Article

On the Bias of the Maximum Likelihood Estimators of Parameters of the Weibull Distribution

Man Chen *, Zheng Zhang and Cen Cui

Department of Mathematics, School of Science, Beijing Jiaotong University, Beijing 100044, China; 13271092@bjtu.edu.cn (Z.Z.); 13271026@bjtu.edu.cn (C.C.)
* Correspondence: 13271049@bjtu.edu.cn

Academic Editor: Fazal M. Mahomed
Received: 18 September 2016; Accepted: 20 January 2016; Published: 16 February 2016

Abstract: Usually, the parameters of a Weibull distribution are estimated by maximum likelihood estimation. To reduce the biases of the maximum likelihood estimators (MLEs) of two-parameter Weibull distributions, we propose analytic bias-corrected MLEs. Two other common estimators of Weibull distributions, least-squares estimators and percentiles estimators, are also introduced. Based on a comparison of their performances in the simulation study, we strongly recommend the analytic bias-corrected MLEs for the parameters of Weibull distributions, especially when the sample size is small.

Keywords: Weibull distribution; maximum likelihood estimators; bias correction; bootstrap; percentiles; least squares estimators

1. Introduction

We investigate several different estimations of the parameters of the two-parameter Weibull distribution, and analyze two different methods to reduce the biases of maximum likelihood estimators (MLEs).

The following formula is the cumulative distribution function of a Weibull distribution:

$$F(x; k, \lambda) = \begin{cases} 1 - e^{-\left(\frac{x}{\lambda}\right)^k}, & x > 0, \lambda > 0, k > 0 \\ 0, & x \leq 0 \end{cases}$$  (1)

and its probability density function is:

$$f(x; k, \lambda) = \begin{cases} \left(\frac{1}{\lambda}\right)kx^{k-1}e^{-\left(\frac{x}{\lambda}\right)^k}, & x > 0, \lambda > 0, k > 0 \\ 0, & x \leq 0 \end{cases}$$  (2)

The Weibull distribution was first identified by Fréchet in 1927, and it was named after Waloddi Weibull, a Swedish mathematician who described this distribution in detail in 1951. Rammler [1] used it to explain a particle size distribution.

The Weibull distribution is not a single distribution, but is in fact a family of distributions, because of the distribution having the shape parameter. The shape parameter enables Weibull distributions to take on various shapes, contingent on the value of the shape parameter. Those distributions are particularly useful in modeling applications since they are flexible enough to model various datasets.

The Weibull distribution has a relatively simple form. Nevertheless, the shape parameter enables the Weibull to assume a variety of shapes. This combination of simplicity and flexibility in the shape of the Weibull distribution makes it an effective distributional model in reliability applications. The ability
to model various distributional shapes by using a relatively simple distributional form is possible with many other distributional families, as well.

The Weibull distribution is associated with various other distributions. For instance, it is well known that a Weibull distribution contains the exponential distribution (when $k = 1$) and the Rayleigh distribution (when $k = 2$). Recently, Ling and Giles [2] studied the Rayleigh distribution and the bias adjustment of the Rayleigh distribution. The Weibull distribution is a special condition of the generalized extreme value distribution, which was proposed by Fréchet [3]. This closely-related Fréchet distribution has the probability density function:

$$f_{\text{Frechet}}(x; k, \lambda) = \begin{cases} \frac{k}{\lambda} (\frac{x}{\lambda})^{-1-k} e^{-(\frac{x}{\lambda})^{-k}} & x > 0, \lambda > 0, k > 0, \\ 0 & x \leq 0. \end{cases} \tag{3}$$

Actually, $f_{\text{Frechet}}(x; k, \lambda) = -f_{\text{Weibull}}(x; -k, \lambda)$. A Weibull distribution is also characterized in terms of a uniform distribution; when $U$ is uniformly distributed on $(0, 1)$, then the random variable $\lambda (-\ln(U))^{1/k}$ is Weibull distributed with parameters $k$ and $\lambda$.

The Weibull distribution has been used in numerous fields, such as production, industry, data analysis, reliability in devices, materials testing and environmental assessment. This distribution was studied and used in an aircraft system to investigate operational reliability by Kaltschmidt et al. [4]. Nikolaj [5] showed that this distribution can be used in oil pollution examination. Singh [6] studied the Weibull distribution and worked in hydrology by using this distribution. Aarset [7] applied a Weibull distribution to test failure times of devices, and the data were further analyzed and modified by Lai et al. [8]. It was used in breakdown voltage estimation by Hirose [9], and Fabiani [10] used it to test electrical breakdown of insulating materials. It was applied to censored data by Ghitany et al. [11]. In the study of wind energy, the Weibull distribution was also applied by Akdağ et al. [12].

Additionally, we have also considered the exponential distribution. When the value of $k$ is one, the Weibull distribution becomes the exponential distribution, which is studied widely. The cumulative distribution function of the exponential distribution is:

$$F(x; k, \lambda) = \begin{cases} 1 - e^{-\lambda x} & x > 0, \lambda > 0, \\ 0 & x \leq 0, \end{cases} \tag{4}$$

and its probability density is:

$$f(x; k, \lambda) = \begin{cases} \lambda e^{-\lambda x} & x > 0, \lambda > 0, \\ 0 & x \leq 0. \end{cases} \tag{5}$$

Maximum-likelihood estimators of the Weibull distribution are commonly used. However, they are biased, and thus, in this paper, we propose analytic bias-corrected MLEs for the two-parameter Weibull distribution. The analytic bias-corrected MLEs are considered in Section 2 and compared with the other bias-corrected MLEs, which are also known as bootstrap bias-corrected MLEs. In Sections 3 and 4, we introduce two other commonly-used estimators. Monte Carlo experiments are conducted for the simulation study. Tabular and graphical representations of the results are presented in this section. The results indicate that the analytic bias-corrected estimators are superior to not only traditional bootstrap bias-corrected estimators, but also the other two commonly-used estimators. Thus, the analytic bias-corrected estimators should be highly recommended. Finally, we apply these estimators to an actual situation and draw conclusions.
2. Maximum Likelihood Estimators and Different Bias-Corrected Maximum Likelihood Estimators

2.1. Maximum Likelihood Estimators

Suppose the sample \((x_1, x_2, ..., x_n)\) is generated from the log-normal distribution. The likelihood function of the Weibull distribution is:

\[
L = \left( \frac{1}{\lambda} \right)^n \left( \prod_{i=1}^{n} x_i \right)^{k-1} e^{-\lambda^{-k} \sum_{i=1}^{n} x_i^k}
\]  

(6)

and the formula for the log-likelihood function is:

\[
l = \log L = -kn \log \lambda + n \log k + (k - 1) \sum_{i=1}^{n} \log x_i - \lambda^{-k} \sum_{i=1}^{n} x_i^k
\]  

(7)

The MLEs can be obtained by solving the following equations:

\[
\frac{\partial l}{\partial \lambda} = -\frac{kn}{\lambda} + k\lambda^{-k-1} \sum_{i=1}^{n} x_i^k = 0
\]

(8)

\[
\frac{\partial l}{\partial k} = -n \log \lambda + \frac{n}{k} + \sum_{i=1}^{n} \log x_i - \lambda^{-k} \sum_{i=1}^{n} x_i^k \log x_i + \lambda^{-k}(\log \lambda) \sum_{i=1}^{n} x_i^k = 0
\]

(9)

The Newton-Raphson method is used to solve these two equations, and solutions \(\hat{k}\) and \(\hat{\lambda}\) were obtained.

2.2. Analytic Bias-Corrected Maximum Likelihood Estimators

Suppose \(l(\sigma)\) is the log-likelihood function of the Weibull distribution based on a sample size \(n\). The joint cumulants of the derivatives of the log-likelihood function are:

\[
h_{ij} = E[\frac{\partial^2 l}{\partial \sigma_i \partial \sigma_j}], i, j = 1, 2, ..., q
\]

(10)

\[
h_{ijm} = E[\frac{\partial^3 l}{\partial \sigma_i \partial \sigma_j \partial \sigma_m}], i, j, m = 1, 2, ..., q
\]

(11)

\[
h_{ij,m} = E[\frac{\partial^2 l}{\partial \sigma_i \partial \sigma_j} \frac{\partial l}{\partial \sigma_m}], i, j, m = 1, 2, ..., q
\]

(12)

\[
h_{ij}^{(m)} = \frac{\partial h_{ij}}{\partial \sigma_m}, i, j, l = 1, 2, ..., q
\]

(13)

In this part, the function \(l(\sigma)\) is regular with respect to all derivatives up to the third order. We assume that Equations (10)–(13) are \(O(n)\). Cox and Snell [13] showed that for sample data that are independent, but not always identically distributed, the bias of the \(s\)-th element of the MLE of \(\sigma\) is:

\[
\text{Bias}(\hat{\sigma}_s) = \sum_{i=1}^{q} \sum_{j=1}^{q} \sum_{m=1}^{q} h^{ij} h^{ij} [(1/2)h_{ijm} + h_{ij,m}] + O(n^{-2}), i, j, m = 1, 2, ..., q
\]

(14)

where \(h^{ij}\) is the \((i, j)\)-th element of the inverse of Fisher’s information matrix, \(H = \{-h^{ij}\}\). Cordeiro and Klein [14] observed that the bias expression still holds even if the observations are non-independent, proved that all of the \(h\) terms are \(O(n)\) and recommended another form for the expression:

\[
\text{Bias}(\hat{\sigma}_s) = \sum_{i=1}^{q} h^{ij} \sum_{j=1}^{q} h^{ij} h^{ij} [(1/2)h_{ijm} + h_{ij,m}] + O(n^{-2}), i, j, m = 1, 2, ..., q
\]

(15)
In Equation (15), we define:
\[ a_{ij}^{(m)} = h_{ij}^{(m)} - \frac{1}{2} h_{ijm}, i, j, m = 1, 2, ..., q \]  
and (15) can be written as:
\[ \text{Bias}(\hat{\sigma}) = H^{-1} \text{Avec}(H^{-1}) + O(n^{-2}) \]  
where \( \text{vec}(H^{-1}) \) is a transformation of the matrix \( H^{-1} \) into a vector form, and the matrix \( A \) is:
\[ A^{(m)} = \{a_{ij}^{(m)}\}; i, j, m = 1, 2, ..., q \]  
\[ A = [A^{(1)} | A^{(2)} | ... | A^{(q)}] \]

Then, the bias-corrected MLEs are:
\[ \text{Bias}\left(\hat{\lambda}_{\text{analytic-corrected}} / \hat{k}_{\text{analytic-corrected}}\right) = \hat{\sigma} - H^{-1} \text{Avec}(H^{-1}) \]  
where \( \hat{H} = (H)_{[\hat{\theta}]} \) and \( \hat{A} = (A)_{[\hat{\theta}]} \). The bias of \( \hat{\sigma} \) can be proven to be \( O(n^{-2}) \). The values of the elements of \( \hat{\sigma} \) can be obtained based on the maximum likelihood equations.

This method was applied to the generalized Pareto distribution by Giles et al. [15] successfully, and Min Wang [16] used it for the weighted Lindley distribution with Wentao Wang.

Specific calculation procedures of the bias-corrected MLEs of the parameters of the Weibull distribution are listed in Appendix A.

2.3. Bootstrap Bias-Corrected Estimators

In this section, the bootstrap bias-corrected estimators of the Weibull distribution are considered, which were suggested by Mackinnon and Smith [17].

The solutions \( \hat{k} \) and \( \hat{\lambda} \) are the particular values of estimators based on MLEs, and we regard \( \hat{k} \) and \( \hat{\lambda} \) as the particular parameters of the log–normal distribution, from which the \( N_B \) replicative samples are generated. Thus, the corresponding estimates \( (\hat{k}_1, \hat{k}_2, ..., \hat{k}_n) \) and \( (\hat{\lambda}_1, \hat{\lambda}_2, ..., \hat{\lambda}_n) \) are obtained.

The bootstrap bias-corrected parameters are:
\[ \hat{k} = 2\hat{k} - (1/N_B) \sum_{j=1}^{N_B} \hat{k}_j, \quad \hat{\lambda} = 2\hat{\lambda} - (1/N_B) \sum_{j=1}^{N_B} \hat{\lambda}_j \]
(the parameters \( \hat{k}_j \) and \( \hat{\lambda}_j \) are the MLEs of \( k \) and \( \lambda \) from the \( j \)-th sample of the \( N_B \) replication samples).

This method is fairly effective and is worth applying to various situations, which causes that many experts have investigated it. Ferrari and Cribari-Neto [18] explored the relationship between bias correction of maximum likelihood estimators through the bootstrap method, and the bootstrap method was presented by Ho and Fernandes da Silva [19] to make improvements in the MLE of Mean Time to Failure (MTTF) and \( p \)-quantiles. Focarelli [20] used the bootstrap bias-correction procedure to estimate long-run relationships from dynamic panels, with an application to money demand in the euro area.

3. Percentile Estimators

The estimators based on percentiles are also investigated. This approach was first studied by Kao [21,22] and then was explored by Mann, Schafer and Singpurwalla [23]. This method is suitable for the Weibull distribution.

Let us consider the situation where both the parameters \( \lambda \) and \( k \) are unknown. From the cumulative distribution function (1), we can obtain:
\[ \lambda[- \log(1 - F(x_i; k, \lambda))]^{1/k} = x_i \]
Suppose that \((x_1, x_2, \ldots, x_n)\) is a random sample of data, and the \((x_{(1)}, x_{(2)}, \ldots, x_{(n)})\) is the ordered observation. Let \(p_i\) be the estimate of \(F(x_{(i)}; k, \lambda)\), then we can obtain estimates of \(\lambda\) and \(k\) by minimizing the equation:

\[
\sum_{i=1}^{n} [x_i - \left(\frac{\log(1 - F(x_{(i)}; k, \lambda))}{(1/k)}\right)]^2
\]

(22)

Generally, some \(p_i\)'s are used to be the estimates of the \(F(x_{(i)}; k, \lambda)\), and the \(p_i = (i/n + 1)\) is most frequently used, because \((i/(n + 1))\) is the expected value of \(F(x_{(i)}; k, \lambda)\). In this article, we also use this \(p_i\).

Equation (22) is a non-linear function. To obtain estimates of the two parameters, we can solve (22) by using some non-linear regression techniques. These estimators are called percentile estimators, which were also applied and introduced by Gupta and Kundu [24].

4. Least Squares Estimators

In this section, we introduce the expressions of another estimation. This method was first applied by Swain, Venkatraman and Wilson [25] to estimate the parameters of the beta distribution, and it can be utilized in other cases, as well. Suppose that \((x_1, x_2, \ldots, x_n)\) is a random sample of size \(n\) from a distribution whose distribution function is \(F(x)\) and that \((x_{(1)}, x_{(2)}, \ldots, x_{(n)})\) denotes the ordered set of sample observations. Jonson, Kotz and Balakrishnan [26] proposed expressions for the expectations and variances,

\[
E(F(x_{(i)})) = \frac{i}{n+1}, \quad i = 1, 2, \ldots, n
\]

(23)

The least squares estimators can be obtained by minimizing:

\[
\sum_{i=1}^{n} \left[F(x_{(i)}) - \frac{i}{n+1}\right]^2
\]

(24)

We call this estimator as least squares estimator.

Therefore, in the case of the Weibull distribution, the least squares estimators of \(k\) and \(\lambda\) are denoted by \(k_{Lse}\) and \(\lambda_{Lse}\), respectively. These values are obtained by minimizing:

\[
\sum_{i=1}^{n} \left[1 - e^{-\frac{x_{(i)}}{\lambda}}\right]^2 - \frac{i}{n+1}\right]^2
\]

(25)

5. Simulation Study

5.1. Procedures

In this part, the Monte Carlo experiment is described, and the simulation results are analyzed. The R 3.2.2 software [27] routines for log-normal random variates are used to generate the sample from the Weibull distribution. Then, the MLE, percentile estimation and least squares estimation can be realized by the non-linear minimization function. In the R software, the non-linear minimization function can solve a minimization problem using a Newton-type algorithm. After substituting the MLEs into the corrected formula mentioned previously, we can obtain the analytic bias-corrected estimators. The bootstrap bias-corrected estimators will be acquired after conducting the replicative MLE experiments.

5.2. Analysis of the Results

5.2.1. Comparison of Maximum Likelihood Estimators and Bias-Corrected Maximum Likelihood Estimators

Comparing the theoretical performances of the different estimators mentioned above is not convincing. Therefore, we performed extensive simulations to compare the performances of the
different methods, mainly with respect to their real absolute percentage biases and percentage mean squared errors (MSEs) for different sample sizes. The percentage MSE is defined as $100 \times (\text{MSE}(\hat{k})/k^2)$, and the absolute percentage bias is defined as $100 \times |(\text{mean}(\hat{k}) - k)/k|$. 

In each section of this experiment, 20,000 Monte Carlo replications are performed, and 10,000 bootstrap samples are used for every replication.

We also investigate general Weibull distributions with some common parameter values. The exponential distribution with particular parameter value ($k = 1$) is also investigated here. We select parameters values for the Weibull distribution ($k = 1$) that are consistent with the data generating process being an exponential distribution.

Absolute percentage biases and MSEs of those estimators for the exponential distribution are presented in Table 1, and those for the Weibull distribution are presented in Table 2.

These tables show that the percentage biases of the analytic bias-corrected MLEs are much smaller than those of the MLEs. Meanwhile, both the analytic bias correction and bootstrap bias correction perform well. However, the Cox-Snell bias correction is more effective than the bootstrap reduction, especially with the modest sample size.

The absolute values of the percentage biases tend to decline as the sample size increases, and the percentage MSEs of the Cox-Snell estimators are smaller than those of the bootstrap estimators, except in a few cases. The MSEs and biases of the analytic bias-corrective estimators are smaller than those of the other estimators.

<table>
<thead>
<tr>
<th>Table 1. Percentage bias and mean squared errors (MSEs) for the exponential distribution.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>$k = 1, \lambda = 4.2$</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>50</td>
</tr>
<tr>
<td>90</td>
</tr>
</tbody>
</table>

| $k = 1, \lambda = 6$ | | | | | |
| 10 | 3.62346 (2.43225) | 4.52579 (2.32371) | 1.93752 (2.37691) | 2.12836 (2.84987) | 1.83741 (2.23637) |
| 20 | 3.48225 (2.34162) | 4.44257 (2.26101) | 1.84247 (2.23775) | 1.93243 (2.76782) | 1.73236 (2.19686) |
| 50 | 2.48276 (1.83342) | 3.33175 (1.71529) | 0.99447 (1.75702) | 1.01764 (2.37791) | 0.96848 (1.67906) |
| 90 | 2.24721 (0.88253) | 2.92384 (0.89496) | 0.83833 (0.89241) | 0.99847 (1.04731) | 0.81751 (0.81043) |

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\hat{\lambda}$</th>
<th>$\lambda_{\text{pr}}$</th>
<th>$\lambda$</th>
<th>$\lambda_{\text{lar}}$</th>
<th>$\hat{\lambda}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k = 1, \lambda = 4.2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.32264 (0.24035)</td>
<td>0.28236 (0.59336)</td>
<td>0.623731 (0.24071)</td>
<td>0.18537 (0.30108)</td>
<td>0.17435 (0.23218)</td>
</tr>
<tr>
<td>20</td>
<td>0.31451 (0.22341)</td>
<td>0.27293 (0.23809)</td>
<td>0.42386 (0.23353)</td>
<td>0.17673 (0.29049)</td>
<td>0.16212 (0.22234)</td>
</tr>
<tr>
<td>50</td>
<td>0.27232 (0.18334)</td>
<td>0.22251 (0.18124)</td>
<td>0.19474 (0.18364)</td>
<td>0.12734 (0.24157)</td>
<td>0.11615 (0.17787)</td>
</tr>
<tr>
<td>90</td>
<td>0.22582 (0.13205)</td>
<td>0.21253 (0.13221)</td>
<td>0.16342 (0.13201)</td>
<td>0.09862 (0.15193)</td>
<td>0.06783 (0.12996)</td>
</tr>
</tbody>
</table>

| $k = 1, \lambda = 6$ | | | | | |
| 10 | 0.34364 (0.22826) | 0.36204 (0.23118) | 0.19424 (0.22719) | 0.35411 (0.23386) | 0.13571 (0.22105) |
| 20 | 0.21527 (0.21376) | 0.34482 (0.21096) | 0.18895 (0.21855) | 0.33206 (0.22596) | 0.12556 (0.21224) |
| 50 | 0.20738 (0.19892) | 0.28401 (0.20515) | 0.17804 (0.19497) | 0.26589 (0.20787) | 0.12404 (0.19272) |
| 90 | 0.19278 (0.09642) | 0.26276 (0.09994) | 0.15497 (0.09036) | 0.24259 (0.10425) | 0.11285 (0.08902) |

In these tables $\hat{\lambda}, \lambda_{\text{pr}}, \lambda, \lambda_{\text{lar}}, \hat{\lambda}$ denote MLEs, percentile, bootstrap, least squares and analytic bias corrected estimators of $\lambda$. 

---

(continued)
5.2.2. Analysis for All Estimators

The absolute values of the percentage biases of all five estimators are shown in the Figures 1–3. Some of the points are clear based on the experimental results. For all methods, as the sample size increases, the biases decrease.
Figure 1. % Bias of the parameters. (a) % Bias of $k (k = 4.5, \lambda = 2.5)$; (b) % bias of $k (k = 3, \lambda = 5)$; (c) % bias of $\lambda (k = 4.5, \lambda = 2.5)$; (d) % bias of $\lambda (k = 3, \lambda = 5)$. In these figures, pere, lse, mle, bmle and mle denote the related curves of percentile estimators, least squares estimators, MLEs, bootstrap corrected MLEs and analytic bias-corrected estimators.
Figure 2. % Bias of the parameters. (a) % bias of $k (k = 5.8, \lambda = 4)$; (b) % bias of $\lambda (k = 5.8, \lambda = 4)$; (c) % bias of $k (k = 3, \lambda = 2.6)$; (d) % bias of $\lambda (k = 3, \lambda = 2.6)$. 
Figure 3. % Bias of the parameters. (a) % bias of $k$ ($k = 16, \lambda = 83.8$); (b) % bias of $\lambda$ ($k = 16, \lambda = 83.8$).

For the most part, the biases of the percentile estimators are larger than those of other estimators, except in specific cases, as shown in Figures 1a and 2a. The bootstrap bias-corrected MLEs and analytic bias-corrected MLEs are generally both superior to MLEs. Furthermore, the analytic bias-corrected MLEs are superior to the other estimators.

The performance of the least squares estimators is uncertain. Sometimes, the least squares estimators perform better than the MLEs, whereas at other times, the latter’s performance stands out. Clearly, the biases of the least squares estimators become extremely similar to those of the bias-corrected MLEs as the sample size increases.

However, the MSEs of the least squares estimators are often larger than those of the other estimators. Meanwhile, the MSEs of the MLEs, analytic bias-corrected MLEs, bootstrap bias-corrected MLEs and percentile estimators are close, as evidenced by the results shown in Figures 4–6.
Figure 4. % MSEs of the parameters. (a) % MSE of $k$ ($k = 4.5, \lambda = 2.5$); (b) % MSE of $k$ ($k = 3, \lambda = 5$); (c) % MSE of $\lambda$ ($k = 4.5, \lambda = 2.5$); (d) % MSE of $\lambda$ ($k = 3, \lambda = 5$).

Figure 5. Cont.
These results indicate that the analytic bias-corrected MLEs can effectively reduce the biases of MLEs. Moreover, to obtain the least biased estimators of the parameters of the Weibull distribution, the analytic bias-corrected MLEs are much more appropriate than the other common estimators mentioned above.

6. Real Illustrative Example

In this section, the practical performance of the Cox–Snell bias correction is explored, and we apply this estimation to real data (i.e., the data provided by Aarest [7]). Aarest tested the times to failure of 50 devices.

We fit the Weibull distribution to these data using MLEs, percentile estimators and least squares estimators. Then, we obtain the bias-corrected parameters by solving the analytic expressions mentioned previously.

Table 3 reports the failure time of 50 devices.
Table 3. Calculated failure time of 50 devices.

<table>
<thead>
<tr>
<th>No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.100</td>
<td>7.00</td>
<td>36.00</td>
<td>67.00</td>
<td>84.00</td>
<td>0.200</td>
<td>11.00</td>
<td>40.00</td>
<td>67.00</td>
<td>84.00</td>
<td>1.00</td>
<td>12.00</td>
<td>45.00</td>
</tr>
<tr>
<td>No.</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>67.00</td>
<td>84.00</td>
<td>1.00</td>
<td>18.00</td>
<td>46.00</td>
<td>67.00</td>
<td>85.00</td>
<td>1.00</td>
<td>18.00</td>
<td>47.00</td>
<td>72.00</td>
<td>85.00</td>
<td>1.00</td>
</tr>
<tr>
<td>No.</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>38</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>18.00</td>
<td>50.00</td>
<td>75.00</td>
<td>85.00</td>
<td>1.00</td>
<td>18.00</td>
<td>55.00</td>
<td>79.00</td>
<td>85.00</td>
<td>2.00</td>
<td>18.00</td>
<td>60.00</td>
<td>82.00</td>
</tr>
<tr>
<td>No.</td>
<td>40</td>
<td>41</td>
<td>42</td>
<td>43</td>
<td>44</td>
<td>45</td>
<td>46</td>
<td>47</td>
<td>48</td>
<td>49</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>85.00</td>
<td>3.00</td>
<td>21.00</td>
<td>63.00</td>
<td>82.00</td>
<td>86.00</td>
<td>6.00</td>
<td>32.00</td>
<td>63.00</td>
<td>83.00</td>
<td>86.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4 reports the maximum likelihood estimates, bootstrap maximum likelihood estimates, analytic bias-corrected maximum likelihood estimates, percentile estimates and least squares estimates after fitting the Weibull distribution to these data.

Table 4. Five estimates of Parameters.

<table>
<thead>
<tr>
<th>$\hat{k}$</th>
<th>$k_{Per}$</th>
<th>$\hat{k}$</th>
<th>$k_{Lse}$</th>
<th>$\hat{k}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\lambda}$</td>
<td>$\lambda_{Per}$</td>
<td>$\hat{\lambda}$</td>
<td>$\lambda_{Lse}$</td>
<td>$\lambda_k$</td>
</tr>
<tr>
<td>83.64601</td>
<td>83.48114</td>
<td>83.655114</td>
<td>83.58114</td>
<td>83.669204</td>
</tr>
</tbody>
</table>

Based on the MLE of the data shown in Table 3, Figure 7 presents the fitted probability density function curve of Weibull distribution.

Figure 7. Fitted Weibull’s probability density function (failure time data).

For these data, the analytic bias correction decreases the height of the Weibull distribution density, and the shape of the density becomes slightly "fatter" and "shorter". Thus, the distribution is more uniform.

7. Conclusions

Usually, the two parameters of a Weibull distribution are estimated using maximum likelihood estimation, which is realized under non-linear and first-order conditions. Nevertheless, after comparing
all of the methods, it is clear that in terms of the minimization, analytic bias-corrected MLEs exhibited the best performance in almost all cases. Therefore, the analytic bias-corrected MLEs of the two-parameter Weibull distribution are proposed to more effectively reduce the biases of the MLEs.

We conduct a Monte Carlo experiment, which is described in the simulation study section. According to the MSEs and absolute percentage biases of the estimates, we explore the effectiveness of these estimations. Clearly, the biases of the analytic bias-corrected MLEs are smaller than those of the other estimators. Consequently, we can conclude that the analytic bias-corrected MLEs are superior to the others tested here.

All in all, the analytic bias-corrected MLEs discussed in this paper are worth popularizing for utilization on actual data. They should be highly recommended, especially for small sample sizes.

**Author Contributions:** Man Chen provided main thoughts, organized language, built programming models and did later modifications. Zheng Zhang did investigations, provided practical data, proposed related ideas and polished the language. Cen Cui recorded data and took part in doing programming models and analyzing data.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Appendix A. Specific Calculation**

Here are the specific calculation procedures of the bias-corrected MLEs of the parameters of the Weibull distribution.

From the previous part, we can know that the likelihood function of the two-parameter Weibull distribution is:

\[
L = \left(\frac{1}{\lambda}\right)^{kn} k^n \left(\prod_{i=1}^{n} x_i\right)^{k-1} e^{-\lambda x} \sum_{i=1}^{n} x_i^k
\]  

(A1)

and the log-likelihood function is:

\[
l = \log L = -kn \log \lambda + n \log k + (k - 1) \sum_{i=1}^{n} \log x_i - \lambda \sum_{i=1}^{n} x_i^k
\]  

(A2)

We obtain these higher-order derivatives:

\[
\frac{\partial l}{\partial \lambda} = -\frac{kn}{\lambda} + k\lambda^{-k-1}\sum_{i=1}^{n} x_i^k
\]  

(A3)

\[
\frac{\partial l}{\partial k} = -n \log \lambda + n + \sum_{i=1}^{n} \log x_i - \lambda^{-k} \sum_{i=1}^{n} x_i^k \log x_i + \lambda^{-k}(\log \lambda) \sum_{i=1}^{n} x_i^k
\]  

(A4)

\[
\frac{\partial^2 l}{\partial \lambda^2} = \frac{kn}{\lambda^2} - k(k+1)\lambda^{-k-2}\sum_{i=1}^{n} x_i^k
\]  

(A5)

\[
\frac{\partial^2 l}{\partial k^2} = \frac{-n}{k^2} + 2\lambda^{-k}(\log \lambda) \sum_{i=1}^{n} x_i^k \log x_i - \lambda^{-k} \sum_{i=1}^{n} x_i^k (\log x_i)^2 - \lambda^{-k}(\log \lambda)^2 \sum_{i=1}^{n} x_i^k
\]  

(A6)

\[
\frac{\partial^2 l}{\partial \lambda \partial k} = \frac{\partial^2 l}{\partial k \partial \lambda} = -\frac{n}{\lambda} + k\lambda^{-k-1}\sum_{i=1}^{n} x_i^k \log x_i + (\lambda^{-k-1} - k\lambda^{-k-1}(\log \lambda)) \sum_{i=1}^{n} x_i^k
\]  

(A7)

\[
\frac{\partial^3 l}{\partial \lambda^3} = \frac{-2kn}{\lambda^3} + k(k+1)(k+2)\lambda^{-k-3}\sum_{i=1}^{n} x_i^k
\]  

(A8)
\[
\frac{\partial^3 I}{\partial k^3} = \frac{2n}{k^3} - 3\lambda^{-k}(\log \lambda)^2 \sum_{i=1}^{n} x_i^k \log x_i + 3\lambda^{-k}(\log \lambda) \sum_{i=1}^{n} x_i^k (\log x_i)^2
\]
\[
-\lambda^{-k} \sum_{i=1}^{n} x_i^k (\log x_i)^3 + \lambda^{-k}(\log \lambda)^3 \sum_{i=1}^{n} x_i^k
\]
\[
(A9)
\]
\[
\frac{\partial^3 I}{\partial \lambda^2 \partial k} = \frac{\partial^3 I}{\partial \lambda \partial k \partial \lambda} = \frac{n}{\lambda^2} + (-2k+1)\lambda^{-k-2} + k(k+1)\lambda^{-k-2}(\log \lambda) \sum_{i=1}^{n} x_i^k
\]
\[
- k(k+1)\lambda^{-k-2} \sum_{i=1}^{n} x_i^k \log \lambda
\]\n\[
(A10)
\]
\[
\frac{\partial^3 I}{\partial k^2 \partial \lambda} = \frac{\partial^3 I}{\partial k \partial \lambda \partial k} = 2(\lambda^{-k-1} - k\lambda^{-k-1}(\log \lambda)) \sum_{i=1}^{n} x_i^k \log x_i + k\lambda^{-k-1} \sum_{i=1}^{n} x_i^k (\log x_i)^2
\]
\[
+ (-2\lambda^{-k-1} \log \lambda + k\lambda^{-k-1}(\log \lambda)^2) \sum_{i=1}^{n} x_i^k
\]\n\[
(A11)
\]

and then, we can obtain the following formulas:

\[
h_{11} = E(\frac{\partial^2 I}{\partial \lambda^2})
\]
\[
h_{12} = k_{21} = E(\frac{\partial^2 I}{\partial \lambda \partial k})
\]
\[
h_{22} = E(\frac{\partial^2 I}{\partial k^2})
\]
\[
(A12)
\]
\[
(A13)
\]
\[
(A14)
\]

The following equations are also easy to get:

\[
E(\sum_{i=1}^{n} x_i^k) = n \int_{0}^{\infty} x^k \left( \frac{1}{\lambda} \right)^k k x^{k-1} e^{-\left( \frac{x}{\lambda} \right)^k} dx = n\lambda^k
\]
\[
(A15)
\]
\[
E(\sum_{i=1}^{n} x_i^k \log x_i) = n \int_{0}^{\infty} x^k \log x \left( \frac{1}{\lambda} \right)^k k x^{k-1} e^{-\left( \frac{x}{\lambda} \right)^k} dx = n\lambda^k (k \log (\lambda) + \frac{\gamma+1}{k})
\]
\[
(A16)
\]
\[
E(\sum_{i=1}^{n} x_i^k (\log x_i)^2) = n \int_{0}^{\infty} x^k (\log x)^2 \left( \frac{1}{\lambda} \right)^k k x^{k-1} e^{-\left( \frac{x}{\lambda} \right)^k} dx
\]
\[
= \frac{n\lambda^k (6k \log (\lambda) (k \log (\lambda) - 2\gamma + 2) + 6\gamma (\gamma - 2) + \pi^2)}{6k^2}
\]
\[
(A17)
\]

and the following equation is required for the next calculation part:

\[
E(\sum_{i=1}^{n} x_i^k (\log x_i)^3) = n \int_{0}^{\infty} x^k (\log x)^3 \left( \frac{1}{\lambda} \right)^k k x^{k-1} e^{-\left( \frac{x}{\lambda} \right)^k} dx
\]
\[
= \frac{n\lambda^k (k \log (\lambda) (2k \log (\lambda) (k \log (\lambda) - 3\gamma + 3) + 6\gamma (\gamma - 2) + \pi^2) - 4\zeta (3) - \gamma (2\gamma (\gamma - 3) + \pi^2) + \pi^2)}{2k^3}
\]
\[
(A18)
\]
Where \( \gamma \) is a constant called eulergamma and its value is about 0.57721; and \( \zeta(3) \) is another constant, and its value is about 1.20206; then, these results can be obtained:

\[
\begin{align*}
    h_{11} &= -\frac{k^2 n}{\lambda^2}, \\
    h_{12} &= h_{21} = \frac{(1 - \gamma) n}{\lambda}, \\
    h_{22} &= -\frac{(6\gamma^2 - 12\gamma + \pi^2 + 6) n}{6k^2}
\end{align*}
\]

Then, based on Equation (16), a series of results is obtained:

\[
\begin{align*}
    h_{11}^1 &= \frac{\partial h_{11}}{\partial \lambda} = \frac{2k^2 n}{\lambda^3}, & (A19) \\
    h_{12}^1 = h_{21}^1 &= \frac{\partial h_{12}}{\partial \lambda} = \frac{(\gamma - 1) n}{\lambda^2}, & (A20) \\
    h_{22}^1 &= \frac{\partial h_{22}}{\partial \lambda} = 0, & (A21) \\
    h_{11}^2 &= \frac{\partial h_{11}}{\partial \gamma} = -\frac{2k n}{\lambda^2}, & (A22) \\
    h_{12}^2 = h_{21}^2 &= \frac{\partial h_{12}}{\partial \gamma} = 0, & (A23) \\
    h_{22}^2 &= \frac{\partial h_{22}}{\partial \gamma} = \frac{(6 - 12\gamma + 6\gamma^2 + \pi^2) n}{3k^3}, & (A24) \\
    h_{111} &= \frac{k^2 (k + 3) n}{\lambda^3}, & (A25) \\
    h_{122} = h_{212} = h_{221} &= \frac{(12 - 24\gamma + 6\gamma^2 + \pi^2) n}{6k\lambda}, & (A26) \\
    h_{222} &= \frac{n (4\zeta(3) + 2\gamma^3 - 6\gamma^2 + \gamma\pi^2 - \pi^2 + 4)}{2k^3}. & (A27)
\end{align*}
\]

then:

\[
\begin{align*}
    a_{11}^{(1)} &= h_{11}^1 - \frac{1}{2} h_{111} = \frac{(k - 1)k^2 n}{2\lambda^3}, & (A28) \\
    a_{12}^{(1)} &= h_{12}^1 - \frac{1}{2} h_{121} = \frac{-\gamma k + 3k + \gamma - 1) n}{2k^2}, & (A29) \\
    a_{21}^{(1)} &= h_{21}^1 - \frac{1}{2} h_{211} = \frac{-\gamma k + 3k + \gamma - 1) n}{2k^2}, & (A30) \\
    a_{22}^{(1)} &= h_{22}^1 - \frac{1}{2} h_{221} = \frac{(12 - 24\gamma + 6\gamma^2 + \pi^2) n}{12k\lambda}, & (A31) \\
    a_{11}^{(2)} &= h_{11}^2 - \frac{1}{2} h_{112} = \frac{(\gamma k + k + \gamma - 1) n}{2k^2}, & (A32) \\
    a_{12}^{(2)} &= h_{12}^2 - \frac{1}{2} h_{122} = \frac{(12 - 24\gamma + 6\gamma^2 + \pi^2) n}{12k\lambda}, & (A33) \\
    a_{21}^{(2)} &= h_{21}^2 - \frac{1}{2} h_{212} = \frac{(12 - 24\gamma + 6\gamma^2 + \pi^2) n}{12k\lambda}, & (A34) \\
    a_{22}^{(2)} &= h_{22}^2 - \frac{1}{2} h_{222} = \frac{n (12\zeta(3) + 3\gamma (8 + \pi^2) + 6\gamma^3 - 30\gamma^2 - 5\pi^2)}{6k^3}. & (A35)
\end{align*}
\]

From the formula \( H_{ij} = \{-h_{ij}\} \), we can know that the information matrix is:
\[ H = \left( \begin{array}{c} \frac{k^2}{n} \lambda^2 - (1-\gamma)n \\ \frac{-(1-\gamma)n}{\lambda} \\ \frac{-(6\gamma - 12\gamma^2 + \pi^2 + 6)n}{6\kappa^2} \end{array} \right) \] (A36)

and:

\[ A = \begin{bmatrix} A^{(1)} \mid A^{(2)} \end{bmatrix} \]

\[ A = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{11}^{(2)} & a_{12}^{(2)} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{21}^{(2)} & a_{22}^{(2)} \end{bmatrix} \]

Using the bias-corrected expressions in the matrix form, we can obtain:

\[ \text{Bias} \left( \begin{bmatrix} \hat{\lambda} \\ \hat{k} \end{bmatrix} \right) = \hat{\sigma} - \tilde{H}^{-1} \tilde{A} \text{vec}(\tilde{H}^{-1}) \] (A37)

The bias-corrected estimates are:

\[ \left( \begin{bmatrix} \hat{\lambda} \\ \hat{k} \end{bmatrix} \right) = \left( \begin{bmatrix} \hat{\lambda} \\ \hat{k} \end{bmatrix} \right) - \text{Bias} \left( \begin{bmatrix} \tilde{\lambda} \\ \tilde{k} \end{bmatrix} \right) \] (A38)

where \( \tilde{H} = H|_{\lambda=\tilde{\lambda},k=\tilde{k}}, \tilde{A} = A|_{\lambda=\tilde{\lambda},k=\tilde{k}} \)

References


27. Murdoch, D. R-3.2.2 for Windows. Available online: https://cran.r-project.org/bin/windows/base/old/3.2.2/ (accessed on 27 August 2015)

© 2017 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).