al. 2018) and the inverse Weibull distribution (Mazucheli et al. 2018) with complete data only. We now extend these results to the Weibull distribution with complete data and Type I censored data.

Theorem 1 *The finite sample bias of the ML estimate (3) for the Weibull distribution with complete data is*

$$\begin{aligned} \text{Bias}(k) &= k \left(\frac{18(\pi^2 - 2\zeta(3))}{n\pi^4} \right) + O(n^{-2}) \\ &\approx k \left(\frac{1.3795}{n} \right) \end{aligned} \tag{12}$$

where $\zeta(\cdot)$ is the Riemann zeta function. ML estimates of k and λ with reduced bias can be obtained from (11).

Proof The proof involves the application of the Cordeiro and Klein (1994) approach [see (A4) and (A5) in Appendix A], to the Weibull distribution (1). It is well known that expected Fisher information matrix for the Weibull distribution, and its inverse, are given by

$$\begin{aligned} \mathbf{K} &= n \begin{pmatrix} \frac{6(\gamma-1)^2 + \pi^2}{6k^2} & \frac{(\gamma-1)}{\lambda} \\ \frac{(\gamma-1)}{\lambda} & \frac{\lambda^2}{k^2} \end{pmatrix}, \\ \mathbf{K}^{-1} &= \frac{1}{n\pi^2} \begin{pmatrix} 6k^2 & -6(\gamma-1)\lambda \\ -6(\gamma-1)\lambda & \frac{(6(\gamma-1)^2 + \pi^2)\lambda^2}{k^2} \end{pmatrix}, \end{aligned}$$

where $\gamma \approx 0.5772$ is the Euler–Mascheroni constant. Direct calculation shows that the 2×4 matrix \mathbf{A} [see (A5) in Appendix A] has entries

$$\begin{aligned} a_{1,1} &= \frac{n(-12\zeta(3) - 3\gamma(2\gamma(\gamma-7) + \pi^2 + 16) + 7\pi^2 + 12)}{12k^3}, \\ a_{1,2} = a_{2,1} &= -\frac{n(6\gamma(\gamma-4) + \pi^2 + 12)}{12k\lambda}, \\ a_{2,2} &= -\frac{n(\gamma k + k + \gamma - 1)}{2\lambda^2}, \\ a_{1,3} &= -\frac{n(6\gamma(\gamma-4) + \pi^2 + 12)}{12k\lambda}, \\ a_{1,4} = a_{2,3} &= \frac{n(-\gamma k + 3k + \gamma - 1)}{2\lambda^2}, \\ a_{2,4} &= -\frac{n(k-1)k^2}{2\lambda^3}. \end{aligned}$$

Substituting \mathbf{K}^{-1} and \mathbf{A} into (A4) and simplifying completes the proof. \square

From (12), we observe that the ML estimate of k is upwardly biased for any finite n . A key advantage of the proposed bias adjusted estimate is that it can be trivially computed in any software that implements ML Weibull estimation. We now

derive a similar correction for the more complex case of type I censoring.

Theorem 2 *The finite sample bias of the maximum likelihood estimate (3) for the Weibull distribution with type I censored data is*

$$\text{Bias}(k) = k \left(\frac{f(p)}{n} \right) + O(n^{-2}), \tag{13}$$

where $z_c = (c/\lambda)^k$, $p = 1 - \exp(-z_c)$ is the proportion of uncensored observations,

$$f(p) = \frac{(6\gamma_2 + \gamma_3)p^2 - 3(2\gamma_1 + \gamma_2)\gamma_1 p + 2\gamma_1^3}{2(\gamma_1^2 - \gamma_2 p)^2}, \tag{14}$$

and $\gamma(\cdot, \cdot)$ is the incomplete gamma function

$$\gamma(z, x) = \int_0^x t^{z-1} \exp(-t) dt \tag{15}$$

whose j th derivative is

$$\gamma^{(j)}(z, x) = \frac{d^j \gamma(z, x)}{dz^j}. \tag{16}$$

For brevity, we use the shorthand notation $\gamma_j \equiv \gamma^{(j)}(1, z_c)$ to denote the j th derivative of the incomplete gamma function evaluated at $(1, z_c)$. As in the case of complete data, the ML estimate of k with reduced bias can be obtained from (11).

Proof The expected Fisher information matrix for the Weibull distribution with type I censoring is (Watkins and John (2004))

$$\begin{aligned} \mathbf{K} &= n \begin{pmatrix} \frac{p+2\gamma_1+\gamma_2}{k^2} & -\frac{p+\gamma_1}{\lambda} \\ -\frac{p+\gamma_1}{\lambda} & \frac{k^2 p}{\lambda^2} \end{pmatrix}, \\ \mathbf{K}^{-1} &= \frac{1}{n(\gamma_2 p - \gamma_1^2)} \begin{pmatrix} k^2 p & \lambda(p + \gamma_1) \\ \lambda(p + \gamma_1) & \frac{\lambda^2(p+2\gamma_1+\gamma_2)}{k^2} \end{pmatrix}. \end{aligned}$$

By direct calculation we have

$$\begin{aligned} a_{1,1} &= \frac{n(2p + 8\gamma_1 + 7\gamma_2 + \gamma_3)}{2k^3}, \\ a_{1,2} = a_{2,1} &= -\frac{n(2p + 4\gamma_1 + \gamma_2)}{2k\lambda}, \\ a_{2,2} &= \frac{n(\gamma_1(k+1) - (k-1)p)}{2\lambda^2}, \\ a_{1,3} &= -\frac{n(2p + 4\gamma_1 + \gamma_2)}{2k\lambda}, \\ a_{1,4} = a_{2,3} &= \frac{n((3k-1)p + \gamma_1(k-1))}{2\lambda^2}, \\ a_{2,4} &= -\frac{n(k-1)k^2 p}{2\lambda^3}. \end{aligned}$$

We note that

$$\begin{aligned} \lim_{p \rightarrow 1} \gamma_1 &= -\gamma, & \lim_{p \rightarrow 1} \gamma_2 &= \gamma^2 + \frac{\pi^2}{6}, \\ \lim_{p \rightarrow 1} \gamma_3 &= -\gamma^3 - \frac{\gamma\pi^2}{2} + \psi^{(2)}(1), \end{aligned}$$

where $\psi^{(2)}(1)$ is the second derivative of the polygamma function evaluated at 1. As expected, the matrix \mathbf{A} for type I censored data converges to the corresponding matrix with complete data as $p \rightarrow 1$. Substituting \mathbf{K}^{-1} and \mathbf{A} into (A4) and simplifying completes the proof. \square

Figure 1 shows the bias adjustment as a function of the proportion of uncensored observations p . As the proportion of uncensored observations $p \rightarrow 1$ (i.e., no censoring), $f(p) \rightarrow (\approx)1.3795$ as expected. Additionally, $f(p) \rightarrow \infty$ as the proportion of censored data is increased (i.e., $p \rightarrow 0$).

Remark As noted in the introduction, the ML estimate of the scale parameter λ has negligible bias even for small sample sizes. For complete data, this finite sample bias, computed using the Cox and Snell methodology, is:

$$\begin{aligned} \text{Bias}(\lambda) &= \lambda \left(\frac{1}{nk^2} \left(\frac{3(\gamma - 1)^2}{\pi^2} + \frac{1}{2} \right) \right. \\ &\quad \left. + \frac{1}{nk} \left(\frac{36(\gamma - 1)\zeta(3)}{\pi^4} + \frac{15 - 12\gamma}{\pi^2} - 1 \right) \right) \\ &\quad + O(n^{-2}) \approx \lambda \left(\frac{0.5543}{nk^2} - \frac{0.3698}{nk} \right) \end{aligned} \quad (17)$$

where $\gamma \approx 0.5772$ is the Euler–Mascheroni constant. For type I censored data, the finite sample bias is:

$$\text{Bias}(\lambda) = \lambda \left(\frac{f_1(p)}{nk^2} + \frac{f_2(p)}{nk} \right) + O(n^{-2}), \quad (18)$$

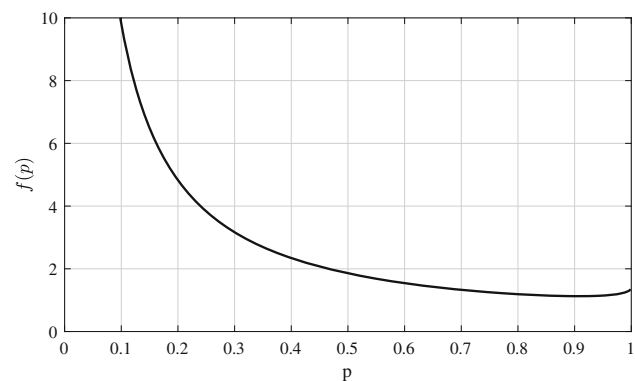


Fig. 1 Bias adjustment $f(p)$ for the maximum likelihood estimate of the Weibull distribution shape parameter k as a function of the proportion of uncensored observations. The estimate with the lowest bias or mean squared error is shown in boldface. $p = 1 - \exp(-(c/\lambda)^k)$

where p is the proportion of uncensored observations, and

$$\begin{aligned} f_1(p) &= -\frac{p + 2\gamma_1 + \gamma_2}{2\gamma_1^2 - 2\gamma_2 p}, \\ f_2(p) &= \frac{(5\gamma_2 + \gamma_3)p^2 + (-5\gamma_1^2 + (\gamma_2 + \gamma_3)\gamma_1 - 2\gamma_2^2)p + (\gamma_2 - 2\gamma_1)\gamma_1^2}{2(\gamma_1^2 - \gamma_2 p)^2}, \end{aligned}$$

with $\gamma_j \equiv \gamma^{(j)}(1, z_c)$ again denoting the j th derivative of the incomplete gamma function (16) evaluated at $(1, z_c)$.

2.1 Simulation

We performed a simulation to examine the finite sample behaviour of the new bias adjusted ML estimates of k for both complete and type I censored data. In all simulations, the scale parameter of the data generating model was set to $\lambda^* = 1$ without loss of generality. Due to the scale invariance of the maximum likelihood estimator and the negligible bias in estimating λ^* , the simulation results for other values of λ^* are expected to yield similar conclusions.

2.1.1 Complete data

For each run of the simulation, we generated n data points from the model Weibull($k^*, \lambda^* = 1$) where $n = \{10, 20, 50\}$ and the shape parameter was set to $k^* \in \{0.5, 1, 5, 10\}$. Regular maximum likelihood (ML) estimates, our proposed bias adjusted maximum likelihood estimates (MMLE), conditional maximum likelihood estimates (MLC) proposed by Yang and Xie (2003), and the profile maximum likelihood estimates of Shen and Yang (MLP) (Shen and Yang 2015) were then computed from the data. We used the second-order bias reduction of Shen and Yang as it was virtually indistinguishable from the third-order formula in our tests. We performed 10^5 simulations for each combination of (k^*, n) and recorded the average bias, mean squared error and Kullback–Leibler (KL) divergence (Kullback and Leibler 1951) from the data generating model (see Appendix B). Simulation results are shown in Table 1 with the KL results omitted for ease of presentation.

All three bias adjusted ML estimates of k result in a significant reduction in bias compared to the usual ML estimate. Compared to MLC, our proposed estimate yields smaller mean squared error and KL divergence, especially as k increases. The profile ML estimate has a slightly smaller bias than our estimate, while the mean squared error and the KL divergence for the two estimates are virtually identical. In contrast to both the MLC and MLP estimates, our bias adjusted ML estimate of k is simple to compute in software via existing Weibull ML estimation procedures and does not require the use of the parametric bootstrap.

Table 1 Bias and mean squared error for maximum likelihood (ML), conditional maximum likelihood of Yang and Xie (MLC), profile maximum likelihood of Shen and Yang (MLP) and our bias adjusted maximum likelihood (MMLE) estimates of k^* computed over 10^5 simulations with $\lambda^* = 1$. The estimate with the lowest bias or mean squared error is shown in boldface.

| n | k^* | Bias | | | | Mean squared error | | | |
|-----|-------|-------|-------|---------------|-------|--------------------|-------|--------------|--------------|
| | | ML | MLC | MLP | MMLE | ML | MLC | MLP | MMLE |
| 10 | 0.5 | 0.085 | 0.008 | 0.001 | 0.004 | 0.038 | 0.023 | 0.022 | 0.023 |
| | 1.0 | 0.170 | 0.017 | 0.002 | 0.009 | 0.151 | 0.092 | 0.089 | 0.090 |
| | 5.0 | 0.852 | 0.088 | 0.014 | 0.045 | 3.775 | 2.305 | 2.241 | 2.268 |
| | 10.0 | 1.701 | 0.172 | 0.026 | 0.087 | 15.102 | 9.218 | 8.978 | 9.079 |
| 20 | 0.5 | 0.038 | 0.004 | 0.001 | 0.001 | 0.012 | 0.009 | 0.009 | 0.009 |
| | 1.0 | 0.077 | 0.009 | 0.001 | 0.003 | 0.048 | 0.037 | 0.037 | 0.037 |
| | 5.0 | 0.382 | 0.042 | 0.004 | 0.011 | 1.203 | 0.928 | 0.916 | 0.917 |
| | 10.0 | 0.755 | 0.075 | 0.000 | 0.014 | 4.769 | 3.686 | 3.636 | 3.639 |
| 50 | 0.5 | 0.014 | 0.002 | 0.000 | 0.000 | 0.004 | 0.003 | 0.003 | 0.003 |
| | 1.0 | 0.029 | 0.003 | -0.000 | 0.000 | 0.015 | 0.013 | 0.013 | 0.013 |
| | 5.0 | 0.143 | 0.016 | 0.001 | 0.001 | 0.367 | 0.329 | 0.328 | 0.327 |
| | 10.0 | 0.290 | 0.036 | 0.005 | 0.006 | 1.458 | 1.308 | 1.300 | 1.299 |

2.1.2 Type I censored data

We also conducted a similar experiment in the setting of type I censored data. For each iteration of the simulation, we generated n data points from the model Weibull($k^*, \lambda^* = 1$) where $n = \{10, 20, 30\}$; the shape parameter was again set to $k^* \in \{0.5, 1, 5, 10\}$. The proportion of uncensored observations was $p \in \{0.3, 0.5, 0.7, 0.9\}$. In addition to the bias and the mean squared error in estimating the shape parameter, we computed the Kullback–Leibler (KL) divergence (Kullback and Leibler 1951) between the data generating model and each estimated model (see Appendix B).

The newly proposed bias adjustment estimate of k (MMLE) was compared to the standard ML estimate, the conditional maximum likelihood estimate (MLC) proposed by Yang and Xie (2003) and the profile maximum likelihood estimate (MLP) of Shen and Yang (2015). The third-order profile ML estimate suffered from issues regarding numerical stability for small n and large amounts of censoring occasionally resulted in a negative estimate of k^* ; hence all the comparisons were made with the second-order variant. We restricted the experiments to exclude data sets where the number of uncensored observations $d (= \sum_i \delta_i) < 2$, as MLC may result in negative estimates of k for $d < 2$, though we note this does not cause a problem for our proposed MMLE method. The results of these simulations, averaged over 10^5 runs for each combination of (n, p, k^*) , are shown in Table 2, with the KL results omitted for ease of presentation.

We observe that our MMLE estimate of k is more efficient and less biased than the standard ML estimate of k for all tested values of (n, p, k^*) . The conditional ML estimate of k is, in general, more biased and has higher mean squared error compared to the MLP and our MMLE estimates. In terms of bias reduction, the profile ML estimate of k is virtually identical to our MMLE for $n \geq 30$. For small sample sizes

($n = 20$) and higher levels of censoring ($p \leq 0.5$), the MMLE estimate appears superior to MLC and MLP in terms of bias, mean squared error and KL divergence. Additionally, in contrast to the profile ML method, our MMLE estimate is easily computed without the need for numerical simulation, and as such can be easily integrated into any software that implements fitting of the Weibull distribution to complete and type I censored data.

2.2 Real data

To illustrate the usefulness of our new bias adjusted maximum likelihood estimates, we consider real data on failure voltages from (Lawless 2002, p. 240) that was also analysed by Shen and Yang (2015). The data consists of failure voltages of two types of electrical cable insulation (type 1 and type 2) of 20 specimens each, and is shown in Table 3 for completeness.

Assuming that the failure voltages can be modelled adequately by the Weibull distribution, the ML estimates of the shape and scale parameters for type 1 cables are $\hat{k}_{ML} = 9.38$ and $\hat{\lambda}_{ML} = 47.78$, respectively, and for type 2 cables, the ML estimates are $\hat{k}_{ML} = 9.14$ and $\hat{\lambda}_{ML} = 59.12$. The newly proposed MMLE estimates of the shape parameter for type 1 and type 2 cables are easily obtained from the corresponding ML estimates using (12):

$$\hat{k}_{MMLE} = 9.38 - 9.38 \left(\frac{1.3795}{20} \right) = 8.74 \quad (\text{type 1}),$$

$$\hat{k}_{MMLE} = 9.14 - 9.14 \left(\frac{1.3795}{20} \right) = 8.51 \quad (\text{type 2}).$$

The estimates of the shape parameter proposed in Yang and Xie (2003) and Shen and Yang (2015) are significantly closer to our bias adjusted estimates than to the original maximum

Table 2 Bias and mean squared error for maximum likelihood (ML), conditional maximum likelihood of Yang and Xie (MLC), profile maximum likelihood of Shen and Yang (MLP) and our bias adjusted maximum likelihood (MMLE) estimates of k^* computed over 10^5

simulations with $\lambda^* = 1$; p denotes the proportion of uncensored observations. The estimate with the lowest bias or mean squared error is shown in boldface

| n | p | k^* | Bias | | | | Mean squared error | | | |
|-----|-----|-------|-------|----------------|--------------|----------------|--------------------|--------------|--------|---------------|
| | | | ML | MLC | MLP | MMLE | ML | MLC | MLP | MMLE |
| 10 | 0.3 | 0.5 | 0.115 | 0.021 | 0.004 | 0.002 | 0.220 | 0.150 | 0.090 | 0.090 |
| | | 1.0 | 0.228 | 0.040 | 0.005 | - 0.001 | 0.605 | 0.401 | 0.303 | 0.301 |
| | | 5.0 | 1.156 | 0.214 | 0.033 | 0.008 | 14.757 | 9.683 | 7.292 | 7.248 |
| | | 10.0 | 2.251 | 0.374 | 0.058 | - 0.007 | 55.591 | 36.196 | 32.112 | 28.799 |
| | 0.5 | 0.5 | 0.051 | 0.001 | 0.001 | -0.003 | 0.037 | 0.028 | 0.028 | 0.028 |
| | | 1.0 | 0.108 | 0.007 | 0.008 | 0.000 | 0.144 | 0.109 | 0.110 | 0.108 |
| | | 5.0 | 0.556 | 0.051 | 0.053 | 0.014 | 3.672 | 2.785 | 2.785 | 2.738 |
| | | 10.0 | 1.095 | 0.084 | 0.085 | 0.009 | 15.172 | 11.561 | 11.571 | 11.377 |
| | 0.7 | 0.5 | 0.034 | - 0.002 | 0.003 | -0.003 | 0.019 | 0.015 | 0.016 | 0.016 |
| | | 1.0 | 0.075 | 0.003 | 0.013 | 0.002 | 0.081 | 0.065 | 0.068 | 0.066 |
| | | 5.0 | 0.381 | 0.023 | 0.075 | 0.017 | 2.048 | 1.660 | 1.730 | 1.676 |
| | | 10.0 | 0.681 | - 0.028 | 0.073 | -0.042 | 7.474 | 6.122 | 6.363 | 6.181 |
| | 0.9 | 0.5 | 0.032 | 0.001 | 0.012 | 0.001 | 0.013 | 0.011 | 0.012 | 0.011 |
| | | 1.0 | 0.063 | 0.001 | 0.023 | 0.002 | 0.051 | 0.042 | 0.045 | 0.042 |
| | | 5.0 | 0.314 | 0.001 | 0.111 | 0.008 | 1.320 | 1.082 | 1.165 | 1.083 |
| | | 10.0 | 0.632 | 0.006 | 0.226 | 0.019 | 5.306 | 4.346 | 4.680 | 4.349 |
| 20 | 0.3 | 0.5 | 0.114 | 0.020 | 0.003 | - 0.000 | 0.262 | 0.177 | 0.185 | 0.102 |
| | | 1.0 | 0.231 | 0.042 | 0.006 | 0.001 | 0.756 | 0.504 | 0.339 | 0.337 |
| | | 5.0 | 1.149 | 0.207 | 0.036 | 0.002 | 16.835 | 11.123 | 9.821 | 7.900 |
| | | 10.0 | 2.312 | 0.427 | 0.108 | 0.014 | 82.069 | 54.570 | 59.320 | 37.053 |
| | 0.5 | 0.5 | 0.054 | 0.003 | 0.003 | - 0.000 | 0.037 | 0.028 | 0.028 | 0.028 |
| | | 1.0 | 0.109 | 0.008 | 0.008 | 0.000 | 0.151 | 0.115 | 0.114 | 0.112 |
| | | 5.0 | 0.534 | 0.030 | 0.029 | - 0.009 | 3.720 | 2.843 | 2.826 | 2.780 |
| | | 10.0 | 1.091 | 0.081 | 0.080 | 0.004 | 15.178 | 11.578 | 11.507 | 11.321 |
| | 0.7 | 0.5 | 0.036 | 0.001 | 0.006 | 0.000 | 0.020 | 0.016 | 0.017 | 0.016 |
| | | 1.0 | 0.072 | 0.001 | 0.011 | - 0.000 | 0.078 | 0.064 | 0.067 | 0.065 |
| | | 5.0 | 0.366 | 0.009 | 0.061 | 0.003 | 1.990 | 1.619 | 1.692 | 1.640 |
| | | 10.0 | 0.724 | 0.012 | 0.114 | - 0.002 | 7.781 | 6.334 | 6.608 | 6.408 |
| | 0.9 | 0.5 | 0.031 | 0.000 | 0.011 | 0.001 | 0.013 | 0.011 | 0.012 | 0.011 |
| | | 1.0 | 0.063 | 0.001 | 0.023 | 0.002 | 0.053 | 0.043 | 0.047 | 0.043 |
| | | 5.0 | 0.309 | -0.004 | 0.106 | 0.003 | 1.298 | 1.066 | 1.146 | 1.066 |
| | | 10.0 | 0.632 | 0.006 | 0.225 | 0.019 | 5.283 | 4.332 | 4.662 | 4.332 |
| 30 | 0.3 | 0.5 | 0.065 | 0.007 | 0.001 | -0.001 | 0.056 | 0.041 | 0.037 | 0.037 |
| | | 1.0 | 0.133 | 0.016 | 0.004 | - 0.000 | 0.336 | 0.257 | 0.262 | 0.179 |
| | | 5.0 | 0.653 | 0.068 | 0.009 | - 0.009 | 5.621 | 4.184 | 3.907 | 3.788 |
| | | 10.0 | 1.334 | 0.161 | 0.045 | 0.008 | 23.086 | 17.180 | 16.069 | 15.252 |
| | 0.5 | 0.5 | 0.034 | 0.001 | 0.002 | - 0.000 | 0.020 | 0.017 | 0.017 | 0.017 |
| | | 1.0 | 0.070 | 0.004 | 0.006 | 0.001 | 0.081 | 0.067 | 0.068 | 0.067 |
| | | 5.0 | 0.341 | 0.015 | 0.025 | 0.000 | 2.019 | 1.680 | 1.698 | 1.679 |
| | | 10.0 | 0.684 | 0.032 | 0.051 | 0.002 | 8.095 | 6.732 | 6.815 | 6.740 |
| | 0.7 | 0.5 | 0.024 | 0.001 | 0.004 | 0.000 | 0.012 | 0.010 | 0.010 | 0.010 |
| | | 1.0 | 0.047 | 0.001 | 0.007 | 0.000 | 0.046 | 0.040 | 0.041 | 0.041 |
| | | 5.0 | 0.235 | 0.004 | 0.038 | 0.001 | 1.170 | 1.020 | 1.050 | 1.030 |
| | | 10.0 | 0.475 | 0.012 | 0.080 | 0.006 | 4.662 | 4.058 | 4.179 | 4.096 |

Table 2 continued

| <i>n</i> | <i>p</i> | <i>k</i> * | Bias | | | | Mean squared error | | | |
|----------|----------|------------|-------|----------------|-------|----------------|--------------------|-------|-------|--------------|
| | | | ML | MLC | MLP | MMLE | ML | MLC | MLP | MMLE |
| | 0.9 | 0.5 | 0.020 | − 0.001 | 0.006 | − 0.000 | 0.008 | 0.007 | 0.007 | 0.007 |
| | | 1.0 | 0.040 | − 0.000 | 0.014 | 0.000 | 0.031 | 0.027 | 0.029 | 0.027 |
| | | 5.0 | 0.198 | − 0.003 | 0.065 | − 0.001 | 0.778 | 0.684 | 0.716 | 0.684 |
| | | 10.0 | 0.402 | − 0.000 | 0.137 | 0.005 | 3.133 | 2.750 | 2.881 | 2.748 |

Table 3 Failure voltages (measured in kV/mm) for two types of electrical cable insulation (type 1 and type 2) of 20 specimens each

| | | | | | | | | | | |
|--------|------|------|------|------|------|------|------|------|------|------|
| Type 1 | 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 | 47.3 | 47.6 | 49.2 | 50.4 | 50.9 | 52.4 | 56.3 |
| Type 2 | 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 |

likelihood estimates, which exhibit significant upward bias. As expected, bias adjusted estimates of the scale parameter are all approximately equal to the corresponding maximum likelihood estimates.

As a further example, we consider the criminal recidivism data first published in ROSSI et al. (1980). This data consists of 432 survival times of individuals released from Maryland state prisons in the 1970s and followed up for 52 weeks after release (i.e., all censored observations were censored at 52 weeks). Approximately 75% of the observations were censored, indicating a relatively high degree of censoring. Assuming that the survival times are Weibull distributed, regular ML estimates of the shape and scale parameter were found to be $\hat{k}_{ML} = 1.37$ and $\hat{\lambda}_{ML} = 123.68$, respectively. In contrast, our bias adjusted MMLE estimates were $\hat{k}_{MMLE} = 1.35$ and $\hat{\lambda}_{MMLE} = 123.68$. As expected, all three bias adjusted estimates examined in this manuscript were similar to the ML estimates due to the relatively large sample size.

We then randomly sampled 20 observations from the original data without replacement; these were 9, 27, 35, 43 and 46 weeks, with the remaining 15 observations censored at 52 weeks. The ML estimate of the shape parameter of this subsample was found to be $\hat{k}_{ML} = 1.72$; note this is substantially higher than the ML estimate of 1.37 obtained on the full data sample. In contrast, our bias adjusted MMLE estimate from this subsample was $\hat{k}_{MMLE} = 1.39$, which is very close to the estimate obtained on the full sample. This again demonstrates that the ML estimate is strongly upwards biased, especially in the case of smaller sample sizes and high degrees of censoring.

3 Discussion

Our proposed MMLE approach to first-order bias correction results in improved performance compared to the standard

ML estimate in small to medium sample sizes for both complete and type I censored data. The methodology introduced here can also be extended to more sophisticated censoring plans. As an example, consider progressive type I interval censoring (PTIC) (Aggarwala 2001). Here, we have *n* items entering a life experiment at time $T_0 = 0$. The items are monitored at $m > 0$ pre-selected times $T_1 < T_2 < \dots < T_m$ only, with the experiment scheduled to terminate at the last observation time T_m . During each inspection time T_i ($i = 1, \dots, m$), the number of failures Y_i for the time interval $(T_{i-1}, T_i]$ is recorded and R_i surviving items are removed from the experiment at random. The number of removed items may be pre-specified as a percentage p_i of the remaining surviving items X_i ; that is, $R_i = \lfloor p_i X_i \rfloor$ where $0 < p_i \leq 1$, $\lfloor z \rfloor$ is the largest integer less than or equal to z and $p_m = 1$ as all surviving items are removed from the experiment at time t_m . Thus, PTIC may be summarised by m triplets $\{Y_i, R_i, T_i\}_{i=1}^m$.

To obtain first order bias adjusted ML estimates for the Weibull distribution under PTIC, we require the expected Fisher information matrix as well as the expected third order derivatives of the log-likelihood function. A general expression for the expected Fisher information matrix under PTIC is given by Theorem 3.3 (Teimouri 2022) while Theorem 3.4 (Teimouri 2022) derives the expected third order derivatives for an arbitrary log-likelihood under PTIC. These two formulas are easily specialised to Weibull distributed survival times.

A limitation of the Cox and Snell first order bias adjustment approach in the case of PTIC is that an analytic expression for the bias correction is not easily available and the estimator must instead be implemented in software. This is because the expected Fisher information matrix and the expected third order derivatives are somewhat long and cumbersome due to interval censoring and the summation over m monitoring times. To fit Weibull distributed data under

PTIC, we recommended the R package `bccp` (Teimouri 2021) which features a numerical implementation of the Cox and Snell bias adjustment approach under progressive type I and type II interval censoring for a wide range of distributions.

Acknowledgements Not applicable.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

Declarations

Conflict of interest Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Cox and Snell approximation

Let $\theta \in \mathbb{R}^p$, where $p > 0$ is the number of free parameters, which is $p = 2$ in the case of the Weibull model. Cox and Snell showed that the bias for the s th element of the ML estimate $\hat{\theta}_{ML}$ can be written as

$$[\text{Bias}(\theta)]_s = \sum_{i=1}^p \sum_{j=1}^p \sum_{l=1}^p \kappa^{s,i} \kappa^{j,l} \left(\frac{1}{2} \kappa_{ijl} + \kappa_{ij,l} \right) + O(n^{-2}) \tag{A1}$$

for $s = 1, \dots, p$, where the cumulants are

$$\kappa_{ij} = \mathbb{E} \left\{ \frac{\partial^2 \ell(\theta)}{\partial \theta_i \partial \theta_j} \right\}, \quad \kappa_{ijl} = \mathbb{E} \left\{ \frac{\partial^3 \ell(\theta)}{\partial \theta_i \partial \theta_j \partial \theta_l} \right\}, \tag{A2}$$

$$\kappa_{ij,l} = \mathbb{E} \left\{ \frac{\partial^2 \ell(\theta)}{\partial \theta_i \partial \theta_j} \frac{\partial \ell(\theta)}{\partial \theta_l} \right\}, \tag{A3}$$

for $i, j = 1, \dots, p$ and $\kappa^{i,j}$ is the (i, j) th entry of the inverse of the expected Fisher information matrix $\mathbf{K} = \{-\kappa_{ij}\}$. Following (Cordeiro and Klein 1994), we can compactly write this in matrix notation as

$$\text{Bias}(\theta) = \mathbf{K}^{-1} \text{Avec}(\mathbf{K}^{-1}) + O(n^{-2}), \tag{A4}$$

where the matrix \mathbf{A} is the $(p \times p^2)$ matrix given by

$$\mathbf{A} = \left[\mathbf{A}^{(1)} | \mathbf{A}^{(2)} | \dots | \mathbf{A}^{(p)} \right], \quad \mathbf{A}^{(l)} = \{a_{ij}^{(l)}\} \tag{A5}$$

$$a_{ij}^{(l)} = \kappa_{ij}^{(l)} - \frac{1}{2} \kappa_{ijl}, \quad \kappa_{ij}^{(l)} = \frac{\partial \kappa_{ij}}{\partial \theta_l} \tag{A6}$$

for $i, j, l = 1, \dots, p$.

Kullback–Leibler divergence

For the case of complete data, the Kullback–Leibler (KL) divergence between the data generating model Weibull(k_0, λ_0) and the approximating model Weibull(k_1, λ_1) is

$$\begin{aligned} \text{KL}(k_0, \lambda_0 || k_1, \lambda_1) &= \left(\frac{\lambda_0}{\lambda_1} \right)^{k_1} \left(\frac{k_1}{k_0} \right) \Gamma \left(\frac{k_1}{k_0} \right) + \left(\frac{k_1}{k_0} - 1 \right) \gamma \\ &\quad + \log \left(\frac{k_0}{k_1} \left(\frac{\lambda_1}{\lambda_0} \right)^{k_1} \right) - 1. \end{aligned}$$

Assuming type I censoring, the KL divergence between two Weibull models Weibull(k_0, λ_0) and Weibull(k_1, λ_1) is

$$\begin{aligned} \text{KL}(k_0, \lambda_0 || k_1, \lambda_1) &= \exp(- (c/\lambda_0)^{k_0}) A_1 \\ &\quad + \left(\frac{\lambda_0}{\lambda_1} \right)^{k_1} A_2 + \left(1 - \frac{k_1}{k_0} \right) A_3 \\ &\quad + \log \left(\frac{k_0}{k_1} \left(\frac{\lambda_1}{\lambda_0} \right)^{k_1} \right) - 1, \end{aligned}$$

where

$$A_1 = \log \left(\frac{k_1}{k_0} c^{k_1-k_0} \lambda_0^{k_0} \lambda_1^{-k_1} \right) + \left(\frac{c}{\lambda_1} \right)^{k_1} + 1,$$

$$A_2 = \Gamma \left(\frac{k_1}{k_0} + 1 \right) - \Gamma \left(\frac{k_1}{k_0} + 1, \left(\frac{c}{\lambda} \right)^k \right),$$

$$A_3 = \text{Ei} \left(- \left(\frac{c}{\lambda_0} \right)^{k_0} \right) - \gamma,$$

and $\text{Ei}(\cdot)$ is the exponential integral function

$$\text{Ei}(z) = - \int_{-z}^{\infty} \frac{\exp(-t)}{t} dt. \tag{B7}$$

References

Aggarwala, R.: Progressive interval censoring: some mathematical results with applications to inference. *Commun. Stat. Theory Methods* **30**(8–9), 1921–1935 (2001). <https://doi.org/10.1081/STA-100105705>

- Choi, K., Park, S.M., Han, S., Yim, D.-S.: A partial imputation EM-algorithm to adjust the overestimated shape parameter of the Weibull distribution fitted to the clinical time-to-event data. *Comput. Methods Programs Biomed.* **197**, 105697 (2020). <https://doi.org/10.1016/j.cmpb.2020.105697>
- Cordeiro, G.M., Klein, R.: Bias correction in ARMA models. *Stat. Probab. Lett.* **19**(3), 169–176 (1994). [https://doi.org/10.1016/0167-7152\(94\)90100-7](https://doi.org/10.1016/0167-7152(94)90100-7)
- Cox, D.R., Reid, N.: Parameter orthogonality and approximate conditional inference. *J. R. Stat. Soc. (Ser. B)* **49**(1), 1–39 (1987)
- Cox, D.R., Reid, N.: A note on the difference between profile and modified profile likelihood. *Biometrika* **79**(2), 408–411 (1992)
- Cox, D.R., Snell, E.J.: A general definition of residuals. *J. R. Stat. Soc. Ser. B (Methodol.)* **30**(2), 248–265 (1968). <https://doi.org/10.1111/j.2517-6161.1968.tb00724.x>
- Hirose, H.: Bias correction for the maximum likelihood estimates in the two-parameter Weibull distribution. *IEEE Trans. Dielectr. Electr. Insul.* **6**(1), 66–68 (1999). <https://doi.org/10.1109/94.752011>
- Kullback, S., Leibler, R.A.: On information and sufficiency. *Ann. Math. Stat.* **22**(1), 79–86 (1951)
- Lawless, J.F.: *Statistical Models and Methods for Lifetime Data*. Wiley, New York (2002). <https://doi.org/10.1002/9781118033005>
- Lin, C.-T., Chou, C.-C., Huang, Y.-L.: Inference for the Weibull distribution with progressive hybrid censoring. *Comput. Stat. Data Anal.* **56**(3), 451–467 (2012). <https://doi.org/10.1016/j.csda.2011.09.002>
- Mazucheli, J., Menezes, A.F.B., Dey, S.: Bias-corrected maximum likelihood estimators of the parameters of the inverse Weibull distribution. *Commun. Stat. Simul. Comput.* **48**(7), 2046–2055 (2018). <https://doi.org/10.1080/03610918.2018.1433838>
- Mazucheli, J., Menezes, A.F.B., Dey, S.: Bias-corrected maximum likelihood estimators of the parameters of the inverse Weibull distribution. *Commun. Stat. Simul. Comput.* **48**(7), 2046–2055 (2018). <https://doi.org/10.1080/03610918.2018.1433838>
- Ng, H.K.T., Wang, Z.: Statistical estimation for the parameters of Weibull distribution based on progressively type-i interval censored sample. *J. Stat. Comput. Simul.* **79**(2), 145–159 (2009). <https://doi.org/10.1080/00949650701648822>
- Ross, R.: Formulas to describe the bias and standard deviation of the ML-estimated Weibull shape parameter. *IEEE Trans. Dielectr. Electr. Insul.* **1**(2), 247–253 (1994). <https://doi.org/10.1109/94.300257>
- Ross, R.: Bias and standard deviation due to Weibull parameter estimation for small data sets. *IEEE Trans. Dielectr. Electr. Insul.* **3**(1), 28–42 (1996). <https://doi.org/10.1109/94.485512>
- Rossi, P.H., Berk, R.A., Lenihan, K.J. (eds.): Academic Press, New York (1980). <https://doi.org/10.1016/B978-0-12-598240-5.50004-1>
- Shen, Y., Yang, Z.: Bias-correction for Weibull common shape estimation. *J. Stat. Comput. Simul.* **85**(15), 3017–3046 (2015). <https://doi.org/10.1080/00949655.2014.949714>
- Shen, Y., Yang, Z.: Improved likelihood inferences for Weibull regression model. *J. Stat. Comput. Simul.* **87**(12), 2349–2371 (2017). <https://doi.org/10.1080/00949655.2017.1331441>
- Starling, J.K., Mastrangelo, C., Choe, Y.: Improving Weibull distribution estimation for generalized type I censored data using modified SMOTE. *Reliab. Eng. Syst. Saf.* **211**, 107505 (2021). <https://doi.org/10.1016/j.res.2021.107505>
- Teimouri, M.: bccp: an r package for life-testing and survival analysis. *Comput. Stat.* **37**(1), 469–489 (2021). <https://doi.org/10.1007/s00180-021-01129-9>
- Teimouri, M.: Bias corrected maximum likelihood estimators under progressive type-I interval censoring scheme. *Commun. Stat. Simul. Comput.* **51**(11), 6854–6865 (2022). <https://doi.org/10.1080/03610918.2020.1819320>
- Teimouri, M., Nadarajah, S.: Bias corrected MLEs for the Weibull distribution based on records. *Stat Methodol* **13**, 12–24 (2013). <https://doi.org/10.1016/j.stamet.2013.01.001>
- Watkins, A.J., John, A.M.: On the expected Fisher information for the Weibull distribution with type I censored data. *Int. J. Pure Appl. Math.* **15**(4), 401–412 (2004)
- Yang, Z., Xie, M.: Efficient estimation of the Weibull shape parameter based on a modified profile likelihood. *J. Stat. Comput. Simul.* **73**(2), 115–123 (2003). <https://doi.org/10.1080/00949650215729>
- Zhu, T.: Statistical inference of Weibull distribution based on generalized progressively hybrid censored data. *J. Comput. Appl. Math.* **371**, 112705 (2020). <https://doi.org/10.1016/j.cam.2019.112705>
- Zhu, T.: Reliability estimation for two-parameter Weibull distribution under block censoring. *Reliab. Eng. Syst. Saf.* **203**, 107071 (2020). <https://doi.org/10.1016/j.res.2020.107071>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.