mathematics

Article

A Note on Weibull Parameter Estimation with Interval Censoring Using the EM Algorithm

Chanseok Park

Applied Statistics Laboratory, Department of Industrial Engineering, Pusan National University, Busan 46241, Republic of Korea; cp@pusan.ac.kr

Abstract: In many engineering applications, it is often the case that the observations are only available in interval form. In this note, by using the expectation-maximization (EM) algorithm, the parameter estimation of the Weibull distribution with interval-censored data is considered. The estimates obtained using the EM algorithm are compared with those obtained using the conventional Newton-type methods, including the Davidon–Fletcher–Powell (DFP) and Berndt–Hall–Hall–Hausman (BHHH) methods. The results indicate that the estimates obtained using the proposed EM method demonstrate superior convergence properties compared to the conventional DFP and BHHH methods. Finally, a numerical study that illustrates the advantages of the proposed method is provided.

Keywords: EM algorithm; censoring; maximum-likelihood estimation

MSC: 62F10; 62N02

Citation: Park, C. A Note on Weibull Parameter Estimation with Interval Censoring Using the EM Algorithm. *Mathematics* **2023**, *11*, 3156. https://doi.org/10.3390/math11143156

Academic Editor: Tzong-Ru Tsai

Received: 9 June 2023
Revised: 10 July 2023
Accepted: 16 July 2023
Published: 18 July 2023

1. Introduction

In various data-driven studies, it is often the case that observations are only available in interval form. For example, in a *go or no-go* inspection scheme, it may be the case that the incoming observations are recorded as pass or no pass according to whether they meet an interval requirement $[a, b]$. For a *go or no-go* inspection, plug gauges are widely used and are usually made in pairs. One side is for the *go* gauge (lower hole gauge) and the other side is for the *no-go* gauge (upper hole gauge). For example, if the *go* gauge can enter the hole, then it is an indication that the hole diameter is above the lower limit. If the *no-go* gauge cannot enter the hole, then it is an indication that the hole diameter is below the upper limit. The part is then accepted because its hole diameter is inside the tolerance band.

Notice that in the *go or no-go* inspection system, if an observation passes the interval inspection scheme, then it is said to be interval-censored at $[a, b]$. This indicates that the value of the observation lies somewhere between the values of a and b. As an illustration of interval-censored observations, a sorting machine example is provided in Figure 1. In addition, if an observation is discarded because it exceeds the allowable limit, then the allowable limit (denoted as d) is lower than the actual but unknown measurement (denoted as X). This type of observation is referred to as right-censored at d (that is, $d \leq X$ or $[d, \infty]$). Similarly, if an observation is discarded because it falls below the allowable limit, it indicates that the allowable limit (denoted as c) is greater than the actual measurement. This type of observation is called left-censored at c (that is, $c \geq X$ or $[0, c]$).

It should be noted that right-censoring is quite common in reliability studies. For example, one may be conducting lifetime tests on light bulbs where most of the lifetimes end up being less than 5000 h, but some of them are still operating after 5000 h. Rather than waiting for the lifetimes of the still-operating light bulbs to end, it will often be the case that these observations are measured as being right-censored with the value of 5000. Therefore, the final sample will consist of both fully observed and right-censored observations.

In the standard case where some observations are fully observed and others are left-, right- or interval-censored, the resulting data set is said to be incomplete. One obvious methodology for dealing with an incomplete data set is to include the fully-observed observations in the subsequent analysis and discard all observations that are in any way censored so that a conventional method that deals only with full observations can be utilized. The obvious disadvantage of the discarding approach is that, if the number of censored observations is large relative to the total number of observations in the sample, then a large amount of information is being discarded. In order to avoid this problem, one can instead use all of the observations and build the likelihood function that needs to be maximized. Clearly, in order to obtain the maximum of the resulting likelihood function, a numerical optimization scheme is required. However, an ordinary numerical method such as Newton-type iterative methods, including the Davidon–Fletcher–Powell (DFP) method [1] and the Berndt–Hall–Hall–Hausman (BHHH) method [2], will often be ineffective when applied to complicated likelihood functions. Furthermore, such a numerical method will often be quite sensitive to the choice of starting values. Recently, the IcenReg R package [3] was developed. This package allows for the analysis of interval censored data using parametric, nonparametric, and semiparametric models, including the Weibull distribution. However, we have observed that the numerical results based on this package often become unstable when the shape parameter of the Weibull distribution is very small or very large.

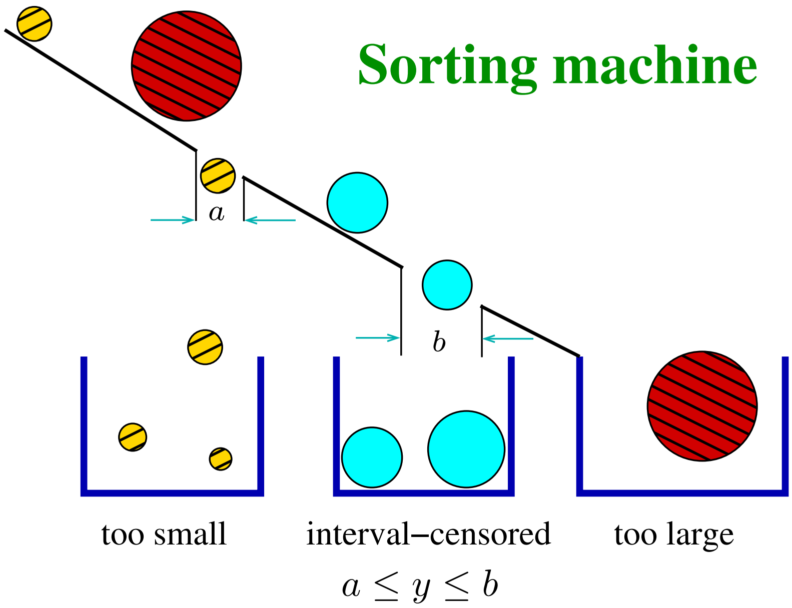

Figure 1. Schematic illustration of a sorting machine.

It is important to emphasize that, when estimating parameters using all of the observations, including those that are censored, the resulting likelihood can be complicated. Therefore, rather than obtaining the maximum-likelihood estimate (MLE) by direct maximization of the likelihood, approximations of the MLE and the best linear unbiased estimate (BLUE) have been provided in the literature. For example, Gupta [4] has provided the BLUE for Type-I and Type-II right-censored samples from a normal distribution. For more details on Type-I and Type-II censoring, one is referred to [5] and Section 7.5 of [6]. Govindarajulu [7] has derived the BLUE for a symmetrically Type-II censored sample from a Laplace distribution, but it can be used only for sample size up to $n = 20$. Balakrishnan [8] and Hassanein et al. [9] also considered the estimation of the scale parameter of the Rayleigh distribution with censored observations. However, the BLUE derived by [9] is limited to the case where the sample sizes are $n = 5, 6, \ldots, 24, 25, 30, 35, \ldots, 45$, and the number of censored observations is limited to $r = 0, 1, \ldots, n-2$. For more details, see Appendix F of [10]. Sultan [11] has given an approximation of the MLE for a Type-II right-censored sample from a normal distribution. Balakrishnan [12] has derived the BLUE for a Type-II right-censored sample for the Laplace distribution. Note that the BLUE requires the coefficients which are tabulated in [12] but the table is provided only for sample sizes up to $n = 20$. In addition, the approximate MLE and the BLUE are not guaranteed to converge to the actual MLE. Another more serious problem is that the methods cited above cannot handle general cases where there are interval-censored observations. The methods cited are restricted to Type-I/Type-II right-censoring or symmetric censoring for sample sizes up to $n = 20$ or $n = 45$.

In order to overcome the issues that come with the analysis of interval-censored data, the use of the EM algorithm [13] for parameter estimation is proposed. However, it is often the case that the implementation of the EM algorithm is quite difficult because the expectation of the log-likelihood in the E-step is generally complex or unavailable in closed form. This problem has been studied by several authors, including Panahi and Asadi [14], Guure et al. [15], Pradhan and Kundu [16], Ferreira and Silva [17], Park [18], Saeed and Elfaki [19], Kurniawan et al. [20], Ameen and Akkash [21], and Almetwally et al. [22].

Panahi and Asadi [14] consider the estimation for Type-II censored samples. This method is limited to the case of right censoring and its estimation is based on the Newton–Raphson iterative procedure. Guure et al. [15] considered the MLE with interval-censored data but obtained it using a simple numerical approximation. Pradhan and Kundu [16] suggested the use of a pseudo-likelihood function in the E-step, Park [18] used the quantile implementation of the E-step, and Saeed and Elfaki [19] used imputation techniques. Kurniawan et al. [20] considered the EM algorithm approach, but they estimated only the shape parameter with Type-II censored observations. Ameen and Akkash [21] obtained the MLE of the three-parameter Weibull with interval-censored data, but they estimated the parameters based on the Newton–Raphson method. Almetwally et al. [22] analyzed progressive Type-II censoring data but the MLE was obtained by utilizing the Newton–Raphson method. Unfortunately, these approaches are not guaranteed to find the maximum of the likelihood function with interval-censored observations. Recently, Ferreira and Silva [17] obtained the MLE of the Weibull using the EM algorithm, but they considered only right-censored data. In this paper, the expectation of the log-likelihood in the E-step is explicitly obtained so that the standard EM algorithm can be used. The use of the standard EM algorithm guarantees that the maximum of the likelihood function with interval-censored observations will be attained.

In this paper, it is assumed that a random sample is generated from the Weibull distribution [23] with the probability density and cumulative distribution functions given below:

$$f(x) = \frac{\kappa x^{\kappa-1}}{\theta^{\kappa}} \exp\left[-\left(\frac{x}{\theta}\right)^{\kappa}\right] \quad \text{and} \quad F(x) = 1 - \exp\left[-\left(\frac{x}{\theta}\right)^{\kappa}\right], \tag{1}$$

where $x > 0$, $\kappa > 0$ and $\theta > 0$.

This paper is organized as follows. In Section 2, the general likelihood function with full and censored observations is provided. In Section 3, the EM algorithm is briefly reviewed. The implementation of the EM algorithm and the parameter estimation method are provided in Section 4. A real-data example illustrating the complexity associated with parameter estimation is provided in Section 5. The paper ends with concluding remarks in Section 6.

2. Likelihood Function for Parameter Estimation in the Interval-Censored Case

As explained in the introduction, a measurement, y_i, is of the interval-censored form if the measurement is only known to fall in an interval:

$$a_i \le y_i \le b_i. \tag{2}$$

The advantage of the interval-censoring formulation in Equation (2) is that it generalizes according to specific censoring conditions. By letting a_i and b_i take on the values $\pm\infty$ in Equation (2), the interval-censoring formulation includes the left-censored and right-censored measurement conditions as special cases. For example, by letting $b_i = \infty$ and a_i equal the censored value, one obtains the right-censored condition. Similarly, by letting

$a_i = -\infty$ and b_i equal to the censored value, one obtains the left-censored condition. Finally, by setting $y_i = a_i = b_i$ one obtains the fully-observed condition. It should be noted that, if the support of the distribution is a positive number and the observation is left-censored, then one can set $a_i = 0$ instead of $a_i = -\infty$. This generality of the interval-censoring formulation is extremely useful in a reliability setting because it is often the case that one sample can contain measurements that arrive under varying censoring conditions. In what follows, the likelihood function in the interval-censored case is provided, but it should be emphasized that the formulation is completely general in that any censoring condition can be handled transparently.

Suppose that the interval-censored data $[a_j, b_j]$ for $j = 1, \ldots, m$ is obtained. Then, by ignoring the normalizing constant, the likelihood function is obtained as

$$L(\boldsymbol{\Theta}) \propto \prod_{j=1}^{m} [F(b_j) - F(a_j)], \tag{3}$$

where $F(\cdot)$ is a cumulative distribution function. For more details regarding the above likelihood function in Equation (3), one is referred to [24].

Similarly, in the more common case when fully-observed and censored data are both present in the same sample, the likelihood function becomes:

$$L(\boldsymbol{\Theta}) \propto \prod_{i=1}^{n} f(y_i) \cdot \prod_{j=1}^{m} [F(b_j) - F(a_j)]. \tag{4}$$

Note that the likelihood in Equation (4) can often be quite complex and difficult to maximize numerically.

3. The EM Algorithm

In this section, we describe the EM algorithm briefly and show that one can use the EM algorithm to estimate the unknown parameters when the likelihood has an interval-censored component. The EM methodology is often a convenient alternative to the more standard estimation approach in which the likelihood of the sample is constructed and then maximized through the use of a numerical optimization such as the Newton–Raphson method. The EM algorithm was proposed by [13] in order to overcome the frequent difficulties associated with more conventional numerical optimization techniques. Good references for the EM algorithm are [25–29].

Suppose that $\mathbf{w} = (w_1, \ldots, w_n)$ are independent and identically distributed (iid) and have a continuous distribution with pdf $f(w)$. Note that it is assumed that the sample $\mathbf{w} = (w_1, \ldots, w_n)$ contains at least one data point that is of interval-censored form. For example, if w_1 is interval-censored at a_1 and b_1, then this implies that all we know is that $a_1 \leq w_1 \leq b_1$. Under the assumption that at least one measurement is of interval-censored form, then the sample is usually referred to as incomplete data in the EM literature.

In what follows, we reformulate the difficult likelihood problem as a missing-data problem, which then allows us to employ the EM algorithm in order to construct the EM sequences. In this manner, we avoid constructing the often complicated likelihood along with the pitfalls associated with the use of conventional numerical optimization.

Let us denote the fully observed (uncensored) part of $w_1, \ldots, w_n$ by $\mathbf{y} = (y_1, \ldots, y_m)$ and the missing (censored) part by $\mathbf{z} = (z_{m+1}, \ldots, z_n)$ with $a_i \leq z_i \leq b_i$. Also, let us denote the vector of unknown parameters as $\boldsymbol{\Theta} = (\kappa, \theta)$. Then, ignoring the normalizing constant, the complete-data likelihood is shown below:

$$L^c(\boldsymbol{\Theta}|\mathbf{y}, \mathbf{z}) \propto \prod_{i=1}^{n} f(w_i). \tag{5}$$

Now, it will often be the case that the complete-data likelihood above is quite difficult to maximize directly because of its inherent complexity. The key idea underlying the EM algorithm is that it solves a difficult incomplete-data problem by constructing two simpler steps, referred to as the E-step and M-step, respectively. These two steps are repeated over and over until the optimal parameter estimates are obtained. The E-step constructs the conditional expectation of the complete-data log-likelihood function, $\ln L^c(\boldsymbol{\Theta}|\mathbf{y},\mathbf{z})$, where the expectation is taken with respect to the missing part $\mathbf{z}$ of the complete data.

Then, in the M-step, one maximizes the expected log-likelihood function that was constructed during the E-step. It is often the case that the EM algorithm will result in an EM sequence that is straightforward to compute. To summarize, the EM algorithm consists of two distinct steps:

- E-step: compute $Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(s)}) = \int \ln L^c(\boldsymbol{\Theta}|\mathbf{y},\mathbf{z})\, f(\mathbf{z}|\mathbf{y},\boldsymbol{\Theta}^{(s)})\, d\mathbf{z}$,
  where $f(\cdot|\cdot)$ is the conditional pdf given the fully-observed $\mathbf{y}$ with the estimated parameter $\boldsymbol{\Theta}^{(s)}$ at the sth step.
- M-step: find the $\boldsymbol{\Theta}^{(s+1)}$ that maximizes $Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(s)})$ in $\boldsymbol{\Theta}$.

The E-step and M-step are repeated successively until the change in the estimates, namely, $\boldsymbol{\Theta}^{(s+1)} - \boldsymbol{\Theta}^{(s)}$, is relatively small. For example, one can stop the EM algorithm if $|\boldsymbol{\Theta}^{(s+1)} - \boldsymbol{\Theta}^{(s)}| < \epsilon|\boldsymbol{\Theta}^{(s+1)}|$, where ϵ is a pre-determined precision of the estimates. The example provided in the following section should elucidate the steps behind the EM algorithm.

4. The Implementation of the EM Algorithm and Parameter Estimation Method

In this section, the explicit $Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(s)})$ function is provided along with the one-dimensional objective function, both of which allow for the estimation of the Weibull parameters using the EM algorithm. First, the explicit form of Q function is derived in the E-step. Deriving this expression requires the use of some complicated calculus and algebra. Next, using the explicit form obtained in the E-step, we obtain the one-dimensional objective function, which allows for a straightforward estimation of the Weibull parameters.

4.1. E-Step

It is immediate upon using Equations (1) and (5) that the complete-data log-likelihood function is given by

$$\begin{aligned}
\ln L^c(\boldsymbol{\Theta}|\mathbf{y},\mathbf{z}) &= \sum_{i=1}^{n} \ln f(w_i) \\
&= n\ln\kappa - n\kappa\ln\theta + \kappa\sum_{i=1}^{m}\ln y_i - \left(\frac{1}{\theta}\right)^{\kappa}\sum_{i=1}^{m} y_i^{\kappa} \\
&\quad + \kappa\sum_{i=m+1}^{n}\ln z_i - \left(\frac{1}{\theta}\right)^{\kappa}\sum_{i=m+1}^{n} z_i^{\kappa} + C,
\end{aligned} \tag{6}$$

where C is a term that does not include the parameters κ and θ.

Because of the iid structure, the conditional pdf of the missing data $\mathbf{z}$, given the fully-observed data $\mathbf{y}$ with the parameter, does not depend on the observed data. Considering $a_i \le z_i \le b_i$, we have

$$f(\mathbf{z}|\mathbf{y},\boldsymbol{\Theta}_s) = f(\mathbf{z}|\boldsymbol{\Theta}_s) = \prod_{i=m+1}^{n} f(z_i|\boldsymbol{\Theta}_s)$$

where $\boldsymbol{\Theta}_s = (\kappa_s, \theta_s)$ and

$$f(z_i|\boldsymbol{\Theta}_s) = \frac{\dfrac{\kappa_s z_i^{\kappa_s - 1}}{{\theta_s}^{\kappa_s}} \exp\left[-\left(\dfrac{z_i}{\theta_s}\right)^{\kappa_s}\right]}{\exp\left[-\left(\dfrac{a_i}{\theta_s}\right)^{\kappa_s}\right] - \exp\left[-\left(\dfrac{b_i}{\theta_s}\right)^{\kappa_s}\right]}, \tag{7}$$

for $a_i < b_i$ and $a_i \le z_i \le b_i$. Then, using Equations (6) and (7), we can compute

$$Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(s)}) = \int \ln L^c(\boldsymbol{\Theta}|\mathbf{y}, \mathbf{z})\, f(\mathbf{z}|\mathbf{y}, \boldsymbol{\Theta}^{(s)})\, d\mathbf{z}.$$

In order to obtain the explicit integral result, the following lemmas, theorems, and corollary are needed.

Lemma 1. *For $a > 0$, we have*

$$\int_a^\infty \ln z \cdot z^{p-1} \exp\Big(-\frac{z^p}{r}\Big) dz = \frac{r \ln a}{p} \exp\Big(-\frac{a^p}{r}\Big) + \frac{r}{p^2}\Gamma\Big(0, \frac{a^p}{r}\Big).$$

where Γ is the upper incomplete gamma function defined as

$$\Gamma(\kappa, \beta) = \int_\beta^\infty t^{\kappa-1} \exp(-t) dt. \tag{8}$$

Proof. We let $t = z^p/r$. Then, we have

$$z = (rt)^{1/p} \quad \text{and} \quad dz = \frac{1}{p}\, r^{1/p}\, t^{1/p-1} dt.$$

The following is obtained using the relation above:

$$\begin{aligned}
&\int_a^\infty \ln z \cdot z^{p-1} \exp\Big(-\frac{z^p}{r}\Big) dz \\
&\quad = \int_{a^p/r}^\infty \frac{r}{p^2} (\ln r + \ln t) \exp(-t) dt \\
&\quad = \frac{r \ln r}{p^2} \int_{a^p/r}^\infty \exp(-t) dt + \frac{r}{p^2} \int_{a^p/r}^\infty \ln t \cdot \exp(-t) dt \\
&\quad = \frac{r \ln r}{p^2} \exp\Big(-\frac{a^p}{r}\Big) + \frac{r}{p^2} \int_{a^p/r}^\infty \ln t \cdot \exp(-t) dt.
\end{aligned} \tag{9}$$

Using integration by parts, the last term in Equation (9) becomes

$$\begin{aligned}
\int_{a^p/r}^\infty \ln t \cdot \exp(-t) dt &= \Big[-\ln t \cdot \exp(-t)\Big]_{a^p/r}^\infty + \int_{a^p/r}^\infty t^{-1} \exp(-t) dt \\
&= (p \ln a - \ln r) \exp\Big(-\frac{a^p}{r}\Big) + \Gamma\Big(0, \frac{a^p}{r}\Big).
\end{aligned} \tag{10}$$

Substituting Equation (10) into the last term in Equation (9), we obtain

$$\int_a^\infty \ln z \cdot z^{p-1} \exp\Big(-\frac{z^p}{r}\Big) dz = \frac{r \ln a}{p} \exp\Big(-\frac{a^p}{r}\Big) + \frac{r}{p^2}\Gamma\Big(0, \frac{a^p}{r}\Big),$$

which completes the proof. □

Lemma 2. *We have*

$$\lim_{a \downarrow 0} \int_a^\infty \ln z \cdot z^{p-1} \exp\Big(-\frac{z^p}{r}\Big) dz = \frac{r}{p^2}(\ln r - \gamma),$$

where γ is the Euler–Mascheroni constant given by

$$\gamma = \lim_{n \to \infty} \Big(-\ln n + \sum_{k=1}^n \frac{1}{k}\Big) \approx 0.57721566490153.$$

Proof. The exponential integral is defined as

$$\mathrm{Ei}(x) = -\int_{-x}^{\infty} t^{-1} \exp(-t)dt,$$

for $x \neq 0$. Then, it is easily seen that the upper incomplete gamma function in Equation (8) satisfies the relation

$$\Gamma(0, \beta) = -\mathrm{Ei}(-\beta).$$

Now, it is well known that the Taylor series expansion of the exponential integral [30] is given by

$$\mathrm{Ei}(x) = \gamma + \ln|x| + \sum_{k=1}^{\infty} \frac{x^k}{k \cdot k!}.$$

Thus, the Taylor series expansion of $\Gamma(0, x)$ is given by

$$\Gamma(0, x) = -\mathrm{Ei}(-x) = -\gamma - \ln|x| - \sum_{k=1}^{\infty} \frac{(-x)^k}{k \cdot k!}.$$

Using this, we have

$$\Gamma\left(0, \frac{a^p}{r}\right) = -\gamma - p\ln a + \ln r - \sum_{k=1}^{\infty} \frac{(-a^p/r)^k}{k \cdot k!}. \tag{11}$$

Substituting Equation (11) into the result of Lemma 1, we have

$$\begin{aligned}\int_a^{\infty} \ln z \cdot z^{p-1} \exp\left(-\frac{z^p}{r}\right)dz = \frac{r\ln a}{p}\left\{\exp\left(-\frac{a^p}{r}\right) - 1\right\} + \frac{r}{p^2}(\ln r - \gamma) \\ - \frac{r}{p^2}\sum_{k=1}^{\infty} \frac{(-a^p/r)^k}{k \cdot k!}.\end{aligned} \tag{12}$$

Next, using the standard Taylor series expansion below

$$\exp(-x) = 1 - x + \frac{1}{2!}x^2 - \frac{1}{3!}x^3 + \cdots,$$

we have

$$\exp\left(-\frac{a^p}{r}\right) - 1 = -\frac{1}{r}a^p + \frac{1}{2!r^2}a^{2p} - \frac{1}{3!r^3}a^{3p} + \cdots.$$

and

$$\ln a \cdot \left\{\exp\left(-\frac{a^p}{r}\right) - 1\right\} = -\frac{1}{r}a^p\ln a + \frac{1}{2!r^2}a^{2p}\ln a - \frac{1}{3!r^3}a^{3p}\ln a + \cdots. \tag{13}$$

Using L'Hôpital's rule, we have

$$\lim_{a\downarrow 0}\left[a^{kp}\ln a\right] = \lim_{a\downarrow 0}\left[\frac{\ln a}{a^{-kp}}\right] = \lim_{a\downarrow 0}\left[-\frac{a^{kp}}{kp}\right] = 0, \tag{14}$$

for $p > 0$ and $k = 1, 2, \ldots$

Using Equations (13) and (14), it is straightforward to show that

$$\lim_{a\downarrow 0}\left[\ln a \cdot \left\{\exp\left(-\frac{a^p}{r}\right) - 1\right\}\right] = 0. \tag{15}$$

Since $(-a^p/r)^k \to 0$ as $a \downarrow 0$ for $p > 0$ and $k = 1, 2, \ldots$, we have

$$\lim_{a \downarrow 0} \sum_{k=1}^{\infty} \frac{(-a^p/r)^k}{k \cdot k!} = 0. \tag{16}$$

Applying Equations (15) and (16) to (12), we have

$$\lim_{a \downarrow 0} \int_a^{\infty} \ln z \cdot z^{p-1} \exp\Big(-\frac{z^p}{r}\Big) dz = \frac{r}{p^2}(\ln r - \gamma),$$

which completes the proof. □

It is noteworthy that the Euler–Mascheroni constant is easily calculated in the R language [31] using the built-in `digamma` function, that is, $\gamma = -$`digamma(1)`.

Theorem 1. *For* $0 < a_i < b_i < \infty$, *we have*

$$\begin{aligned}\int_{a_i}^{b_i} \ln z_i \cdot f(z_i|\boldsymbol{\Theta}_s) dz_i = & \frac{1}{D_{i,s}} \Big[\ln a_i \cdot \exp\Big[-\Big(\frac{a_i}{\theta_s}\Big)^{\kappa_s}\Big] + \frac{1}{\kappa_s} \Gamma\Big(0, \big(\frac{a_i}{\theta_s}\big)^{\kappa_s}\Big)\Big] \\ & - \frac{1}{D_{i,s}} \Big[\ln b_i \cdot \exp\Big[-\Big(\frac{b_i}{\theta_s}\Big)^{\kappa_s}\Big] + \frac{1}{\kappa_s} \Gamma\Big(0, \big(\frac{b_i}{\theta_s}\big)^{\kappa_s}\Big)\Big],\end{aligned}$$

where $D_{i,s} = \exp\big[-(a_i/\theta_s)^{\kappa_s}\big] - \exp\big[-(b_i/\theta_s)^{\kappa_s}\big]$.

Proof. It is immediate upon using Equation (7) that we have

$$\begin{aligned}\int_{a_i}^{b_i} \ln z_i \cdot f(z_i|\boldsymbol{\Theta}_s) dz_i &= \frac{\int_{a_i}^{b_i} \ln z_i \cdot \frac{\kappa_s z_i^{\kappa_s - 1}}{\theta_s{}^{\kappa_s}} \exp\Big[-\Big(\frac{z_i}{\theta_s}\Big)^{\kappa_s}\Big] dz_i}{\exp\Big[-\Big(\frac{a_i}{\theta_s}\Big)^{\kappa_s}\Big] - \exp\Big[-\Big(\frac{b_i}{\theta_s}\Big)^{\kappa_s}\Big]} \\ &= \frac{\kappa_s}{\theta_s{}^{\kappa_s} D_{i,s}} \int_{a_i}^{b_i} \ln z_i \cdot z_i^{\kappa_s - 1} \exp\Big(-\frac{z_i^{\kappa_s}}{\theta_s{}^{\kappa_s}}\Big) dz_i. \end{aligned} \tag{17}$$

Using Lemma 1, we obtain

$$\int_{a_i}^{\infty} \ln z_i \cdot z_i^{\kappa_s - 1} \exp\Big(-\frac{z_i^{\kappa_s}}{\theta_s{}^{\kappa_s}}\Big) dz_i = \frac{\theta_s^{\kappa_s} \ln a_i}{\kappa_s} \cdot \exp\Big[-\Big(\frac{a_i}{\theta_s}\Big)^{\kappa_s}\Big] + \frac{\theta_s^{\kappa_s}}{\kappa_s^2} \Gamma\Big(0, \Big(\frac{a_i}{\theta_s}\Big)^{\kappa_s}\Big) \tag{18}$$

and

$$\int_{b_i}^{\infty} \ln z_i \cdot z_i^{\kappa_s - 1} \exp\Big(-\frac{z_i^{\kappa_s}}{\theta_s{}^{\kappa_s}}\Big) dz_i = \frac{\theta_s^{\kappa_s} \ln b_i}{\kappa_s} \cdot \exp\Big[-\Big(\frac{b_i}{\theta_s}\Big)^{\kappa_s}\Big] + \frac{\theta_s^{\kappa_s}}{\kappa_s^2} \Gamma\Big(0, \Big(\frac{b_i}{\theta_s}\Big)^{\kappa_s}\Big). \tag{19}$$

Substituting Equations (18) and (19) into (17), we have

$$\begin{aligned}\int_{a_i}^{b_i} \ln z_i \cdot f(z_i|\boldsymbol{\Theta}_s) dz_i = & \frac{1}{D_{i,s}} \Big[\ln a_i \cdot \exp\Big[-\Big(\frac{a_i}{\theta_s}\Big)^{\kappa_s}\Big] + \frac{1}{\kappa_s} \Gamma\Big(0, \big(\frac{a_i}{\theta_s}\big)^{\kappa_s}\Big)\Big] \\ & - \frac{1}{D_{i,s}} \Big[\ln b_i \cdot \exp\Big[-\Big(\frac{b_i}{\theta_s}\Big)^{\kappa_s}\Big] + \frac{1}{\kappa_s} \Gamma\Big(0, \big(\frac{b_i}{\theta_s}\big)^{\kappa_s}\Big)\Big],\end{aligned}$$

which completes the proof. □

Corollary 1. *For* $a_i > 0$ *and* $b_i \to \infty$, *we have*

$$\begin{aligned}\int_{a_i}^{\infty} \ln z_i \cdot f(z_i|\mathbf{\Theta}_s)dz_i &= \lim_{b_i\to\infty}\int_{a_i}^{b_i} \ln z_i \cdot f(z_i|\mathbf{\Theta}_s)dz_i \\ &= \frac{1}{D_s(a_i)}\left[\ln a_i \cdot \exp\Big(-\Big(\frac{a_i}{\theta_s}\Big)^{\kappa_s}\Big) + \frac{1}{\kappa_s}\Gamma\Big(0, \big(\frac{a_i}{\theta_s}\big)^{\kappa_s}\Big)\right],\end{aligned}$$

where $D_s(a_i) = \exp\big[-(a_i/\theta_s)^{\kappa_s}\big]$. *In addition, for* $a_i \downarrow 0$ *and* $b_i > 0$, *we have*

$$\begin{aligned}\int_{0}^{b_i} \ln z_i \cdot f(z_i|\mathbf{\Theta}_s)dz_i &= \lim_{a_i\downarrow 0}\int_{a_i}^{b_i} \ln z_i \cdot f(z_i|\mathbf{\Theta}_s)dz_i \\ &= \frac{1}{D^*_{i,s}(b_i)}\left[\ln\theta_s - \frac{\gamma}{\kappa_s} - \ln b_i \cdot \exp\Big(-\Big(\frac{b_i}{\theta_s}\Big)^{\kappa_s}\Big) - \frac{1}{\kappa_s}\Gamma\Big(0, \big(\frac{b_i}{\theta_s}\big)^{\kappa_s}\Big)\right],\end{aligned}$$

where $D^*_s(b_i) = 1 - \exp\big[-(b_i/\theta_s)^{\kappa_s}\big]$.

Proof. The first result will be derived as follows. It is easily seen from L'Hôpital's rule that

$$\lim_{b_i\to\infty} \ln b_i \cdot \exp\Big[-\Big(\frac{b_i}{\theta_s}\Big)^{\kappa_s}\Big] = \lim_{b_i\to\infty}\frac{\ln b_i}{\exp\big[\big(\frac{b_i}{\theta_s}\big)^{\kappa_s}\big]} = 0.$$

Also, for $\kappa_s > 0$ and $\theta_s > 0$, we have $\lim_{b_i\to\infty}\Gamma\big(0,(b_i/\theta_s)^{\kappa_s}\big) = 0$ from Equation (8). Thus, using the result of Theorem 1, we have

$$\begin{aligned}\int_{a_i}^{\infty} \ln z_i \cdot f(z_i|\mathbf{\Theta}_s)dz_i &= \lim_{b_i\to\infty}\int_{a_i}^{b_i} \ln z_i \cdot f(z_i|\mathbf{\Theta}_s)dz_i \\ &= \frac{1}{D_s(a_i)}\left[\ln a_i \cdot \exp\Big[-\Big(\frac{a_i}{\theta_s}\Big)^{\kappa_s}\Big] + \frac{1}{\kappa_s}\Gamma\Big(0, \big(\frac{a_i}{\theta_s}\big)^{\kappa_s}\Big)\right].\end{aligned} \tag{20}$$

Next, the second result is derived as follows.

$$\begin{aligned}\int_{0}^{b_i} \ln z_i \cdot f(z_i|\mathbf{\Theta}_s)dz_i &= \lim_{a_i\downarrow 0}\int_{a_i}^{b_i} \ln z_i \cdot f(z_i|\mathbf{\Theta}_s)dz_i \\ &= \frac{1}{D^*_s(b_i)}\lim_{a_i\downarrow 0}\int_{a_i}^{b_i}\ln z_i \cdot \frac{\kappa_s z_i^{\kappa_s-1}}{\theta_s^{\kappa_s}}\exp\Big[-\Big(\frac{z_i}{\theta_s}\Big)^{\kappa_s}\Big]dz_i \\ &= \frac{1}{D^*_s(b_i)}\lim_{a_i\downarrow 0}\int_{a_i}^{\infty}\ln z_i \cdot \frac{\kappa_s z_i^{\kappa_s-1}}{\theta_s^{\kappa_s}}\exp\Big[-\Big(\frac{z_i}{\theta_s}\Big)^{\kappa_s}\Big]dz_i \\ &\quad - \frac{1}{D^*_s(b_i)}\int_{b_i}^{\infty}\ln z_i \cdot \frac{\kappa_s z_i^{\kappa_s-1}}{\theta_s^{\kappa_s}}\exp\Big[-\Big(\frac{z_i}{\theta_s}\Big)^{\kappa_s}\Big]dz_i\end{aligned} \tag{21}$$

It is immediate upon using Equation (7) that we have

$$\lim_{a_i\downarrow 0}\int_{a_i}^{\infty}\ln z_i \cdot \frac{\kappa_s z_i^{\kappa_s-1}}{\theta_s^{\kappa_s}}\exp\Big[-\Big(\frac{z_i}{\theta_s}\Big)^{\kappa_s}\Big]dz_i = \frac{\kappa_s}{{\theta_s}^{\kappa_s}}\lim_{a_i\downarrow 0}\int_{a_i}^{\infty}\ln z_i \cdot z_i^{\kappa_s-1}\exp\Big(-\frac{z_i^{\kappa_s}}{{\theta_s}^{\kappa_s}}\Big)dz_i. \tag{22}$$

Using Lemma 2, we have

$$\begin{aligned}\lim_{a_i\downarrow 0}\int_{a_i}^{\infty}\ln z_i \cdot z_i^{\kappa_s-1}\exp\Big(-\frac{z_i^{\kappa_s}}{{\theta_s}^{\kappa_s}}\Big)dz_i &= \frac{\theta_s^{\kappa_s}}{\kappa_s^2}(\ln\theta_s^{\kappa_s} - \gamma) \\ &= \frac{\theta_s^{\kappa_s}}{\kappa_s}\Big(\ln\theta_s - \frac{\gamma}{\kappa_s}\Big).\end{aligned} \tag{23}$$

Substituting Equation (23) into (22), we have

$$\lim_{a_i \downarrow 0} \int_{a_i}^{\infty} \ln z_i \cdot \frac{\kappa_s z_i^{\kappa_s - 1}}{\theta_s^{\kappa_s}} \exp\Big[-\Big(\frac{z_i}{\theta_s}\Big)^{\kappa_s}\Big] dz_i = \ln \theta_s - \frac{\gamma}{\kappa_s}. \tag{24}$$

Using Lemma 1, we have

$$\begin{aligned}\int_{b_i}^{\infty} \ln z_i \cdot \frac{\kappa_s z_i^{\kappa_s - 1}}{\theta_s^{\kappa_s}} \exp\Big[-\Big(\frac{z_i}{\theta_s}\Big)^{\kappa_s}\Big] dz_i &= \frac{\kappa_s}{\theta_s^{\kappa_s}} \int_{b_i}^{\infty} \ln z_i \cdot z_i^{\kappa_s - 1} \exp\Big[-\Big(\frac{z_i}{\theta_s}\Big)^{\kappa_s}\Big] dz_i \\ &= \ln b_i \cdot \exp\Big(-\Big(\frac{b_i}{\theta_s}\Big)^{\kappa_s}\Big) + \frac{1}{\kappa_s}\Gamma\Big(0, \Big(\frac{b_i}{\theta_s}\Big)^{\kappa_s}\Big)\end{aligned} \tag{25}$$

Thus, substituting Equations (24) and (25) into (21), we have the second result stated in this Theorem. This completes the proof. □

Lemma 3. *We have*

$$\int_a^{\infty} z^{p-1} \exp\Big(-\frac{z^q}{r}\Big) dz = \frac{r^{p/q}}{q} \cdot \Gamma\Big(\frac{p}{q}, \frac{a^q}{r}\Big).$$

Proof. Let $t = z^q/r$. Then, we have $z = (rt)^{1/q}$ and $dz = \frac{1}{q}\, r^{1/q}\, t^{1/q-1} dt$. Using this substitution, we derive the relation below:

$$\begin{aligned}\int_a^{\infty} z^{p-1} \exp\Big(-\frac{z^q}{r}\Big) dz &= \int_{a^q/r}^{\infty} (rt)^{(p-1)/q} \exp(-t) \cdot \frac{1}{q} r^{1/q}\, t^{1/q-1} dt \\ &= \frac{r^{p/q}}{q} \int_{a^q/r}^{\infty} t^{p/q-1} \exp(-t) dt \\ &= \frac{r^{p/q}}{q} \cdot \Gamma\Big(\frac{p}{q}, \frac{a^q}{r}\Big),\end{aligned}$$

where Γ is the upper incomplete gamma function again. This completes the proof. □

Theorem 2. *We have*

$$\int_{a_i}^{b_i} z_i^{\kappa} f(z_i|\mathbf{\Theta}_s) dz_i = \frac{{\theta_s}^{\kappa}}{D_{i,s}} \Big[\Gamma\Big(\frac{\kappa+\kappa_s}{\kappa_s}, \Big(\frac{a_i}{\theta_s}\Big)^{\kappa_s}\Big) - \Gamma\Big(\frac{\kappa+\kappa_s}{\kappa_s}, \Big(\frac{b_i}{\theta_s}\Big)^{\kappa_s}\Big)\Big],$$

where $D_{i,s} = \exp\big[-(a_i/\theta_s)^{\kappa_s}\big] - \exp\big[-(b_i/\theta_s)^{\kappa_s}\big]$.

Proof. It is immediate upon using Equation (7) that we have

$$\begin{aligned}\int_{a_i}^{b_i} z_i^{\kappa} f(z_i|\mathbf{\Theta}_s) dz_i &= \frac{\displaystyle\int_{a_i}^{b_i} z_i^{\kappa} \cdot \frac{\kappa_s z_i^{\kappa_s - 1}}{{\theta_s}^{\kappa_s}} \exp\Big[-\Big(\frac{z_i}{\theta_s}\Big)^{\kappa_s}\Big] dz_i}{\exp\Big[-\Big(\frac{a_i}{\theta_s}\Big)^{\kappa_s}\Big] - \exp\Big[-\Big(\frac{b_i}{\theta_s}\Big)^{\kappa_s}\Big]} \\ &= \frac{\kappa_s {\theta_s}^{-\kappa_s}}{D_{i,s}} \int_{a_i}^{b_i} z_i^{\kappa+\kappa_s-1} \exp\Big(-\frac{z_i^{\kappa_s}}{{\theta_s}^{\kappa_s}}\Big) dz_i,\end{aligned} \tag{26}$$

where $D_{i,s} = \exp\big[-(a_i/\theta_s)^{\kappa_s}\big] - \exp\big[-(b_i/\theta_s)^{\kappa_s}\big]$ again.

Using Lemma 3, we have

$$\int_{a_i}^{\infty} z_i^{\kappa+\kappa_s-1} \exp\Big(-\frac{z_i^{\kappa_s}}{{\theta_s}^{\kappa_s}}\Big) dz_i = \frac{{\theta_s}^{\kappa+\kappa_s}}{\kappa_s} \Gamma\Big(\frac{\kappa+\kappa_s}{\kappa_s}, \frac{a_i^{\kappa_s}}{{\theta_s}^{\kappa_s}}\Big) \tag{27}$$

and

$$\int_{b_i}^{\infty} z_i^{\kappa+\kappa_s-1} \exp\Big(-\frac{z_i^{\kappa_s}}{{\theta_s}^{\kappa_s}}\Big) dz_i = \frac{{\theta_s}^{\kappa+\kappa_s}}{\kappa_s} \Gamma\Big(\frac{\kappa+\kappa_s}{\kappa_s}, \frac{b_i^{\kappa_s}}{{\theta_s}^{\kappa_s}}\Big). \tag{28}$$

It is immediate upon substituting Equations (27) and (28) into (26) that we have

$$\begin{aligned}\int_{a_i}^{b_i} z_i^{\kappa} f(z_i|\boldsymbol{\Theta}_s)dz_i &= \int_{a_i}^{\infty} z_i^{\kappa} f(z_i|\boldsymbol{\Theta}_s)dz_i - \int_{b_i}^{\infty} z_i^{\kappa} f(z_i|\boldsymbol{\Theta}_s)dz_i \\ &= \frac{{\theta_s}^{\kappa}}{D_{i,s}}\left[\Gamma\left(\frac{\kappa+\kappa_s}{\kappa_s}, \left(\frac{a_i}{\theta_s}\right)^{\kappa_s}\right) - \Gamma\left(\frac{\kappa+\kappa_s}{\kappa_s}, \left(\frac{b_i}{\theta_s}\right)^{\kappa_s}\right)\right],\end{aligned}$$

which completes the proof. □

For notational convenience going forward, we let

$$U_{i,s} = \begin{cases} \displaystyle\int_{a_i}^{b_i} \ln z_i \cdot f(z_i|\boldsymbol{\Theta}_s)dz_i & \text{if } b_i > a_i \\ \ln a_i & \text{if } b_i = a_i \end{cases}$$

and

$$V_{i,s}(\kappa) = \begin{cases} \displaystyle\int_{a_i}^{b_i} z_i^{\kappa} f(z_i|\boldsymbol{\Theta}_s)dz_i & \text{if } b_i > a_i \\ a_i^{\kappa} & \text{if } b_i = a_i \end{cases}.$$

Note that the value of $V_{i,s}(\kappa)$ depends only on the parameter κ. Thus, we need to solve for κ in the M-step, the details of which will be described later. However, as shown in Theorem 1, $U_{i,s}$ does not include either κ or θ Thus, using Theorem 1 and Corollary 1, we can easily calculate the value of $U_{i,s}$ at each E-step. The value of $V_{i,s}(\kappa)$ can be calculated by using Theorem 2. This allows for the evaluation of the expected log likelihood at the sth step of the EM sequence. Thus we obtain the expected log-likelihood at the sth step in the EM sequence:

$$\begin{aligned}Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}_s) &= \int \ln L^c(\boldsymbol{\Theta}|\mathbf{y},\mathbf{z}) f(\mathbf{z}|\boldsymbol{\Theta}_s)d\mathbf{z} \\ &= n\ln\kappa - n\kappa\ln\theta + \kappa\sum_{i=1}^{m}\ln y_i + \kappa\sum_{i=m+1}^{n} U_{i,s} \\ &\quad - \left(\frac{1}{\theta}\right)^{\kappa}\sum_{i=1}^{m} y_i^{\kappa} - \left(\frac{1}{\theta}\right)^{\kappa}\sum_{i=m+1}^{n} V_{i,s}(\kappa) + C \\ &= n\ln\kappa - n\kappa\ln\theta + \kappa\sum_{i=1}^{n} U_{i,s} - \left(\frac{1}{\theta}\right)^{\kappa}\sum_{i=1}^{n} V_{i,s}(\kappa) + C.\end{aligned} \tag{29}$$

4.2. M-Step

Differentiating the expected log-likelihood function in Equation (29) with respect to θ and setting it equal to zero, we obtain

$$\frac{\partial Q(\boldsymbol{\Theta}|\boldsymbol{\Theta}_{(s)})}{\partial\theta} = -\frac{n\kappa}{\theta} + \kappa\left(\frac{1}{\theta}\right)^{\kappa+1}\sum_{i=1}^{n} V_{i,s}(\kappa) = 0.$$

Solving the above for θ, we have

$$\theta = \left[\frac{1}{n}\sum_{i=1}^{n} V_{i,s}(\kappa)\right]^{1/\kappa}. \tag{30}$$

Clearly, given Equation (30), θ is only a function of κ. Substituting Equation (30) into (29), we obtain the expected log-likelihood as a function of κ only at each step. This expression reduces to

$$Q(\mathbf{\Theta}|\mathbf{\Theta}_s) = n \ln \kappa + \kappa \sum_{i=1}^{n} U_{i,s} - n \ln\left(\sum_{i=1}^{n} V_{i,s}(\kappa) \right) + n \ln n - n + C.$$

Notice that the problem of maximizing the expected log-likelihood has been transformed into a one-dimensional optimization problem with respect to κ. Thus, it suffices to maximize the following:

$$n \ln \kappa + \kappa \sum_{i=1}^{n} U_{i,s} - n \ln\left(\sum_{i=1}^{n} V_{i,s}(\kappa) \right). \tag{31}$$

5. A Real-Data Example

In this section, we investigate the performance of the proposed EM-based method using a real-data set from the literature. The data was taken from a study of patients with breast cancer [32,33]. This data set can also be found in Table 3.10 of [34]. The observed value of interest was the time it took until there was a cosmetic deterioration of the breast. For convenience, the interval-censored observations are also provided in Table 1.

Table 1. Interval-censored observations.

$[8,12]$	$[0,22]$	$[24,31]$	$[17,27]$	$[17,23]$	$[24,30]$	$[16,24]$	$[13,\infty]$	$[11,13]$	$[16,20]$
$[18,25]$	$[17,26]$	$[32,\infty]$	$[23,\infty]$	$[44,48]$	$[10,35]$	$[0,5]$	$[5,8]$	$[12,20]$	$[11,\infty]$
$[33,40]$	$[31,\infty]$	$[13,39]$	$[19,32]$	$[34,\infty]$	$[13,\infty]$	$[16,24]$	$[35,\infty]$	$[15,22]$	$[11,17]$
$[22,32]$	$[48,\infty]$	$[30,34]$	$[13,\infty]$	$[10,17]$	$[8,21]$	$[4,9]$	$[11,\infty]$	$[14,19]$	$[4,8]$
$[34,\infty]$	$[30,36]$	$[18,24]$	$[16,60]$	$[35,39]$	$[21,\infty]$	$[11,20]$			

Assuming that time to cosmetic deterioration has a Weibull distribution, one can estimate the Weibull parameters of the breast cancer data set. This in turn allows for the comparison of the proposed EM-based estimation method with the estimation method based on direct maximization of the likelihood using the DFP and BHHH methods. Some specific details behind the two methods are provided below.

In order to optimize the log-likelihood function in Equation (4) using the DFP and BHHH methods, we use the R packages `pracma` [35] and `maxLik` [36,37] respectively. As a precursor to maximizing the likelihood, we sketch the likelihood function using the perspective and contour plots in Figure 2. Both of the plots suggest that the reasonable parameter estimates $(\hat{\kappa}, \hat{\theta})$ are in the general vicinity of $\hat{\kappa} = 2$ and $\hat{\theta} = 30$. In fact, it turns out that the exact MLE, which is a component of the object returned by the call to the contour plot function in R is $\hat{\kappa} = 2.026$ and $\hat{\theta} = 28.34$. In the case of the EM-based method, it will be shown that the optimization during the M-step is quite straightforward because the derived EM sequence is a one-dimensional function of κ.

Finally, when comparing the algorithms, the sole criterion investigated in this note is the sensitivity of the respective algorithms to the respective starting values used. In order to investigate the sensitivity efficiently, we generated a symmetric and circular set of starting value (κ_0, θ_0) pairs whose center is equal to $(2, 30)$. An illustration of the circular set of starting values is provided in Figure 3a. Notice that the circular set of two-dimensional starting values is generated by letting

$$\kappa_0 = 2 + 1.5 \cdot \cos(k \times 10^\circ) \quad \text{and} \quad \theta_0 = 30 + 25 \cdot \sin(k \times 10^\circ)$$

where $k = 1, 2, \cdots, 36$. Essentially, the equations above generate (κ_0, θ_0) pairs that span the circumference of a circle with spaces of 10 degrees with the center at $(2, 30)$. Obviously, any center point that is relatively close to the true MLE could have been chosen as the center, but we felt that using whole integer values of 2 and 30 was convenient for illustration.

Next, for each of the two-dimensional starting values along the circumference of the circle, we obtained the parameter estimates using (i) the DFP method, (ii) the BHHH method, and (iii) the EM method. Additionally, the values of the likelihood function associated with the respective estimates were calculated using Equation (4). These results are summarized in Table 2.

Table 2. The parameter estimates with their corresponding likelihood values using the DFP and BHHH methods and the proposed EM algorithm.

Angle	**Starting Values**		**DFP**			**BHHH**			**EM**		
Degree (°)	κ_0	θ_0	$\hat{\kappa}$	$\hat{\theta}$	$L \times 10^{32}$	$\hat{\kappa}$	$\hat{\theta}$	$L \times 10^{32}$	$\hat{\kappa}$	$\hat{\theta}$	$L \times 10^{32}$
10	3.477	34.34	2.026	28.34	1.515	2.026	28.33	1.515	2.026	28.34	1.515
20	3.410	38.55	2.026	28.34	1.515	2.026	28.33	1.515	2.026	28.34	1.515
30	3.299	42.50	2.026	28.34	1.515	2.027	28.33	1.515	2.026	28.34	1.515
40	3.149	46.07	2.026	28.34	1.515	2.026	28.33	1.515	2.026	28.34	1.515
50	2.964	49.15	1.298	46.78	1.140×10^{-4}	2.026	28.34	1.515	2.026	28.34	1.515
60	2.750	51.65	2.114	39.26	2.858×10^{-3}	2.027	28.33	1.515	2.026	28.34	1.515
70	2.513	53.49	2.026	28.34	1.515	2.027	28.34	1.515	2.026	28.34	1.515
80	2.260	54.62	1.142	44.07	1.178×10^{-4}	2.026	28.33	1.515	2.026	28.34	1.515
90	2.000	55.00	2.069	41.71	4.192×10^{-4}	2.026	28.34	1.515	2.026	28.34	1.515
100	1.740	54.62	2.312	34.39	6.794×10^{-2}	2.026	28.34	1.515	2.026	28.34	1.515
110	1.487	53.49	1.298	52.47	8.944×10^{-6}	2.026	28.33	1.515	2.026	28.34	1.515
120	1.250	51.65	1.334	50.64	2.082×10^{-5}	2.026	28.34	1.515	2.026	28.34	1.515
130	1.036	49.15	1.382	48.14	7.009×10^{-5}	2.026	28.34	1.515	2.026	28.34	1.515
140	0.851	46.07	2.026	28.34	1.515	2.026	28.34	1.515	2.026	28.34	1.515
150	0.701	42.50	1.572	41.50	2.651×10^{-3}	2.026	28.34	1.515	2.026	28.34	1.515
160	0.591	38.55	2.026	28.34	1.515	2.026	28.33	1.515	2.026	28.34	1.515
170	0.523	34.34	2.026	28.34	1.515	2.027	28.34	1.515	2.026	28.34	1.515
180	0.500	30.00	2.026	28.34	1.515	2.026	28.33	1.515	2.026	28.34	1.515
190	0.523	25.66	2.026	28.34	1.515	2.027	28.34	1.515	2.026	28.34	1.515
200	0.591	21.45	2.026	28.34	1.515	2.026	28.33	1.515	2.026	28.34	1.515
210	0.701	17.50	2.026	28.34	1.515	2.027	28.34	1.515	2.026	28.34	1.515
220	0.851	13.93	2.026	28.34	1.515	2.026	28.33	1.515	2.026	28.34	1.515
230	1.036	10.85	2.026	28.34	1.515	2.026	28.34	1.515	2.026	28.34	1.515
240	1.250	8.35	2.026	28.34	1.515	2.027	28.34	1.515	2.026	28.34	1.515
250	1.487	6.51	0.885	9.43	6.281×10^{-17}	2.026	28.33	1.515	2.026	28.34	1.515
260	1.740	5.38	NA	NA	NA	NA	NA	NA	2.026	28.34	1.515
270	2.000	5.00	NA	NA	NA	NA	NA	NA	2.026	28.34	1.515
280	2.260	5.38	NA	NA	NA	NA	NA	NA	2.026	28.34	1.515
290	2.513	6.51	NA	NA	NA	NA	NA	NA	2.026	28.34	1.515
300	2.750	8.35	NA	NA	NA	NA	NA	NA	2.026	28.34	1.515
310	2.964	10.85	NA	NA	NA	NA	NA	NA	2.026	28.34	1.515
320	3.149	13.93	NA	NA	NA	NA	NA	NA	2.026	28.34	1.515
330	3.299	17.50	2.026	28.34	1.515	2.027	28.34	1.515	2.026	28.34	1.515
340	3.410	21.45	NA	NA	NA	2.026	28.34	1.515	2.026	28.34	1.515
350	3.477	25.66	2.026	28.34	1.515	2.026	28.34	1.515	2.026	28.34	1.515
360	3.500	30.00	2.026	28.34	1.515	2.026	28.33	1.515	2.026	28.34	1.515

Refer back to Figure 3. Notice that, by connecting the starting value to its resulting estimate, we have constructed what is referred to as the diagram of paths, as shown in Figure 3b–d. Unfortunately, in the cases of the DFP and BHHH methods, the diagram paths indicate a number of starting values where convergence was not obtained. For the DFP method, most of these occurred when the angle of the starting value was between 50° and 150° and between 250° and 340°. For the BHHH method, they occurred when the angle of the starting value was between 260° and 320°. Conversely, in the case of the EM algorithm method, all of the starting values converged to the unique estimate which is essentially the true MLE. It is noteworthy that if the likelihood function is unimodal, then all EM sequences converge to the unique MLE [38]. Given the EM convergence results and the two-dimensional nature of the objective function, the frequency of nonconvergence when using the DFP and BHHH methods is concerning.

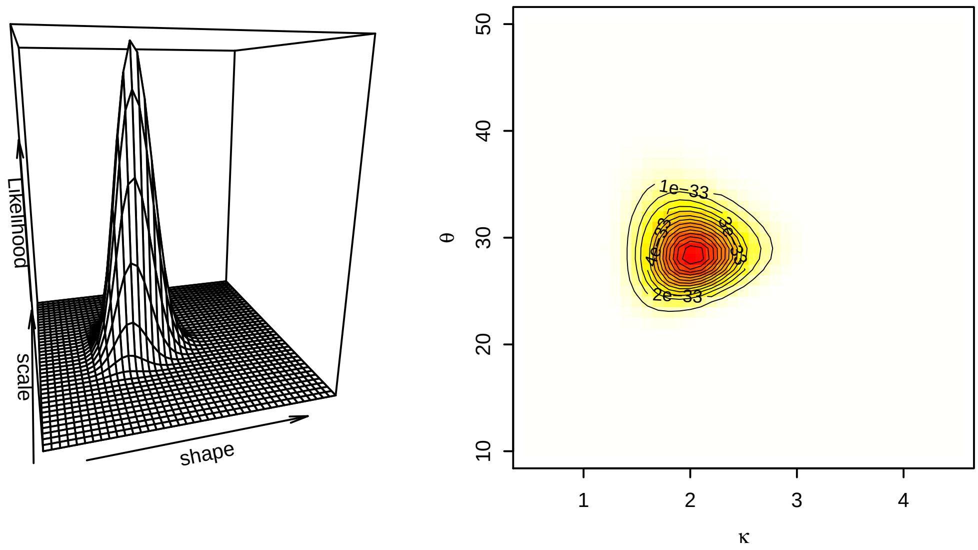

Figure 2. Perspective and contour plots of the likelihood function.

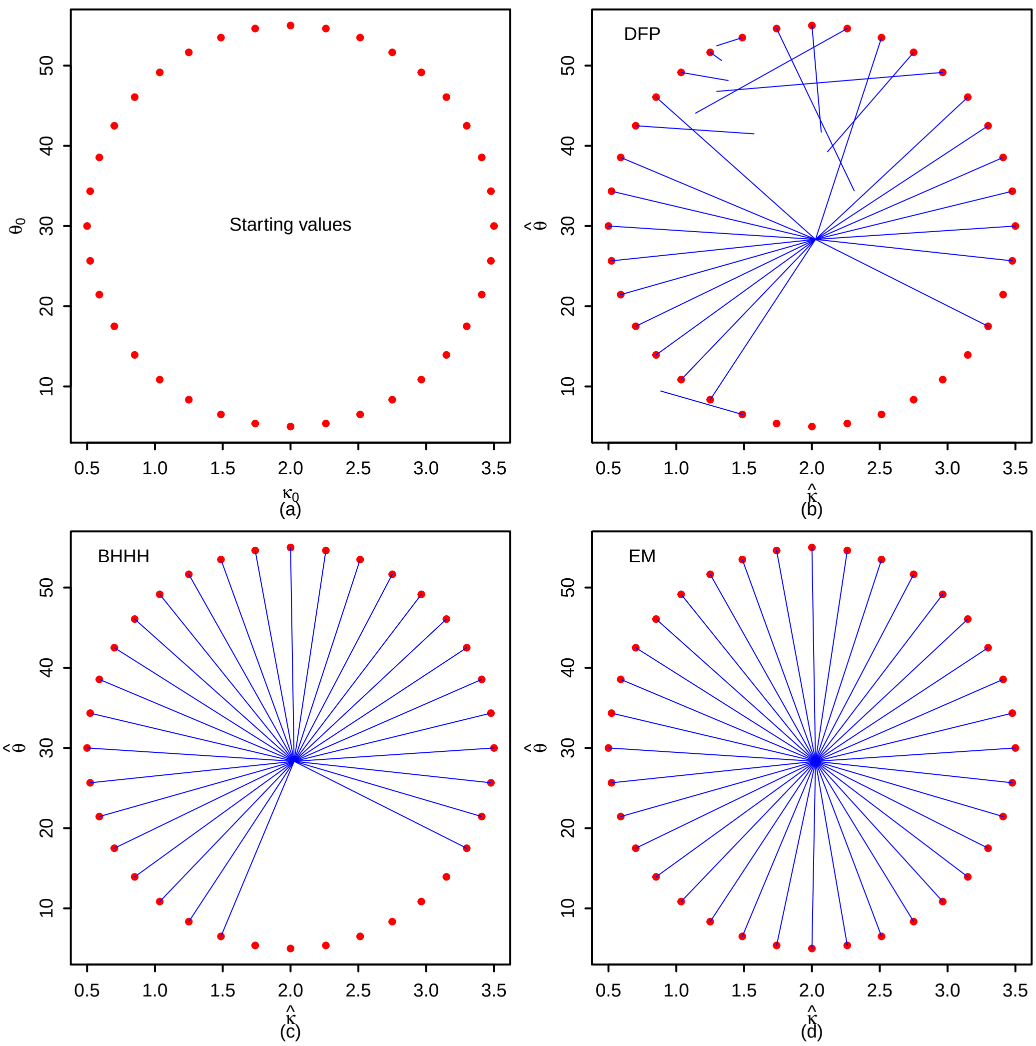

Figure 3. Diagram of paths connecting the starting values to their resulting estimates. (**a**) Starting values given by $(\kappa_0, \theta_0) = (2 + 1.5 \cdot \cos(k \times 10^\circ), 30 + 25 \cdot \sin(k \times 10^\circ))$ for $k = 1, 2, \ldots, 36$. (**b**) Diagram path using the DFP method. (**c**) Diagram path using the BHHH method. (**d**) Diagram path using the EM method.

The R code for the Weibull parameter estimation based on the EM algorithm and the other Newton-type method using the `nlm` R function are available at https://github.com/AppliedStat/R-code/tree/master/2023b, accessed on 15 July 2023. In this URL, additional illustrative examples are also provided.

6. Concluding Remarks

An EM-based parameter estimation method was developed using the EM algorithm in the case of the Weibull distribution with interval-censored data. The sensitivity of the results to various starting values was also investigated. The findings indicate that the suggested technique is not affected by the choice of starting values, unlike the conventional DFP and BHHH methods, which occasionally struggle to reach the maximum-likelihood estimation. By employing the EM-based approach instead of conventional numerical optimization, analysts have a much larger choice of starting values for estimation problems of a similar nature. The R code for the EM algorithm is provided with practical examples, allowing field engineers and practitioners to utilize them according to their needs.

Funding: This research was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (Nos. 2022R1A2C1091319, RS-2023-00242528).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are provided in this paper.

Acknowledgments: This research is dedicated to the memory and honor of Byung Ho Lee of Nuclear Engineering at Seoul National University and KAIST. The author's interests in mathematics and engineering were formed under his strong influence. Lee passed away peacefully in July 2001. The author also wishes to thank anonymous referees and Mark Leeds for their valuable comments that led to the improvement of the original manuscript.

Conflicts of Interest: The author declares no conflict of interest.

Nomenclature

a_i	Lower end of an interval-censored observation
b_i	Upper end of an interval-censored observation
$f(\cdot)$	Probability density function of Weibull distribution
$F(\cdot)$	Cumulative distribution function of Weibull distribution
$\mathrm{Ei}(\cdot)$	Exponential integral function
$L(\cdot)$	Likelihood function
$L^c(\cdot)$	Complete-data likelihood function
$Q(\cdot\|\cdot)$	Q function in the E-step
γ	Euler–Mascheroni constant
$\Gamma(\cdot,\cdot)$	Upper incomplete gamma function
κ	Shape parameter of Weibull distribution
θ	Scale parameter of Weibull distribution

References

1. Bonnans, J.F.; Gilbert, J.C.; Lemarechal, C.; Sagastizabal, C.A. *Numerical Optimization: Theoretical and Practical Aspects*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2006.
2. Berndt, E.; Hall, B.; Hall, R.; Hausman, J. Estimation and Inference in Nonlinear Structural Models. *Ann. Econ. Soc. Meas.* **1974**, *3*, 653–665.
3. Anderson-Bergman, C. icenReg: Regression Models for Interval Censored Data in R. *J. Stat. Softw.* **2017**, *81*, 1–23. [CrossRef]
4. Gupta, A.K. Estimation of the mean and standard deviation of a normal population from a censored sample. *Biometrika* **1952**, *39*, 260–273. [CrossRef]
5. Balakrishnan, N.; Aggarwala, R. *Progressive Censoring: Theory, Methods, and Applications*; Birkhäuser: Boston, MA, USA, 2000.
6. Leemis, L.M. *Reliability: Probabilistic Models and Statistical Methods*, 2nd ed.; Lawrence M. Leemis: Williamsburg, VA, USA, 2009.

7. Govindarajulu, Z. Best linear estimates under symmetric censoring of the parameters of a double exponential population. *J. Am. Stat. Assoc.* **1966**, *61*, 248–258.
8. Balakrishnan, N. Approximate MLE of the scale parameter of the Rayleigh distribution with censoring. *IEEE Trans. Reliab.* **1989**, *38*, 355–357. [CrossRef]
9. Hassanein, K.M.; Saleh, A.K.; Brown, E. *Best Linear Unbiased Estimate and Confidence Interval for Rayleigh's Scale Parameter When the Threshold Parameter Is Known for Data with Censored Observations from the Right*; Report; University of Kansas Medical School, Kansas City, KS, USA, 1995.
10. Elsayed, E.A. *Reliability Engineering*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2012.
11. Sultan, A.M. New Approximation for parameters of normal distribution using Type-II censored sampling. *Microelectron. Reliab.* **1997**, *37*, 1169–1171. [CrossRef]
12. Balakrishnan, N. BLUEs of location and scale parameters of Laplace distribution based on Type-II censored samples and associated inference. *Microelectron. Reliab.* **1996**, *36*, 371–374. [CrossRef]
13. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B* **1977**, *39*, 1–22.
14. Panahi, H.; Asadi, S. Estimation of the Weibull Distribution Based on Type-II Censored Samples. *Appl. Math. Sci.* **2011**, *5*, 2549–2558.
15. Guure, C.B.; Ibrahim, N.A.; Adam, M.B. Bayesian Inference of the Weibull Model Based on Interval-Censored Survival Data. *Comput. Math. Methods Med.* **2013**, *2013*, 849520. [CrossRef]
16. Pradhan, B.; Kundu, D. Analysis of Interval-Censored Data with Weibull Lifetime Distribution. *Sankhya B* **2014**, *76*, 120–139. [CrossRef]
17. Ferreira, L.A.; Silva, J.L. Parameter estimation for Weibull distribution with right censored data using EM algorithm. *Eksploat. I Niezawodn.-Mainteance Reliab.* **2017**, *19*, 310–315. [CrossRef]
18. Park, C. A Quantile Variant of the Expectation-Maximization Algorithm and its Application to Parameter Estimation with Interval Data. *J. Algorithms Comput. Technol.* **2018**, *12*, 253–272. [CrossRef]
19. Saeed, N.M.; Elfaki, F.A.M. Parametric Weibull Model Based on Imputations Techniques for Partly Interval Censored Data. *Austrian J. Stat.* **2020**, *49*, 30–37. [CrossRef]
20. Kurniawan, A.; Avicena, N.; Ana, E. Estimation of the shape parameter of Weibull distribution based on type-II censored data using EM algorithm. *AIP Conf. Proc.* **2020**, *2264*, 030011. [CrossRef]
21. Ameen, A.A.; Akkash, U.A. Maximum Likelihood Estimation and Bayesian Estimation of three-parameter Weibull Distribution Based on Interval-Censored Data. *J. Physics: Conf. Ser.* **2021**, *1818*, 012199. [CrossRef]
22. Almetwally, E.M.; Jawa, T.M.; Sayed-Ahmed, N.; Park, C.; Zakarya, M.; Dey, S. Analysis of unit-Weibull based on progressive type-II censored with optimal scheme. *Alex. Eng. J.* **2023**, *63*, 321–338. [CrossRef]
23. Weibull, W. A Statistical Distribution Function of Wide Applicability. *ASME J. Appl. Mech.* **1951**, *18*, 293–297. [CrossRef]
24. Lee, S.B.; Park, C. Development of Robust Design Optimization using Incomplete Data. *Comput. Ind. Eng.* **2006**, *50*, 345–356. [CrossRef]
25. Tanner, M.A. *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*; Springer: New York, NY, USA, 1996.
26. McLachlan, G.J.; Krishnan, T. *The EM Algorithm and Extensions*; John Wiley & Sons: New York, NY, USA, 1997.
27. Schafer, J.L. *Analysis of Incomplete Multivariate Data*; Chapman & Hall: Boca Raton, FL, USA, 1997.
28. Little, R.J.A.; Rubin, D.B. *Statistical Analysis with Missing Data*, 2nd ed.; John Wiley & Sons: New York, NY, USA, 2002.
29. Do, C.B.; Batzoglou, S. What is the expectation maximization algorithm? *Nat. Biotechnol.* **2008**, *26*, 897–899. [CrossRef]
30. Harris, F.E. Tables of the Exponential Integral $\mathrm{Ei}(x)$. *Math. Tables Other Aids Comput.* **1957**, *11*, 9–16. [CrossRef]
31. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2023. Available online: http://www.r-project.org (accessed on 3 January 2023).
32. Finkelstein, D.M. A proportional hazards model for interval-censored failure time data. *Biometrics* **1986**, *42*, 845–854. [CrossRef]
33. Lindsey, J.C.; Ryan, L.M. Methods for interval-censored data. *Stat. Med.* **1998**, *17*, 219–238. [CrossRef]
34. Lawless, J.F. *Statistical Models and Methods for Lifetime Data*, 2nd ed.; John Wiley & Sons: New York, NY, USA, 2003.
35. Borchers, H.W. Pracma: Practical Numerical Math Functions. R Package Version 2.4.2. (Published on 22 September 2022). 2022. Available online: https://CRAN.R-project.org/package=pracma (accessed on 15 July 2023).
36. Henningsen, A.; Toomet, O. maxLik: A package for maximum likelihood estimation in R. *Comput. Stat.* **2011**, *26*, 443–458. [CrossRef]
37. Toomet, O.; Henningsen, A.; Graves, S.; Croissant, Y.; Hugh-Jones, D.; Scrucca, L. maxLik: Maximum Likelihood Estimation and Related Tools. R Package Version 1.5.2 (Published on 26 July 2021). 2021. Available online: https://CRAN.R-project.org/package=maxLik (accessed on 15 July 2023).
38. Wu, C.F.J. On the convergence properties of the EM algorithm. *Ann. Stat.* **1983**, *11*, 95–103. [CrossRef]