

# Estimation of the Shape Parameter of a Wear-Out Failure Period for a Three-Parameter Weibull Distribution in a Small Sample

Toru Ogura<sup>1</sup>, Takatoshi Sugiyama<sup>2</sup> & Nariaki Sugiura<sup>3</sup>

<sup>1</sup> Clinical Research Support Center, Mie University Hospital, Mie, Japan

<sup>2</sup> Department of applied Mathematics, Science Graduate School of Science, Tokyo University of Science, Tokyo, Japan

<sup>3</sup> Department of Mathematics, University of Tsukuba, Ibaraki, Japan

Correspondence: Toru Ogura, Clinical Research Support Center, Mie University Hospital, Mie, Japan.  
E-mail: t-ogura@clin.medic.mie-u.ac.jp

Received: August 4, 2020 Accepted: September 7, 2020 Online Published: September 24, 2020

doi:10.5539/ijsp.v9n6p39 URL: <https://doi.org/10.5539/ijsp.v9n6p39>

## Abstract

We propose a method to estimate a shape parameter for a three-parameter Weibull distribution. The proposed method first derives an unbiased estimator for the shape parameter independent of the location and scale parameters and then estimates the shape parameter using a minimum-variance linear unbiased estimator. Since the proposed method is expressed using a hyperparameter, its optimal hyperparameter is searched using Monte Carlo simulations. The recommended hyperparameter used for estimating the shape parameter depends on the sample size, and this causes no problems since the sample size is known when data is obtained. The proposed method is evaluated using a bias and a root mean squared error, and the results are very promising when the population shape parameter is 2 or more in the Weibull distribution representing the wear-out failure period. A numerical dataset is analyzed to demonstrate the practical use of the proposed method.

**Keywords:** estimation, shape parameter, small sample, three-parameter Weibull distribution

## 1. Introduction

Weibull distribution is often used in lifetime and reliability studies. The probability density function and cumulative distribution function of the three-parameter Weibull distribution are expressed as follows:

$$g(x; m, \eta, \gamma) = \frac{m}{\eta} \left( \frac{x - \gamma}{\eta} \right)^{m-1} \exp \left[ - \left( \frac{x - \gamma}{\eta} \right)^m \right], \quad (1)$$

$$G(x; m, \eta, \gamma) = 1 - \exp \left[ - \left( \frac{x - \gamma}{\eta} \right)^m \right], \quad (2)$$

where  $\eta > 0$ ,  $m > 0$ , and  $\gamma < x$  are the scale, shape, and location parameters, respectively. Although maximum likelihood estimation (MLE) is often used for parameter estimation, it may not be possible to be used for a three-parameter Weibull distribution. Therefore, the parameter estimation method that extends MLE have proposed for a three-parameter Weibull distribution. Hall & Wang (2005) proposed the Bayesian likelihood (BL) method, which multiplies MLE by an empirical prior. Cousineau (2009a) proposed the weighted maximum likelihood estimation (w-MLE) method that added the three weights to MLE. Nagatsuka et al. (2013) proposed the location and scale parameters free maximum likelihood estimators (LSPF-MLE) method, which estimates the shape parameter using the independent statistics of the location and scale parameters. The parameter estimation method of the three-parameter Weibull distribution has also been studied by other researchers (Smith, 1985; Lawless, 2002; Murthy et al., 2004; Nagatsuka & Kamakura, 2005; Cousineau, 2009b; Ng et al., 2012; Sugiyama et al., 2019).

We propose a method to estimate the shape parameter, independent of the location and scale parameters. Therefore, the shape parameter can be estimated first for the three-parameter Weibull distribution. Estimating the shape parameter with a high accuracy is expected to improve the estimation accuracy of the location and scale parameters.

In Section 2, we derive an unbiased estimator of the shape parameter that does not depend on the location and scale parameters. In Section 3, we search for the recommended hyperparameters for each sample size, and compare the effectiveness of the proposed method with the existing methods. Section 4 describes an attempt at estimating the shape parameter using a numerical example. Our conclusions are presented in Section 5.

## 2. Unbiased Estimator of the Shape Parameter Independent of the Location and Scale Parameters

Let  $X_1, \dots, X_n$  be random samples from the three-parameter Weibull distribution, and let  $X_{(i)}$  denote the  $i$ -th order statistic,  $\gamma < X_{(1)} < \dots < X_{(n)}$  ( $X_{(i)} \neq X_{(j)}$  for some  $i \neq j$ ),  $i = 1, \dots, n$  ( $n \geq 2$ ). Using the inverse transform method, the following formula is established:

$$1 - \exp\left[-\left(\frac{X_{(i)} - \gamma}{\eta}\right)^m\right] = U_{(i)}, \tag{3}$$

where  $U_{(i)}$  is the  $i$ -th order statistics of the uniform distribution,  $0 \leq U_{(1)} < \dots < U_{(n)} \leq 1$  ( $U_{(i)} \neq U_{(j)}$ , for some  $i \neq j$ ). The following results are obtained by double logarithmic conversion of both sides:

$$m \log\left(\frac{X_{(i)} - \gamma}{\eta}\right) = \log[-\log(1 - U_{(i)})], \quad i = 1, \dots, n. \tag{4}$$

Subtracting the formula in which  $i = 1$  is substituted into (4) from the formula in which  $i = 2, \dots, n$  is substituted into (4), it leads to:

$$m \log\left(\frac{X_{(i)} - \gamma}{X_{(1)} - \gamma}\right) = \log\left[\frac{\log(1 - U_{(i)})}{\log(1 - U_{(1)})}\right], \quad i = 2, \dots, n. \tag{5}$$

The right-hand side of (5) is independent of  $\eta$ . The mean of both sides in (5) is given by:

$$E\left[\log\left(\frac{X_{(i)} - \gamma}{X_{(1)} - \gamma}\right)\right] = \frac{1}{m} E\left\{\log\left[\frac{\log(1 - U_{(i)})}{\log(1 - U_{(1)})}\right]\right\} = \frac{k(i, n) - k(1, n)}{m}, \tag{6}$$

where

$$k(i, n) = \int_0^1 \log[-\log(1 - u)] \frac{n!}{(i-1)!(n-i)!} u^{i-1} (1-u)^{n-i} du. \tag{7}$$

The expected value of (6) is the function of only  $i$  and  $n$ . The unbiased estimator of  $1/m$  is expressed as follows:

$$T_{(i)} = \frac{1}{m} = \log\left(\frac{X_{(i)} - \gamma}{X_{(1)} - \gamma}\right) \Big/ [k(i, n) - k(1, n)]. \tag{8}$$

Furthermore, let

$$\frac{1}{X_{(1)} - \gamma} = c \left(\frac{1}{U} - 1\right) \tag{9}$$

be the empirical prior of  $\gamma$ , where  $U$  is the standard uniform distribution, and  $c > 0$  is a hyperparameter. Both sides of (9) have the same range  $(0, \infty)$ . The hyperparameter  $c$  is chosen to obtain an optimal estimator. Substituting (9) into (8), we obtain:

$$\begin{aligned} T_{(i)} &= \log\left(\frac{X_{(i)} - X_{(1)} + X_{(1)} - \gamma}{X_{(1)} - \gamma}\right) \Big/ [k(i, n) - k(1, n)] = \log\left(\frac{X_{(i)} - X_{(1)}}{X_{(1)} - \gamma} + 1\right) \Big/ [k(i, n) - k(1, n)] \\ &= \log\left[\frac{c(1 - U)(X_{(i)} - X_{(1)})}{U} + 1\right] \Big/ [k(i, n) - k(1, n)] = \frac{\log[c(1 - U)(X_{(i)} - X_{(1)}) + U] - \log(U)}{k(i, n) - k(1, n)}. \end{aligned} \tag{10}$$

Using the minimum-variance linear unbiased estimator (Ahsanullah et al., 2013),  $m$  is estimated by:

$$\hat{m} = \frac{1}{T} = \mathbf{1}' \mathbf{S}^{-1} \mathbf{1} \Big/ \mathbf{1}' \mathbf{S}^{-1} \begin{pmatrix} \int_0^1 \{\log[c(1 - U)(X_{(2)} - X_{(1)}) + U] - \log(U)\} dU / [k(i, n) - k(1, n)] \\ \vdots \\ \int_0^1 \{\log[c(1 - U)(X_{(n)} - X_{(1)}) + U] - \log(U)\} dU / [k(i, n) - k(1, n)] \end{pmatrix}, \tag{11}$$

where  $\mathbf{1}' = (1 \dots 1)$  is  $(n - 1)$ -dimensional vector,

$$S = (S_{(ij)}) = \text{cov}(T_{(i)}, T_{(j)}) = \frac{1}{k(i, n) - k(1, n)} \frac{1}{k(j, n) - k(1, n)} [\text{cov}(\log(X_{(i)} - \gamma), \log(X_{(j)} - \gamma)) - \text{cov}(\log(X_{(i)} - \gamma), \log(X_{(1)} - \gamma)) - \text{cov}(\log(X_{(j)} - \gamma), \log(X_{(1)} - \gamma)) + \text{var}(\log(X_{(1)} - \gamma))], \quad i, j = 2, \dots, n, \tag{12}$$

$$\text{cov}(\log(X_{(i)} - \gamma), \log(X_{(j)} - \gamma)) = \int_u^1 \int_0^1 \left\{ \log[-\log(1 - u)] - \frac{k(i, n)}{m} \right\} \left\{ \log[-\log(1 - v)] - \frac{k(j, n)}{m} \right\} \times \frac{n!}{(i - 1)!(j - i - 1)!(n - j)!} u^{i-1} (v - u)^{j-i-1} (1 - v)^{n-j} dudv, \tag{13}$$

$$\text{var}(\log(X_{(i)} - \gamma)) = \int_0^1 \left\{ \log[-\log(1 - u)] - \frac{k(i, n)}{m} \right\}^2 \frac{n!}{(i - 1)!(n - i)!} u^{i-1} (1 - u)^{n-i} du. \tag{14}$$

Since  $\eta$  and  $\gamma$  are not included in (11),  $m$  can be estimated first among the three parameters. Note that  $S$  appears to be expressed with  $\gamma$ , but substituting (13) and (14) into (12) gives an expression without  $\gamma$ . Although (13) and (14) seem to depend on  $m$ , they are canceled by substituting into (11), and  $\hat{m}$  can then be estimated without depending on  $m$ . Since a proof by the expansion of (11) is complicated calculations, we show that  $\hat{m}$  has the same value regardless of the setting of  $m$  on the right-hand side of (13) and (14) in Appendix B. The optimal hyperparameter  $c$  in (10) depends on the sample size; thus, its optimal value is searched using Monte Carlo simulations in Section 3.

### 3. Monte Carlo Simulations

#### 3.1 Search for Optimal $c$

We set five cases ( $\gamma = 0, \eta = 1, m = 1, 2, 3, 4, 5$ ) of sample size  $n = 5, 10, 15, 20$  to search for the optimal value of the hyperparameter  $c$ . Table 1 shows the Monte Carlo simulation study results of the bias and root mean squared error (RMSE =  $\sqrt{\text{variance} + \text{bias}^2}$ ) based on 10 000 Monte Carlo simulations for each hyperparameter  $c$ . We used the software R version 3.6.2 (R Core Team, 2019) for the Monte Carlo simulations. It can be seen that the optimal hyperparameter  $c$  values vary greatly with respect to  $n$ , and also change depending on  $m$ , however, the differences are not large. Since  $n$  is known and  $m$  is unknown when data is obtained, considering the bias and RMSE from the Monte Carlo simulation study results, we use the values of  $c = 1.5, 2.3, 3.4,$  and  $4.3$  for  $n = 5, 10, 15,$  and  $20,$  respectively.

Table 1. Bias and RMSE of the shape parameter estimated by the proposed method with various hyperparameter  $c$  values (10 000 Monte Carlo simulations)

$n$	$m$		$c$								
			1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
5	1	Bias	1.135	1.045	0.982	0.932	0.884	0.825	0.789	0.750	0.729
		RMSE	1.411	1.288	1.233	1.162	1.125	1.043	1.014	0.968	0.952
	2	Bias	0.965	0.850	0.754	0.660	0.585	0.497	0.461	0.400	0.326
		RMSE	1.328	1.197	1.115	1.011	0.959	0.865	0.843	0.801	0.731
	3	Bias	0.595	0.458	0.329	0.203	0.104	0.002	-0.067	-0.141	-0.206
		RMSE	1.198	1.089	1.006	0.924	0.877	0.803	0.810	0.794	0.797
	4	Bias	0.127	-0.014	-0.189	-0.322	-0.454	-0.566	-0.662	-0.729	-0.836
		RMSE	1.179	1.159	1.111	1.101	1.069	1.108	1.132	1.161	1.187
	5	Bias	-0.340	-0.563	-0.773	-0.896	-1.049	-1.181	-1.278	-1.373	-1.484
		RMSE	1.427	1.392	1.411	1.474	1.524	1.590	1.643	1.714	1.770
$n$	$m$		$c$								
			1.9	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7
10	1	Bias	0.777	0.751	0.735	0.702	0.676	0.656	0.638	0.616	0.598
		RMSE	0.890	0.860	0.847	0.811	0.782	0.767	0.748	0.726	0.705
	2	Bias	0.653	0.615	0.571	0.538	0.488	0.451	0.421	0.383	0.351
		RMSE	0.859	0.826	0.781	0.755	0.710	0.683	0.655	0.619	0.600
	3	Bias	0.259	0.198	0.139	0.075	0.036	-0.014	-0.046	-0.091	-0.133
		RMSE	0.716	0.678	0.639	0.631	0.611	0.595	0.596	0.598	0.595
	4	Bias	-0.270	-0.326	-0.413	-0.470	-0.541	-0.586	-0.629	-0.691	-0.745
		RMSE	0.796	0.798	0.832	0.842	0.877	0.894	0.919	0.955	0.986
	5	Bias	-0.841	-0.935	-1.019	-1.095	-1.167	-1.228	-1.279	-1.335	-1.398
		RMSE	1.173	1.236	1.293	1.337	1.385	1.434	1.477	1.519	1.564

n	m		c								
			3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8
15	1	Bias	0.574	0.560	0.550	0.535	0.520	0.516	0.506	0.495	0.481
		RMSE	0.648	0.633	0.622	0.606	0.590	0.590	0.582	0.568	0.554
	2	Bias	0.432	0.408	0.384	0.362	0.347	0.322	0.306	0.293	0.262
		RMSE	0.606	0.589	0.575	0.552	0.539	0.521	0.514	0.503	0.480
	3	Bias	-0.009	-0.039	-0.065	-0.086	-0.119	-0.140	-0.165	-0.201	-0.211
		RMSE	0.537	0.526	0.527	0.522	0.531	0.524	0.530	0.536	0.539
	4	Bias	-0.577	-0.601	-0.640	-0.675	-0.702	-0.733	-0.761	-0.787	-0.823
		RMSE	0.839	0.855	0.879	0.901	0.914	0.944	0.958	0.975	1.001
	5	Bias	-1.182	-1.243	-1.270	-1.328	-1.361	-1.387	-1.430	-1.474	-1.488
		RMSE	1.365	1.417	1.442	1.480	1.508	1.531	1.566	1.604	1.614

n	m		c								
			3.9	4.0	4.1	4.2	4.3	4.4	4.5	4.6	4.7
20	1	Bias	0.487	0.480	0.473	0.463	0.458	0.450	0.443	0.435	0.429
		RMSE	0.543	0.535	0.528	0.517	0.514	0.506	0.501	0.491	0.486
	2	Bias	0.344	0.329	0.315	0.306	0.285	0.274	0.262	0.253	0.238
		RMSE	0.509	0.497	0.488	0.478	0.464	0.451	0.447	0.435	0.429
	3	Bias	-0.083	-0.108	-0.126	-0.137	-0.162	-0.185	-0.205	-0.215	-0.240
		RMSE	0.484	0.490	0.490	0.495	0.497	0.497	0.507	0.512	0.517
	4	Bias	-0.668	-0.681	-0.709	-0.731	-0.753	-0.773	-0.784	-0.813	-0.827
		RMSE	0.871	0.881	0.898	0.916	0.933	0.947	0.959	0.981	0.988
	5	Bias	-1.307	-1.327	-1.356	-1.384	-1.411	-1.433	-1.457	-1.472	-1.503
		RMSE	1.455	1.466	1.492	1.515	1.541	1.560	1.578	1.592	1.622

### 3.2 Existing Shape Parameter Estimation Methods

We compare the shape parameter estimated by the proposed method to those by three existing methods. The first method is the BL method proposed by Hall & Wang (2005), which multiplies MLE by an empirical prior. The location and shape parameters are estimated by maximizing

$$BL(m, \gamma) = \frac{X_{(1)} - \gamma}{X_{(2)} - \gamma} \frac{\Gamma(n)m^n \prod_{i=1}^n (X_{(i)} - \gamma)^{m-1}}{\left[ \sum_{i=1}^n (X_{(i)} - \gamma)^m \right]^n}. \tag{15}$$

The second method is the w-MLE method proposed by Cousineau (2009a), which extends MLE technique by incorporating three weights into the solutions to the maximum likelihood equations. The author states that the weight of  $W_3$  is not a constant but changes during the search as new values of the estimating shape parameter are explored. However, the paper used the linear interpolation, because of the lack of a fast computational algorithm. The w-MLE method used for comparison in this paper is based on Cousineau (2009a), but the weight of  $W_3$  is calculated by the following five steps based on Nagatsuka et al. (2013). (i) A temporary shape parameter,  $\hat{m}_0$ , is estimated by MLE for the two-parameter Weibull distribution in  $(X_{(i)} - X_{(1)})$ ,  $i = 2, \dots, n$ . (ii) Uniform random samples,  $(u_1, \dots, u_n)$ , are generated from the standard uniform distribution. (iii) Using  $\hat{m}_0$  and  $(u_1, \dots, u_n)$ , we calculate

$$Y_3 = \frac{1}{n} \sum_{i=1}^n \log \frac{1}{1 - u_i} \frac{\frac{1}{n} \sum_{i=1}^n \left( \log \frac{1}{1 - u_i} \right)^{-1/\hat{m}_0}}{\frac{1}{n} \sum_{i=1}^n \left( \log \frac{1}{1 - u_i} \right)^{1-1/\hat{m}_0}}. \tag{16}$$

(iv) Steps (i)–(iii) are repeated independently 30 000 times. (v) Let  $W_3$  be the median of  $Y_3$ . In the Monte Carlo simulations, Steps (i)–(v) are performed every time random samples from the three-parameter Weibull distribution. The location and shape parameters were estimated by minimizing

$$wMLE(m, \gamma) = \left[ \frac{W_2}{m} + \frac{1}{n} \sum_{i=1}^n \log (X_{(i)} - \gamma) - \frac{\sum_{i=1}^n \log (X_{(i)} - \gamma) (X_{(i)} - \gamma)^m}{\sum_{i=1}^n (X_{(i)} - \gamma)^m} \right]^2 + \left[ \frac{1}{n} \sum_{i=1}^n \frac{1}{X_{(i)} - \gamma} \times \frac{\sum_{i=1}^n (X_{(i)} - \gamma)^m}{\sum_{i=1}^n (X_{(i)} - \gamma)^{m-1}} - W_3 \right]^2. \tag{17}$$

The third method is the LSPF-MLE method (Nagatsuka et al., 2013), in which the shape parameter is estimated by maximizing

$$LSPF(m) = n!m^n \int_0^\infty \int_0^\infty v^{n-2} \left[ \prod_{i=1}^n (u + Z_{(i)}v) \right]^{m-1} \exp \left\{ - \left[ \sum_{i=1}^n (u + Z_{(i)}v)^m \right] \right\} dudv, \tag{18}$$

where  $Z_{(i)} = \frac{X_{(i)} - X_{(1)}}{X_{(n)} - X_{(1)}}$ .

As a feature of each estimation method, the BL and w-MLE methods first estimate the location and shape parameters simultaneously without the scale parameter. The LSPF-MLE method first estimates the shape parameter without the location and scale parameters, which is also the feature of the proposed method.

### 3.3 Effectiveness of the Proposed Method

To demonstrate the validity of the estimations of the shape parameter, five cases ( $\gamma = 0, \eta = 1, m = 1, 2, 3, 4, 5$ ) of sample size  $n = 5, 10, 15, 20$  were investigated. We compared the shape parameter estimated by the proposed method with the shape parameters estimated by the BL, w-MLE, and LSPF-MLE methods in the Monte Carlo simulations. When the shape parameters estimated by the BL, w-MLE, and LSPF-MLE methods were 10 or more, we treated them as estimated at 10 in the Monte Carlo simulations. Since the Monte Carlo simulations were set with  $m \leq 5$ , the bias and RMSE became too large using the estimation result of 10 or more. However, even if the shape parameter estimated by the proposed method was 10 or more, the estimation result was used to calculate the bias and RMSE. The bias and RMSE were calculated from 10 000 Monte Carlo simulations for each set of configurations, and the results are summarized in Table 2. The RMSE of the shape parameter estimated by the proposed method was the smallest of the four methods at  $m = 2, 3, 4, 5$  in the Weibull distribution representing the wear-out failure period. Table 2 also shows the frequencies of the shape parameter estimation of 9.999 or more. In the maximization or minimization of the function using the software R, the frequencies of 9.999 or more were shown because a value slightly less than 10 may be output, when searching in the range where 10 is the upper limit. The shape parameters estimated by the BL and LSPF-MLE methods often exhibited high frequencies of being estimated at 9.999 or more. Figure 1 shows the histogram of the Monte Carlo simulation study results for  $m = 5$  and  $n = 10, 20$ . The biases of the shape parameters estimated by the BL and LSPF-MLE methods seem to be small when  $m = 5$ , but it can be seen from the histogram that the bias became small by setting the maximum value to 10. The proposed method gave estimates that were closest to the population shape parameter at  $m \geq 2$  in the Weibull distribution representing the wear-out failure period.

Table 2. Comparison of the bias and RMSE of the shape parameters estimated by the four methods (10 000 Monte Carlo simulations)

<i>n</i>	<i>m</i>	Proposed			BL			w-MLE			LSPF-MLE		
		Bias	RMSE	Freq	Bias	RMSE	Freq	Bias	RMSE	Freq	Bias	RMSE	Freq
5	1	0.888	1.126	0	1.609	3.257	982	0.572	1.379	36	1.410	3.283	1088
	2	0.581	0.943	0	2.451	4.300	2304	0.340	1.754	114	2.003	4.188	2349
	3	0.108	0.870	4	2.399	4.416	3282	-0.282	1.980	190	1.780	4.249	3078
	4	-0.458	1.083	8	1.913	4.173	3789	-1.029	2.340	232	1.217	4.088	3563
	5	-1.022	1.544	24	1.286	3.940	4319	-1.888	2.885	315	0.517	3.956	3838
10	1	0.672	0.781	0	0.341	1.181	100	0.132	0.458	0	0.332	1.249	111
	2	0.481	0.709	0	1.332	3.127	1135	-0.133	0.792	1	1.004	2.794	884
	3	0.032	0.608	0	1.902	3.945	2566	-0.711	1.223	3	1.299	3.468	1866
	4	-0.976	1.147	0	0.864	3.527	2449	-1.721	1.983	1	0.211	3.156	1772
	5	-1.156	1.384	0	1.027	3.641	3704	-2.255	2.602	13	0.590	3.552	3214
15	1	0.529	0.602	0	0.134	0.454	7	0.069	0.309	0	0.146	0.623	15
	2	0.343	0.539	0	0.540	1.948	343	-0.177	0.587	0	0.556	1.928	316
	3	-0.119	0.529	0	0.947	2.857	1160	-0.721	1.054	1	0.932	2.879	1152
	4	-0.705	0.921	0	1.197	3.362	2293	-1.422	1.700	1	0.941	3.226	2040
	5	-1.354	1.505	0	0.975	3.418	3270	-2.215	2.456	2	0.618	3.304	2817
20	1	0.457	0.512	0	0.081	0.237	0	0.041	0.237	0	0.076	0.319	0
	2	0.285	0.466	0	0.093	1.180	100	-0.181	0.503	0	0.338	1.386	0
	3	-0.159	0.498	0	0.508	2.254	612	-0.684	0.955	0	0.676	2.296	0
	4	-0.742	0.925	0	1.022	3.081	1805	-1.365	1.590	0	0.682	2.654	0
	5	-1.413	1.544	0	0.873	3.206	2786	-2.146	2.349	0	0.339	2.714	0

The proposed method used the values of  $c = 2.3, 3.4,$  and  $4.3$  for  $n = 10, 15,$  and  $20,$  respectively. Freq: Frequencies with which the shape parameter was estimated at 9.999 or more.

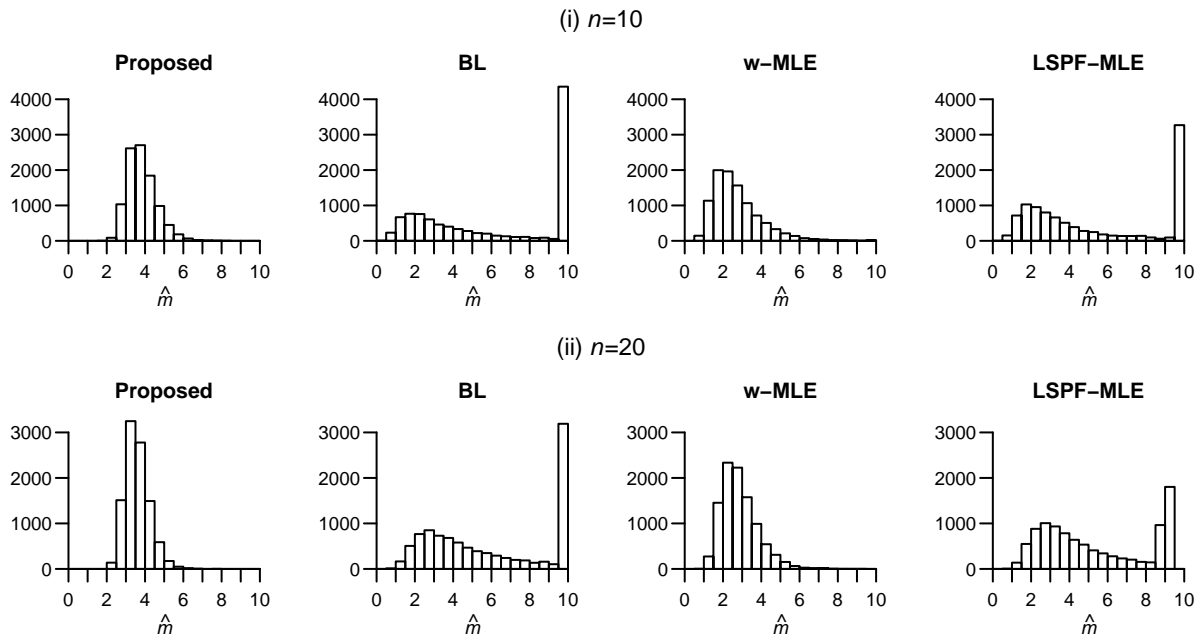


Figure 1. Histogram of the shape parameters estimated by the four methods ( $m = 5, n = 10, 20$ )

### 3.4 Estimation of the Location and Scale Parameters

The location and scale parameters were estimated by the BL, w-MLE, and LSPF-MLE methods using the shape parameter estimated by the proposed method. To distinguish the location and scale parameters estimated by those methods from the original methods, in this paper they are referred to as the modified BL, modified w-MLE, and modified LSPF-MLE methods, respectively. The location and scale parameters estimated by the BL, modified BL, w-MLE, modified w-MLE, LSPF-MLE, and modified LSPF-MLE were compared in the Monte Carlo simulations with the same conditions as in Section 3.3. The bias and RMSE of the location and scale parameters were calculated from 10 000 Monte Carlo simulations for each set of configurations, and the results are summarized in Tables 3 and 4, respectively. Using the shape parameter estimated by the proposed method, the RMSE of the location and scale parameters are often smaller than the RMSE of the original methods at  $m = 2, 3, 4, 5$ . By estimating the shape parameter with a high accuracy, the estimations of the location and scale parameters became highly accurate, too.

Table 3. Comparison of the bias and RMSE of the location parameters estimated by the six methods (10 000 Monte Carlo simulations)

<i>m</i>	BL		M. BL		w-MLE		M. w-MLE		LSPF-MLE		M. LSPF-MLE	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
<i>n</i> = 5												
1	-0.332	1.102	-0.142	0.284	-0.201	0.605	-0.325	0.395	-0.118	0.450	-0.131	0.271
2	-0.403	1.212	0.009	0.270	-0.086	0.651	-0.123	0.289	0.049	0.417	0.093	0.288
3	-0.200	0.939	0.161	0.293	0.119	0.583	0.045	0.245	0.240	0.432	0.271	0.368
4	-0.033	0.769	0.252	0.338	0.228	0.586	0.146	0.267	0.354	0.477	0.376	0.440
5	0.092	0.664	0.324	0.385	0.321	0.613	0.226	0.307	0.443	0.528	0.459	0.505
<i>n</i> = 10												
1	-0.022	0.372	-0.138	0.191	-0.021	0.144	-0.260	0.285	-0.049	0.205	-0.137	0.184
2	-0.291	0.856	-0.044	0.191	0.085	0.249	-0.117	0.212	0.003	0.307	0.012	0.195
3	-0.337	0.887	0.086	0.211	0.214	0.334	0.028	0.187	0.117	0.337	0.161	0.261
4	-0.335	0.883	0.085	0.210	0.214	0.331	0.028	0.186	0.119	0.334	0.160	0.260
5	-0.053	0.614	0.266	0.320	0.401	0.469	0.221	0.280	0.321	0.422	0.362	0.409
<i>n</i> = 15												
1	0.041	0.123	-0.093	0.132	-0.003	0.088	-0.179	0.198	-0.024	0.120	-0.098	0.131
2	-0.066	0.505	-0.028	0.154	0.088	0.193	-0.075	0.162	0.001	0.246	0.004	0.162
3	-0.127	0.624	0.091	0.191	0.207	0.288	0.057	0.170	0.082	0.296	0.134	0.228
4	-0.154	0.643	0.189	0.254	0.306	0.369	0.161	0.228	0.174	0.330	0.240	0.304
5	-0.085	0.571	0.273	0.317	0.392	0.440	0.247	0.292	0.267	0.374	0.330	0.376
<i>n</i> = 20												
1	0.038	0.064	-0.072	0.101	0.001	0.062	-0.140	0.155	-0.013	0.072	-0.077	0.100
2	0.069	0.300	-0.025	0.134	0.083	0.164	-0.061	0.139	0.003	0.205	-0.001	0.142
3	-0.036	0.492	0.085	0.174	0.194	0.259	0.061	0.157	0.063	0.265	0.114	0.205
4	-0.133	0.578	0.187	0.242	0.296	0.346	0.168	0.223	0.152	0.299	0.221	0.282
5	-0.093	0.539	0.265	0.306	0.375	0.416	0.249	0.288	0.235	0.341	0.304	0.350

M. BL: Modified BL, M. w-MLE: Modified w-MLE, M. LSPF-MLE: Modified LSPF-MLE.

Table 4. Comparison of the bias and RMSE of the scale parameters estimated by the six methods (10 000 Monte Carlo simulations)

<i>m</i>	BL		M. BL		w-MLE		M. w-MLE		LSPF-MLE		M. LSPF-MLE	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
<i>n</i> = 5												
1	0.291	0.928	0.270	0.552	0.266	0.918	0.498	0.717	0.142	0.820	0.261	0.580
2	0.303	1.012	-0.013	0.285	0.068	0.718	0.138	0.321	-0.058	0.540	-0.089	0.306
3	0.159	0.915	-0.184	0.296	-0.154	0.622	-0.057	0.242	-0.260	0.498	-0.283	0.369
4	0.010	0.780	-0.275	0.344	-0.261	0.616	-0.161	0.266	-0.371	0.519	-0.387	0.440
5	-0.107	0.680	-0.348	0.395	-0.357	0.631	-0.244	0.311	-0.460	0.558	-0.471	0.508
<i>n</i> = 10												
1	0.033	0.438	0.296	0.454	0.047	0.408	0.434	0.561	0.056	0.462	0.295	0.470
2	0.252	0.824	0.055	0.224	-0.109	0.317	0.136	0.253	-0.009	0.406	0.006	0.229
3	0.317	0.901	-0.096	0.218	-0.244	0.373	-0.033	0.193	-0.118	0.394	-0.163	0.264
4	0.314	0.899	-0.097	0.216	-0.245	0.371	-0.034	0.190	-0.123	0.390	-0.164	0.262
5	0.040	0.630	-0.281	0.328	-0.427	0.496	-0.232	0.285	-0.321	0.440	-0.367	0.409
<i>n</i> = 15												
1	-0.022	0.282	0.238	0.367	0.017	0.308	0.334	0.439	0.025	0.334	0.243	0.380
2	0.049	0.512	0.037	0.189	-0.112	0.254	0.089	0.201	-0.008	0.328	0.009	0.197
3	0.116	0.640	-0.101	0.199	-0.233	0.321	-0.062	0.177	-0.080	0.343	-0.139	0.234
4	0.144	0.659	-0.203	0.262	-0.332	0.396	-0.172	0.234	-0.169	0.358	-0.249	0.309
5	0.076	0.585	-0.287	0.327	-0.415	0.464	-0.259	0.300	-0.262	0.389	-0.339	0.381
<i>n</i> = 20												
1	-0.025	0.238	0.209	0.319	0.007	0.258	0.285	0.376	0.012	0.268	0.214	0.329
2	-0.085	0.313	0.036	0.168	-0.103	0.217	0.075	0.176	-0.008	0.272	0.015	0.175
3	0.027	0.505	-0.094	0.184	-0.217	0.289	-0.067	0.166	-0.060	0.305	-0.120	0.212
4	0.125	0.594	-0.200	0.252	-0.318	0.370	-0.179	0.231	-0.147	0.322	-0.231	0.288
5	0.086	0.552	-0.279	0.316	-0.396	0.438	-0.261	0.297	-0.230	0.353	-0.314	0.357

M. BL: Modified BL, M. w-MLE: Modified w-MLE, M. LSPF-MLE: Modified LSPF-MLE.

#### 4. Numerical Example

We analyzed a dataset of 20 voltage breakdowns (Kako, 1986). The dataset, which is also discussed by Hirose & Lai (1997), is as follows:

2.1, 3.4, 3.7, 2.7, 3.2, 3.5, 2.7, 2.7, 3.4, 3.3, 3.0, 2.9, 3.0, 3.4, 2.8, 3.5, 3.5, 3.4, 3.4, 3.6 ( $\times 100$  kV).

We used the software R for estimating the shape parameters by the four methods (the proposed, BL, w-MLE, and LSPF-MLE methods), and a sample code is shown in Appendix A. From the Monte Carlo simulation study results at  $n = 20$  in Section 3.1, we adopted  $c = 3.4$  as the hyperparameter of the proposed method. The shape parameter was estimated to be 4.048 by the proposed method and 4.772 by the w-MLE method, respectively. Note that since the w-MLE method calculates the median by the Monte Carlo simulations, the estimated value will change slightly each time. The shape parameters were not estimated by the BL and LSPF-MLE methods, because  $BL(m, \gamma)$  and  $LSPF(m)$  increased monotonically with  $m$ . Since the upper limit in the sample code was set to 10, they were estimated at 10 or less, but it can be seen that the shape parameter was estimated to be infinite, when the initial value and the upper limit were changed. A graph can be drawn to make sure they are monotonically increasing. Since the shape parameters estimated by the proposed and w-MLE methods had a negative bias from the Monte Carlo simulation study results at  $m \geq 3$  and  $n = 20$ , the population shape parameter was considered to be  $m \geq 4$ . As  $m \geq 2$  was assumed from the numerical example, we can recommend the use of the shape parameter estimated by the proposed method.

#### 5. Conclusions

We proposed an unbiased estimator of the shape parameter that did not depend on the location and scale parameters. The proposed method was expressed including the hyperparameter, and the optimal hyperparameter depended on the sample size. The procedure of the proposed method first looks for the optimal hyperparameter from the data obtained, and then estimates the shape parameter. The Monte Carlo simulation study results showed that the RMSE of the shape parameter estimated by the proposed method was smallest of the four methods when the population shape parameter was 2 or more in the Weibull distribution representing the wear-out failure period. In particular, the difference between the RMSE of the shape parameter estimated by the proposed method and the RMSE of the shape parameters estimated by the other three methods increased with  $m$ . The proposed method is recommended, when the population shape parameter of 2 or more is assumed, as in the numerical example.

#### Acknowledgements

The authors wish to thank the Editor and two reviewers for their valuable comments and suggestions, which have led to an improved version of the paper.

#### References

- Ahsanullah, M., Nevzorov, V., & Shakil, M. (2013). Minimum Variance Linear Unbiased Estimators. *In: An Introduction to Order Statistics* (pp.139-151), Paris: Atlantis Press. [https://doi.org/10.2991/978-94-91216-83-1\\_13](https://doi.org/10.2991/978-94-91216-83-1_13)
- Cousineau, D. (2009a). Nearly unbiased estimators for the three-parameter Weibull distribution with greater efficiency than the iterative likelihood method. *British Journal of Mathematical and Statistical Psychology*, 62(1), 167-191. <https://doi.org/10.1348/000711007X270843>
- Cousineau, D. (2009b). Fitting the three-parameter Weibull distribution: Review and evaluation of existing and new methods. *IEEE Transactions on Dielectrics and Electrical Insulation*, 16(1), 281-288. <https://doi.org/10.1109/TDEI.2009.4784578>
- Hall, P., & Wang, J. Z. (2005). Bayesian likelihood methods for estimating the end point of a distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(5), 717-729. <https://doi.org/10.1111/j.1467-9868.2005.00523.x>
- Hirose, H., & Lai, T. L. (1997). Inference from grouped data in three-parameter Weibull models with applications to breakdown-voltage experiments. *Technometrics*, 39(2), 199-210. <https://doi.org/10.2307/1270908>
- Kako, Y. (1986). Estimation of Weibull parameters for BDV data of transformer oil by most likelihood methods *In Conference Internationale des Grands Reseaux Electriques* (Vol. 2).
- Lawless, J. F. (2012). *Statistical models and methods for lifetime data* 2nd edition. Hoboken: John Wiley & Sons.
- Murthy, D. N. P., Xie, M., & Jiang, R. (2004). *Weibull models*. New York: John Wiley & Sons.
- Nagatsuka, H., & Kamakura, T. (2005). Parameter estimation of the shape parameter of the Castillo-Hadi model. *Communications in Statistics-Theory and Methods*, 33(1), 15-27. <https://doi.org/10.1081/STA-120026574>



- Nagatsuka, H., Kamakura, T., & Balakrishnan, N. (2013). A consistent method of estimation for the three-parameter Weibull distribution. *Computational Statistics & Data Analysis*, 58, 210-226.  
https://doi.org/10.1016/j.csda.2012.09.005
- Ng, H. K. T., Luo, L., Hu, Y., & Duan, F. (2012). Parameter estimation of three-parameter Weibull distribution based on progressively type-II censored samples. *Journal of Statistical Computation and Simulation*, 82(11), 1661-1678.  
https://doi.org/10.1080/00949655.2011.591797
- R Core Team. (2019). *R: A language and environment for statistical computing* (Version 3.6.2). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/
- Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72(1), 67-90.  
https://doi.org/10.1093/biomet/72.1.67
- Sugiyama, T., Ogura, T., & Sugiyama, T. (2019). Estimation of the threshold parameter of a wear-out failure period in the case of a three-parameter Weibull distribution. *International Journal of Knowledge Engineering and Soft Data Paradigms*, 6(2), 110-119. https://doi.org/10.1504/IJKESDP.2019.102818
- Yamamoto, W., Kumazaki, C., & Suzuki, K. (2010). Estimation of archival lifetime distribution for writable optical disks from accelerated testings. In *Mathematical and Statistical Models and Methods in Reliability* (pp. 307-313), Birkhäuser Boston. https://doi.org/10.1007/978-0-8176-4971-5\_24

## Appendix A

### Sample code of the software R

We show a sample code of the software R for estimating the shape parameter using the numerical example in Section 4 by the proposed, BL, w-MLE, and LSPF-MLE methods. When maximizing or minimizing a function, it is advisable to search several times with different initial values.

```
X<-sort(c(2.1,3.4,3.7,2.7,3.2,3.5,2.7,2.7,3.4,3.3,3.0,2.9,3.0,3.4,2.8,3.5,3.5,3.4,3.4,3.6))
n<-length(X)

#Proposed method
k<-function(i,n){f<-function(u){log(-log(1-u))*factorial(n)/factorial(i-1)/
  factorial(n-i)*u^(i-1)*(1-u)^(n-i)}; integrate(f,0,1)$value}
cov_ij<-function(i,j,n,m){ki<-function(u){
  log(-log(1-u))*factorial(n)/factorial(i-1)/factorial(n-i)*u^(i-1)*(1-u)^(n-i)}
  kj<-function(u){log(-log(1-u))*factorial(n)/factorial(j-1)/factorial(n-j)*u^(j-1)*
  (1-u)^(n-j)}; kin<-integrate(ki,0,1)$value; kjn<-integrate(kj,0,1)$value
  g<-function(u){f<-function(v){(log(-log(1-u))-kin/m)*(log(-log(1-v))-kjn/m)*
  factorial(n)/factorial(i-1)/factorial(j-1)/factorial(n-j)*u^(i-1)*(v-u)^(j-i-1)*
  (1-v)^(n-j)}; integrate(f,u,1)$value}; integrate(Vectorize(g),0,1)$value}
var_i<-function(i,n,m){ki<-function(u){log(-log(1-u))*factorial(n)/factorial(i-1)/
  factorial(n-i)*u^(i-1)*(1-u)^(n-i)}; kin<-integrate(ki,0,1)$value
  f<-function(u){(log(-log(1-u))-kin/m)^2*factorial(n)/
  factorial(i-1)/factorial(n-i)*u^(i-1)*(1-u)^(n-i)}; integrate(f,0,1)$value}
S<-function(n,m){S0<-matrix(NA,n-1,n-1); for(i in 2:n){for(j in 2:n){if(i<j){S0[i-1,j-1]
  ]<-(cov_ij(i,j,n,m)-cov_ij(1,i,n,m)-cov_ij(1,j,n,m)+var_i(1,n,m))/(k(i,n)-k(1,n))/(k(
  j,n)-k(1,n)); if(j<i){S0[i-1,j-1]<-(cov_ij(j,i,n,m)-cov_ij(1,i,n,m)-cov_ij(1,j,n,m)+
  var_i(1,n,m))/(k(i,n)-k(1,n))/(k(j,n)-k(1,n)); if(i==j){S0[i-1,j-1]<-(var_i(i,n,m)-
  2*cov_ij(1,i,n,m)+var_i(1,n,m))/(k(i,n)-k(1,n))^2}}; return(S0)}
proposed<-function(c,m0){T<-function(i){T0<-function(U){(log(c*(1-U)*(X[i]-X[1])+U)-log(
  U))/(k(i,n)-k(1,n))}; integrate(T0,0,1)$value}; vT<-NULL; for(i in 2:n){vT<-c(vT,T(i))}
  rep(1,n-1)%%solve(S(n,m0))%rep(1,n-1)/rep(1,n-1)%%solve(S(n,m0))%vT}
c(proposed(4.3,3),proposed(4.3,5),proposed(4.3,7),proposed(4.3,9))

#BL method
BL_f<-function(a){(X[1]-a[2])/(X[2]-a[2])*gamma(n)*a[1]^n*prod((X-a[2])^(a[1]-1))/
  (sum((X-a[2])^a[1]))^n}
optim(c(5,1),BL_f,control=list(fnscale=-1),method="L-BFGS-B",lower=c(0,-Inf),
  upper=c(10,X[1]))$par[1]

#w-MLE method
```

```

library(MASS); MLE0<-fitdistr((X-X[1])[-1],densfun="weibull",lower=c(0,0))
m0<-MLE0$estimate[1]; eta0<-MLE0$estimate[2]; gamma0<-X[1]-eta0/n^(1/m0)*gamma(1+1/m0)
simul2<-100000; simul3<-30000; Y2<-NULL; Y3<-NULL; for(i in 1:simul2){u<-1/(1-runif(n))
  Y2<-c(Y2,sum(log(u)*log(log(u)))/sum(log(u))-1/n*sum(log(log(u))))}; W2<-median(Y2)
for(i in 1:simul3){u<-1/(1-runif(n)); Y3<-c(Y3,mean(log(u))*sum((log(u))^(1/m0))/
  sum((log(u))^(m0-1/m0)))}; W3<-median(Y3)
wMLE_f<-function(a){W2/a[1]+1/n*(sum(log(X-a[2]))-sum(log(X-a[2])*(X-a[2])^a[1])/sum(
  (X-a[2])^a[1]))^2+(1/n*sum(1/(X-a[2]))*sum((X-a[2])^a[1])/sum((X-a[2])^(a[1]-1))-W3)^2}
optim(c(m0,gamma0),wMLE_f,method="L-BFGS-B",lower=c(0,-Inf),upper=c(10,X[1]))$par[1]

#LSPF-MLE method
library(cubature); Z<-(X-X[1])/(X[n]-X[1])
LSPF_f<-function(m){g<-function(a){a[2]^(n-2)*(prod(a[1]+Z*a[2]))^(m-1)*exp(-(sum(
  a[1]+Z*a[2])^m))}; adaptIntegrate(g,c(0,0),c(Inf,Inf))$integral*m^n*factorial(n)}
optimize(LSPF_f,c(0,10),maximum=T)$maximum

```

## Appendix B

### Proof that the same result can be obtained regardless of the setting of $m$ on the right-hand side of (13) and (14)

In the sample code of the software R in Appendix A, the last line of the proposed method is set to  $m = 3, 5, 7, 9$  on the right-hand side of (13) and (14). In every case, the shape parameter by the proposed method is estimated to be  $\hat{m} = 4.048$ , and, therefore, the right-hand side of (11) does not depend on  $m$ .

## Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).