Communications for Statistical Applications and Methods
2023, Vol. 30, No. 4, 411–421

https://doi.org/10.29220/CSAM.2023.30.4.411
Print ISSN 2287-7843 / Online ISSN 2383-4757

Novel estimation based on a minimum distance under the progressive Type-II censoring scheme

Young Eun Jeon[a], Suk-Bok Kang[a], Jung-In Seo[1,b]

[a]Department of Statistics, Yeungnam University, Korea;
[b]Department of Information Statistics, Andong National University, Korea

Abstract

This paper provides a new estimation equation based on the concept of a minimum distance between the empirical and theoretical distribution functions under the most widely used progressive Type-II censoring scheme.

For illustrative purposes, simulated and real datasets from a three-parameter Weibull distribution are analyzed. For comparison, the most popular estimation methods, the maximum likelihood and maximum product of spacings estimation methods, are developed together. In the analysis of simulated datasets, the excellence of the provided estimation method is demonstrated through the degree of the estimation failure of the likelihood-based method, and its validity is demonstrated through the mean squared errors and biases of the estimators obtained from the provided estimation equation. In the analysis of the real dataset, two types of goodness-of-fit tests are performed on whether the observed dataset has the three-parameter Weibull distribution under the progressive Type-II censoring scheme, through which the performance of the new estimation equation provided is examined.

Keywords: Cramer-von Mises distance, maximum product of spacings, progressive Type-II censored sample, Weibull distribution

1. Introduction

To estimate unknown parameters in a probability distribution, the maximum likelihood estimation method with useful properties such as consistency and sufficiency may generally be considered. However, the maximum likelihood estimation method can break down not only for heavy-tailed continuous distributions with unknown location and scale parameters as discussed in Pitman (1979) but also for mixtures of continuous distributions as mentioned in Ranneby (1984). The maximum likelihood estimation method is also bound to fail for distributions with unknown endpoints and J-shaped density function as mentioned in Cheng and Amin (1983). Moreover, for a certain three-parameter distribution where one parameter is an unknown shifted origin, the likelihood can become infinite, and it can lead to inconsistent estimators of other parameters.

A typical example is a Weibull distribution which is extensively used for reliability analysis. Suppose that a random variable X has the three-parameter Weibull distribution with the following probability density function (PDF) and cumulative distribution function (CDF)

$$f(x;\mu,\sigma,\lambda) = \lambda\sigma^{-\lambda}(x-\mu)^{\lambda-1}\, e^{-\left(\frac{x-\mu}{\sigma}\right)^{\lambda}},$$

This work was supported by a grant from 2022 Research Fund of Andong National University

[1] Corresponding author: Department of Information Statistics, Andong National University, 1375 Gyeongdong-ro (SongCheon-dong), Andong, Gyeongsangbuk-do 36729, Korea. E-mail: leehoo1928@gmail.com

Published 31 July 2023 / journal homepage: http://csam.or.kr

and

$$F(x;\mu,\sigma,\lambda) = 1 - e^{-\left(\frac{x-\mu}{\sigma}\right)^{\lambda}}, \quad x > \mu,\ \sigma > 0,\ \lambda > 0, \tag{1.1}$$

respectively, where μ is the location parameter, σ is the scale parameter, and λ is the shape parameter that allows various shapes. Then, the Weibull distribution does not satisfy one of the regularity conditions since the domain of the random variable X depends on the location parameter μ (Cousineau, 2009). In particular, the Weibull distribution has the J-shaped density function when the shape parameter λ is less than 1, which may yield invalid maximum likelihood estimators (MLEs) or inconsistency of the MLEs. In addition, the MLEs do not satisfy asymptotic normality when the shape parameter is between 1 and 2.

Recently, as an alternative to resolving these problems, several methods have been introduced. In particular, the maximum product of spacings (MPS) estimation method has received considerable attention because of providing consistent estimators with asymptotic efficiency. The MPS estimation method was independently introduced by Cheng and Amin (1983) and Ranneby (1984). Moreover, Cheng and Amin (1983) proved that the MPS estimation method gives consistent and asymptotically efficient estimators not only when the MLEs exist but also when the maximum likelihood estimation method is bound to fail, and they argued that the MPS estimators (MPSEs) are as efficient as the MLEs. Ranneby (1984) studied the MPS estimation method as an approximation of the Kullback-Leibler information and showed the consistency of the MPSE. In addition, the consistency was generalized by Ekström (1998). Anatolyev and Kosenok (2005) demonstrated that the MPSE is more efficient than the MLE in small samples for a heavy tail and/or skewed distribution. Based on this, Singh *et al.* (2014) and Singh *et al.* (2018) applied the MPS estimation method to the generalized inverted exponential and exponentiated Pareto distributions in the case where there is no censoring, respectively, and they proved the excellence of the MPS estimation method through a simulation. Studies related to the censoring scheme have also proven the excellence of the MPS estimation method by several authors. Basu *et al.* (2017) studied the MPS estimation method for the inverse Lindley distribution under the Type-I censoring scheme. Basu *et al.* (2018) studied the MPS estimation method for the same distribution under a progressive hybrid Type-I censoring scheme with the probability of removals. Jeon *et al.* (2022) provided the MPSEs for the three-parameter Weibull distribution based on a generalized Type-II progressive hybrid censored sample.

Another alternative is to use the minimum distance estimation, which finds unknown parameters where the distance (difference) between the empirical distribution function (EDF) and the theoretical distribution function is minimized. This approach is especially useful for estimating the location parameter as well as being closely related to goodness-of-fit tests including Kolmogorov-Smirnov (K-S), Cramer-von Mises (CVM), and Anderson-Darling (A-D) tests. Successful cases of applying the minimum distance approach in estimating the location parameter of the probability distributions have been reported in some literature. Hobbs *et al.* (1984) studied the minimum distance estimator (MDE) of the location parameter in a three-parameter gamma distribution using the K-S, CVM, and A-D distances. Similarly, Hobbs *et al.* (1985) applied the minimum distance estimation method to the three-parameter Weibull distribution and stated that it provides a valid approach for s-robust and s-consistent estimation of the location parameter. Gallagher and Moore (1990) investigated robust estimating techniques for the location parameter in the three-parameter Weibull distribution.

As mentioned above, the minimum distance estimation method is related to goodness-of-fit tests such as K-S, CVM, and A-D tests. These goodness-of-fit tests use a statistic based on the distance between the EDF and the theoretical distribution function, which indicates that the smaller the statistic, the closer the distance between the two distribution functions is. Among them, the K-S test is the most popular goodness-of-fit test based on the EDF, but empirical evidence suggests that the CVM test is usually more powerful than the K-S test.

Due to this reason, we propose a novel approach based on the CVM distance by incorporating the minimum distance estimation method into the progressive Type-II censoring scheme that plays a very important role in planning duration experiments in reliability studies to save time and money. Moreover, a comparison is made with the maximum likelihood and MPS estimation methods to show the excellence of the proposed approach.

The rest of the paper is organized as follows. Section 2 provides the minimum distance estimation method for estimating unknown parameters in the three-parameter Weibull distribution under the progressive Type-II censoring scheme. Section 3 deploys the maximum likelihood and MPS estimation methods for comparison with the proposed approach. Section 4 provides two types of goodness-of-fit tests based on the replicated data and pivotal quantity. Section 5 provides the results of the simulation and real data analysis, and Section 6 concludes this paper.

2. Minimum distance estimation

The progressive Type-II censoring scheme is explained as follows: In the first failure, R_1 units out of $n-1$ survival units are randomly withdrawn (or censored) from the lifetime experiment. In the same manner, when the next failure occurs, R_2 units among the remaining $n-2-R_1$ survival units are randomly withdrawn. This process takes place up to m^{th} failure, and all remaining $R_m = n - m - R_1 - \cdots - R_{m-1}$ units are censored at the time of the m^{th} failure. Here, $R_1, \ldots, R_m$ are fixed. It is worth noting that it is the complete sample situation for the case $m = n$ where $R_1 = \cdots = R_m = 0$, while it is the conventional Type-II censoring scheme in the case of $R_1 = \cdots = R_{m-1} = 0$, $R_m = n - m$.

To obtain the MDE under the progressive Type-II censoring scheme, the CVM distance for the considered censoring scheme is required. It can be derived by extending the following CVM distance under the no censoring situation

$$W^2(\theta) = n \int_{-\infty}^{\infty} [H_n(x) - H(x;\theta)]^2 \, dH(x;\theta),$$

where $H_n(x)$ and $H(x;\theta)$ are the EDF and the theoretical distribution function with an unknown parameter θ, respectively.

Let

$$X_{1:m:n} \leq \cdots \leq X_{m:m:n} \tag{2.1}$$

be a progressive Type-II censored sample with the censoring scheme $\mathscr{R} = (R_1, \ldots, R_m)$. The EDF based on the progressive Type-II censored sample (2.1) stated in Pakyari and Balakrishnan (2012) is given by

$$H_{m:n}(x) = \begin{cases} 0, & \text{if } x < x_{1:m:n}, \\ p_{i:m:n}, & \text{if } x_{i:m:n} \leq x < x_{i+1:m:n}, \quad i = 1, \ldots, m-1, \\ p_{m:m:n}, & \text{if } x \geq x_{m:m:n}, \end{cases}$$

where $p_{i:m:n} = E(U_{i:m:n}) = 1 - \prod_{j=1}^{i} \gamma_j (1+\gamma_j)^{-1}$ is the mean of the i^{th} order statistic from a standard uniform distribution with the interval $(0, 1)$ under the progressive Type-II censoring scheme. Here, $\gamma_j = \sum_{k=j}^{m} (1 + R_k)$. It is noted that $H_{m:n}(x) \neq 1$ for $x \geq x_{m:m:n}$ since $n - m$ samples are censored, unlike the no censoring case.

Then, the CVM distance can be derived under the progressive Type-II censoring scheme according to Pakyari and Balakrishnan (2012) as

$$W^2_{m:n}(\theta) = n\sum_{i=1}^{m-1} p_{i:m:n}\left(H(x_{i+1:m:n};\theta) - H(x_{i:m:n};\theta)\right)\left(p_{i:m:n} - H(x_{i+1:m:n};\theta) - H(x_{i:m:n};\theta)\right) + \frac{n}{3}\left(H(x_{m:m:n};\theta)\right)^3, \tag{2.2}$$

where $H(x_{i:m:n};\theta)$ denotes the theoretical distribution function with the progressive Type-II censored sample. The MDE of the unknown parameter θ can be derived by minimizing the CVM distance (2.2). That is, $\hat{\theta}_{\text{CVM}} = \arg\min_{\theta} W^2_{m:n}(\theta)$. For the three-parameter Weibull distribution with the CDF (1.1), the CVM distance (2.2) is written as

$$W^2_{m:n}(\mu,\sigma,\lambda) = n\sum_{i=1}^{m-1} p_{i:m:n}\left(e^{-z^{\lambda}_{i:m:n}} - e^{-z^{\lambda}_{i+1:m:n}}\right)\left(p_{i:m:n} - 2 + e^{-z^{\lambda}_{i+1:m:n}} + e^{-z^{\lambda}_{i:m:n}}\right) + \frac{n}{3}\left(1 - e^{-z^{\lambda}_{m:m:n}}\right)^3 \tag{2.3}$$

by letting $Z_{i:m:n} = (X_{i:m:n} - \mu)/\sigma$. To obtain the MDEs of μ, σ, and λ from the CVM distance (2.3), two types of algorithms are applied: One is the algorithm introduced by Hobbs *et al.* (1985), and the other is the algorithm introduced by Hooke and Jeeves (1961). The former algorithm proceeds as follows: First, μ is estimated by minimizing $W^2_{m:n}(\mu, \hat{\sigma}_{\text{MLE}}, \hat{\lambda}_{\text{MLE}})$, where $\hat{\sigma}_{\text{MLE}}$ and $\hat{\lambda}_{\text{MLE}}$ are the MLEs of σ and λ, respectively. Then, σ and λ are re-estimated using the maximum likelihood estimation method provided in Section 3, given the MDE of μ. The resulting estimators are denoted as $\hat{\mu}_{\text{CVM1}}$, $\hat{\sigma}_{\text{CVM1}}$, and $\hat{\lambda}_{\text{CVM1}}$. The latter is the derivative-free algorithm without requiring the derivative to find unknown parameters where the given function is minimized. The resulting estimators are denoted as $\hat{\mu}_{\text{CVM2}}$, $\hat{\sigma}_{\text{CVM2}}$, and $\hat{\lambda}_{\text{CVM2}}$.

3. Classical estimation

For comparison with the proposed approach, this section provides the maximum likelihood estimation method as well as the MPS estimation method which gives consistent and asymptotically efficient estimators.

The MLEs $\hat{\mu}_{\text{MLE}}$, $\hat{\sigma}_{\text{MLE}}$, and $\hat{\lambda}_{\text{MLE}}$ can be found by maximizing the natural logarithm of the likelihood function based on the progressive Type-II censored sample (2.1), given by

$$L(\mu,\sigma,\lambda) = C\prod_{i=1}^{m} f(x_{i:m:n};\mu,\sigma,\lambda)\left[1 - F(x_{i:m:n};\mu,\sigma,\lambda)\right]^{R_i} \propto \lambda^m\sigma^{-\lambda m} e^{-\sum_{i=1}^{m}(1+R_i)z^{\lambda}_{i:m:n}} \prod_{i=1}^{m}(x_{i:m:n} - \mu)^{\lambda-1},$$

where $C = n(n-1-R_1)\cdots(n-m+1-R_1-\cdots-R_{m-1})$.

To derive the MPSE, assume that the PDF $h(x;\theta)$ corresponding to the CDF $H(x;\theta)$ is absolutely positive in the interval (a,b) and 0 elsewhere. Additionally, let $x_1 < \cdots < x_n$ be an ordered random sample. According to Cheng and Amin (1983), in any continuous univariate distribution, the MPSE of the unknown parameter θ is derived by maximizing the geometric mean of the spacings defined as $G(\theta) = (\prod_{i=1}^{n+1} S_i(\theta))^{1/(n+1)}$, where $S_i(\theta) = H(x_i;\theta) - H(x_{i-1};\theta)$, $i = 1,\ldots,n+1$, $x_0 \equiv a$ and $x_{n+1} \equiv b$ under the no censoring situation. However, since censoring can occur by various causes, a new product of spacings function is required to obtain an appropriate MPSE.

For the progressive Type-II censoring scheme, the sample information induces the partitions in the interval $(0,\infty)$ as follows: $(0, x_{1:m:n}], (x_{1:m:n}, x_{2:m:n}], \ldots, (x_{m:m:n}, \infty)$. Based on the partition $(x_{i-1:m:n}, x_{i:m:n}]$, the spacing is given by $S_{i:m:n}(\theta) = H(x_{i:m:n};\theta) - H(x_{i-1:m:n};\theta)$, $i = 1,\ldots,m$ with $H(x_{0:m:n};\theta) = 0$. However, the spacing should be modified because R_i units are removed at $x_{i:m:n}$. Moreover, the information of R_i censored units may be considered as ties in terms of the survival function $(1 - H(x_{i:m:n};\theta))$, assigning the same probability to R_i units by the idea of Shao and Hahn (1999). Using this fact, the product of spacings function based on the progressive Type-II censored sample (2.1) is derived as

$$G^*(\theta) = \prod_{i=1}^{m} S_{i:m:n}(\theta) \left[\frac{1 - H(x_{i:m:n};\theta)}{R_i}\right]^{R_i}. \tag{3.1}$$

Then, for the three-parameter Weibull distribution, the MPSEs $\hat{\mu}_{\text{MPS}}$, $\hat{\sigma}_{\text{MPS}}$, and $\hat{\lambda}_{\text{MPS}}$ are derived by maximizing the natural logarithm of the product of spacings function (3.1), given by

$$\log G^*(\mu,\sigma,\lambda) \propto \sum_{i=1}^{m} \log\left(e^{-z_{i-1:m:n}^{\lambda}} - e^{-z_{i:m:n}^{\lambda}}\right) - \sum_{i=1}^{m} R_i z_{i:m:n}^{\lambda}.$$

4. Goodness-of-fit test

The validity of the proposed method can be demonstrated through the goodness-of-fit test of the observed progressive Type-II censored sample. This section provides two approaches, which are based on the techniques introduced by Seo *et al*. (2020).

The provided goodness-of-fit test methods are applied to real data analysis to illustrate in Section 5.2.

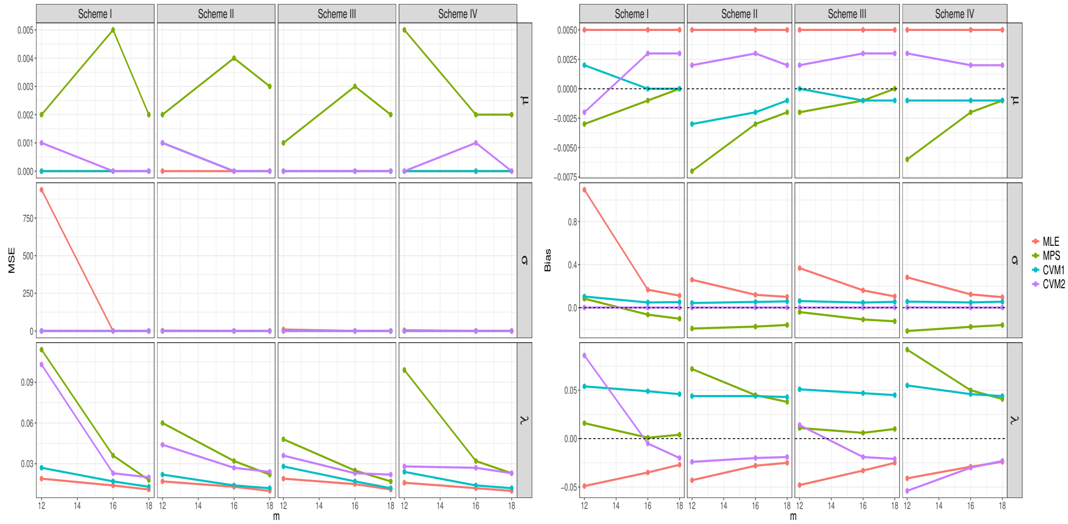
4.1. Replicated data-based

Let X^{rep} be a replicated data which means a replication of an observed data x. Then, the replicated data X^{rep} is generated from the distribution $p(x^{\text{rep}}|x,\tilde{\theta})$ that is a fitted model with the observed data x and an estimator $\tilde{\theta}$. This indicates that the replicated data is very similar to the observed data if the fitted model has a high goodness-of-fit. Here, the degree of the goodness-of-fit depends on how close the true value of the parameter to be estimated and an estimated value of the parameter are.

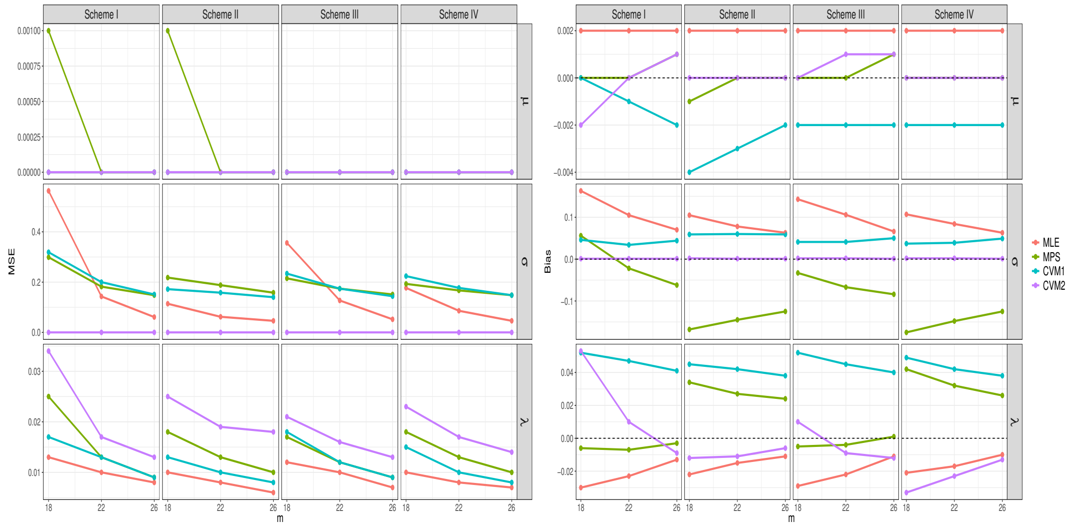
Under the progressive Type-II censoring scheme, the replicated data is denoted as $X_{i:m:n}^{\text{rep}}$, and it is randomly generated from the fitted model $p(x_{i:m:n}^{\text{rep}}|x_{i:m:n},\tilde{\theta})$. Then, the goodness-of-fit test can be performed by plotting the scatter plot between the observed progressive Type-II censored sample $\{x_{1:m:n},\ldots,x_{m:m:n}\}$ and the expectation of the replicated data

$$E\left(X_{i:m:n}^{\text{rep}}\right) = \frac{1}{N}\sum_{k=1}^{N} X_{i:m:n}^{\text{rep},k}, \quad i = 1,\ldots,m,$$

where N denotes the number of the replicated data generated. If the points are close to a straight line in the scatter plot, then it indicates that a correlation between the observed progressive Type-II censored



(a)



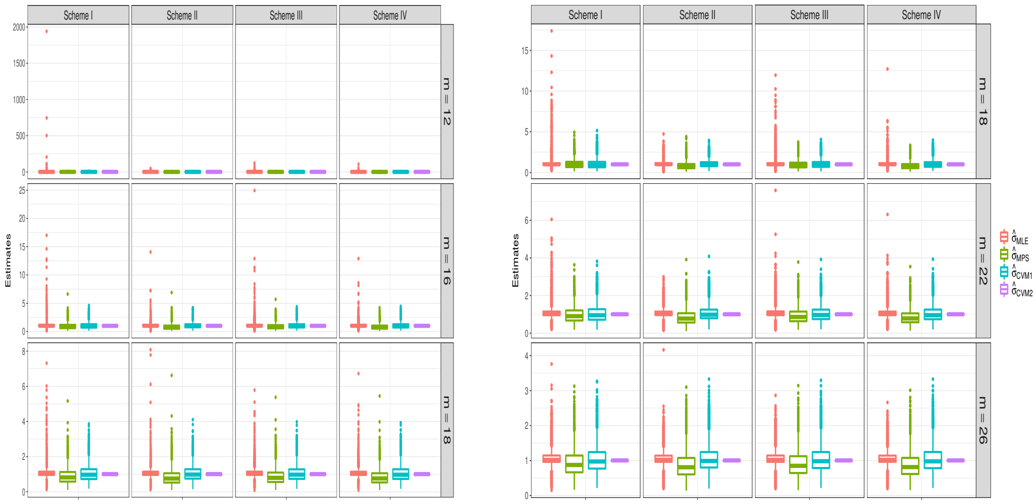
(b)

Figure 1: *The MSEs and biases of the estimators for the sample size (a) $n = 20$, (b) $n = 30$.*

sample and the expectation of the replicated data is high. Furthermore, to examine the uncertainty, the $100(1-\alpha)\%$ predictive region for $X_{i:m:n}^{\text{rep}}$ is computed, which is given by

$$\left(X_{i:m:n}^{\text{rep},([(\alpha/2)N])},\ X_{i:m:n}^{\text{rep},([(1-\alpha/2)N])}\right),$$

where $X_{i:m:n}^{\text{rep},([\alpha N])}$ denotes the $[\alpha N]^{th}$ smallest of $\{X_{i:m:n}^{\text{rep},k}\}$.



(a) (b)

Figure 2: *Boxplots for the estimates of* σ *for the sample size (a)* $n = 20$, *(b)* $n = 30$.

4.2. Pivotal-based

Let

$$Q_{i:m:n} = -\log\left[1 - H\left(x_{i:m:n};\theta\right)\right], \quad i = 1, \ldots, m. \tag{4.1}$$

Then, it is a progressive Type-II censored sample having a standard exponential distribution with a mean $E(Q_{i:m:n}) = \sum_{j=1}^{i} 1/\gamma_j$. Note that it is a pivotal quantity because its distribution does not depend on the unknown parameter θ. Then, by plotting the scatter plot between the pivotal quantity $q_{i:m:n}$ and its mean $E(Q_{i:m:n})$, the goodness-of-fit test can be conducted. The corresponding algorithm can be summarized as follows:

(a) Estimate the unknown parameter θ.

(b) Compute $q_{i:m:n}$ for $i = 1, \ldots, m$ from (4.1).

(c) Plot scatter plot for m pairs
$\{(q_{1:m:n}, E(Q_{1:m:n})), (q_{2:m:n}, E(Q_{2:m:n})), \ldots, (q_{m:m:n}, E(Q_{m:m:n}))\}$.

5. Application

In this section, to examine the performance of the estimation methods discussed in Sections 2 and 3, the Monte Carlo simulations and the real data analysis are conducted.

5.1. Simulation result

As mentioned earlier, the maximum likelihood estimation method can break down for $\lambda < 1$ in the three-parameter Weibull distribution. To resolve this problem, the minimum distance estimation

Table 1: Breaking strength of jute fiber at gauge length 20 mm

71.46	419.02	284.64	585.57	456.60	113.85	187.85	688.16	662.66	45.58	578.62	756.70	594.29	166.49	99.72
707.36	765.14	187.13	145.96	350.70	547.44	116.99	375.81	581.60	119.86	48.01	200.16	36.75	244.53	83.55

Table 2: Progressive Type-II censored sample generated from the breaking strength dataset

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$x_{i:15:30}$	0.3675	0.4558	0.4801	0.7146	0.8355	0.9972	1.1385	1.1699	1.1986	1.6649	1.8713	1.8785	2.0016	3.5070	4.1902
R_i	0	0	0	0	0	2	2	2	2	2	0	0	0	0	5

method is developed, and its excellence can be demonstrated through the Monte Carlo simulation. For the simulation, the unknown parameters μ, σ, and λ of the three-parameter Weibull distribution are assigned 0, 1, and 0.5, respectively. Then, to generate the progressive Type-II censored sample, the following censoring schemes are employed:

$$\begin{aligned} \text{Scheme} \quad \text{I} &: \mathscr{R} = \left(0^{*m-1},\ n-m\right), \\ \text{II} &: \mathscr{R} = \left(n-m,\ 0^{*m-1}\right), \\ \text{III} &: \mathscr{R} = \left(\frac{n-m}{2},\ 0^{*m-2},\ \frac{n-m}{2}\right), \\ \text{IV} &: \mathscr{R} = \left(0^{*\frac{m}{2}-1},\ n-m,\ 0^{*\frac{m}{2}}\right). \end{aligned}$$

Here, 0^{*m-1} denotes a vector of zeros of the size $m-1$. The mean squared errors (MSEs) and biases are obtained from 5,000 replications for each censoring scheme.

The results are reported in Figure 1, and it is summarized as follows: For μ, the MSEs and biases of all estimators have values close to zero, which indicates that they have very satisfactory performance. For λ, $\hat{\lambda}_{\text{CVM2}}$ has generally more efficient bias than the MLE counterpart in the censoring schemes II and III, while $\hat{\lambda}_{\text{MLE}}$ has generally more efficient MSE than other estimators for λ. For σ, $\hat{\sigma}_{\text{CVM1}}$ has generally more efficient bias than the MLE and MPS counterparts, while $\hat{\sigma}_{\text{MLE}}$ often yields absurd MSE values for $m = 12$ compared to other estimators for σ. In addition, $\hat{\sigma}_{\text{CVM2}}$ is expected to have stable performance as the MSE and bias values are very close to zero. To verify this, the boxplots for all estimates of σ used to calculate the MSE and bias in Figure 1 are presented in Figure 2. As expected, Figure 2 shows that $\hat{\sigma}_{\text{CVM2}}$ is very stable with a small deviation, while $\hat{\sigma}_{\text{MLE}}$ is often much larger than the true value 1 with a big deviation.

In conclusion, the maximum likelihood estimation method yields absurd MSE values for σ, while the MLE of λ has generally more efficient MSE than other estimators. However, the maximum likelihood estimates of μ, σ, and λ are simultaneously obtained using a numerical optimization method, so they affect each other. Based on this fact, it is questionable whether the MSE values for the MLE of λ are reliable.

5.2. Real data

Jute fiber is an environmentally friendly natural fiber that can be used as a good alternative in the reinforcement of composite materials, but it exhibits brittle. So, strength reliability is one of the critical factors to restrict the broader use of jute fiber, and the strength dataset is often characterized by the Weibull distribution.

This subsection considers the dataset on the breaking strength of jute fiber at gauge length 20 mm reported by Xia *et al*. (2009) to illustrate the real dataset application of the proposed approach. For the analysis, the dataset is divided by 100, then the progressive Type-II censored samples are generated

Table 3: Estimates of μ, σ, and λ for the breaking strength dataset

$\hat{\mu}_{MLE}$	$\hat{\mu}_{MPS}$	$\hat{\mu}_{CVM1}$	$\hat{\mu}_{CVM2}$	$\hat{\sigma}_{MLE}$	$\hat{\sigma}_{MPS}$	$\hat{\sigma}_{CVM1}$	$\hat{\sigma}_{CVM2}$	$\hat{\lambda}_{MLE}$	$\hat{\lambda}_{MPS}$	$\hat{\lambda}_{CVM1}$	$\hat{\lambda}_{CVM2}$
0.367	0.305	0.367	0.365	6.528	3.168	3.714	3.605	0.336	0.860	0.664	0.791

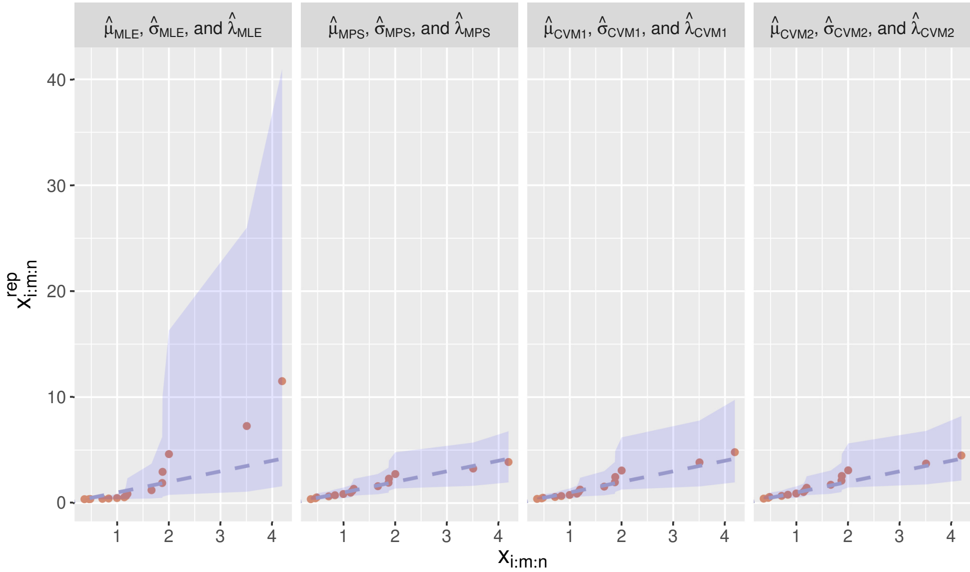


Figure 3: *95% predictive regions for* $X^{rep}_{i:m:n}$ *and the scatter plots between the progressive Type-II censored sample* $x_{i:m:n}$ *and the expectation* $E(X^{rep}_{i:m:n})$.

by setting $m = 15$ with the censoring scheme $\mathscr{R} = (0^{*5}, 2^{*5}, 0^{*4}, 5)$. The breaking strength dataset and the progressive Type-II sample are reported in Tables 1 and 2, respectively. Furthermore, the analysis results are reported in Table 3.

To examine that the progressive Type-II censored sample reported in Table 2 has the three-parameter Weibull distribution, two types of goodness-of-fit tests described in Section 4 are conducted. In the case of the replicated data, the expectation $E(X^{\text{rep}}_{i:m:n})$ and 95% predictive region are obtained based on the 20,000 replicated data. These results are reported in Figure 3, which reveals that the progressive Type-II censored sample reported in Table 2 has the three-parameter Weibull distribution well for all estimates because the points are close to a straight line. In addition, the MLEs have the worst performance in terms of the uncertainty due to a big predictive region.

Figure 4 plots the relationship between the pivotal quantity $q_{i:m:n}$ and its mean $E(Q_{i:m:n})$ for $i = 1, \ldots, m$, which also reveals that the progressive Type-II censored sample reported in Table 2 has the three-parameter Weibull distribution well for all estimates as in the case of Figure 3. Moreover, it shows that the MLEs have the worst fit because the points are slightly away from the straight line, compared to the MPSEs and MDEs.

6. Conclusion

In this study, a minimum distance estimation method was proposed to yield stable estimation results under the progressive Type-II censoring scheme. This approach was achieved by deriving a

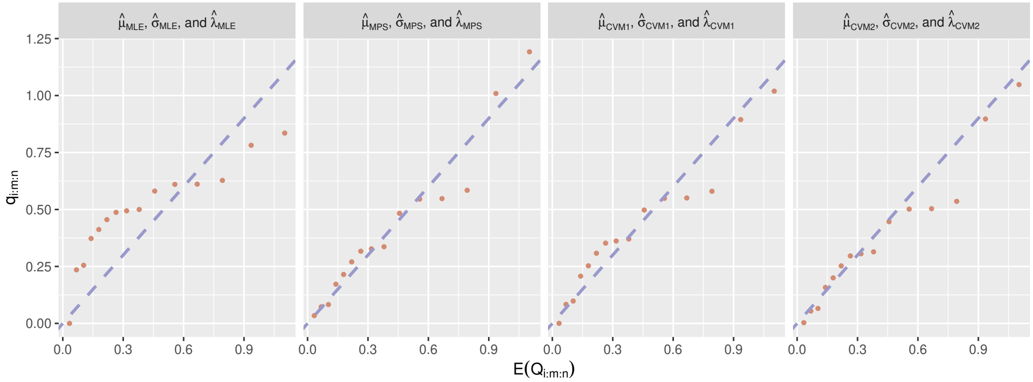


Figure 4: *The scatter plots between the pivotal quantity* $q_{i:m:n}$ *and its mean* $E(Q_{i:m:n})$.

new estimation equation based on the CVM distance. As an example of practical applications, the three-parameter Weibull distribution for the progressive Type-II censored sample was assumed, and a method for estimating unknown parameters was provided based on this.

The robustness and excellence of the proposed method were demonstrated by comparison with the maximum likelihood and MPS estimation methods through Monte Carlo simulations and analysis of the breaking strength of jute fiber, especially it showed very stable estimation performance for the scale parameter, compared to the maximum likelihood estimation method. Furthermore, two types of goodness-of-fit tests provided were shown that the breaking strength of jute fiber considered is very well fitted to the three-parameter Weibull distribution under the progressive Type-II censoring scheme, which reveals the usefulness and applicability of the proposed method.

In conclusion, it can be used as an alternative in various situations where the maximum likelihood estimation method may fail or the nuisance parameter is an obstacle to estimating the parameter of interest. Accordingly, the proposed minimum distance estimation method is expected to be applicable to more general censoring schemes such as hybrid or adaptive censoring schemes in addition to the progressive Type-II censoring scheme, and additional research on its scalability is being conducted.

References

Anatolyev S and Kosenok G (2005). An alternative to maximum likelihood based on spacings, *Econometric Theory*, **21**, 472–476.

Basu S, Singh SK, and Singh U (2017). Parameter estimation of inverse Lindley distribution for Type-I censored data, *Computational Statistics*, **32**, 367–385.

Basu S, Singh SK, and Singh U (2018). Bayesian inference using product of spacings function for progressive hybrid Type-I censoring scheme, *Statistics*, **52**, 345–363.

Cousineau D (2009). Fitting the three-parameter Weibull distribution: Review and evaluation of existing and new methods, *IEEE Transactions on Dielectrics and Electrical Insulation*, **16**, 281–288.

Cheng RCH and Amin NAK (1983). Estimating parameters in continuous univariate distributions with a shifted origin, *Journal of the Royal Statistical Society: Series B (Methodological)*, **45**, 394–403.

Ekström M (1998). On the consistency of the maximum spacing method, *Journal of Statistical Planning and Inference*, **70**, 209–224.

Gallagher MA and Moore AH (1990). Robust minimum-distance estimation using the 3-parameter

Weibull distribution, *IEEE Transactions on Reliability*, **39**, 575–580.

Hooke R and Jeeves TA (1961). "Direct search" solution of numerical and statistical problems, *Journal of the Association for Computing Machinery*, **8**, 212–229.

Hobbs JR, Moore AH, and James W (1984). Minimum distance estimation of the three parameters of the gamma distribution, *IEEE Transactions on Reliability*, **33**, 237–240.

Hobbs JR, Moore AH, and Miller RM (1985). Minimum-distance estimation of the parameters of the 3-parameter Weibull distribution, *IEEE Transactions on Reliability*, **34**, 495–496.

Jeon YE, Kang SB, and Seo JI (2022). Maximum product of spacings under a generalized Type-II progressive hybrid censoring scheme, *Communications for Statistical Applications and Methods*, **29**, 665–677.

Pitman EJG (1979). *Some Basic Theory for Statistical Inference*, Chapman and Hall, London.

Pakyari R and Balakrishnan N (2012). A general purpose approximate goodness-of-fit test for progressively Type-II censored data, *IEEE Transactions on Reliability*, **61**, 238–244.

Ranneby B (1984). The maximum spacing method: An estimation method related to the maximum likelihood method, *Scandinavian Journal of Statistics*, **11**, 93–112.

Shao Y and Hahn MG (1999). Maximum product of spacings method: A unified formulation with illustration of strong consistency, *Illinois Journal of Mathematics*, **43**, 489–499.

Seo JI, Jeon YE, and Kang SB (2020). New approach for a Weibull distribution under the progressive Type-II censoring scheme, *Mathematics*, **8**, 1713.

Singh RK, Kaushik A, Singh SK, and Singh U (2018). Product spacings for the estimation of the parameters of the exponentiated Pareto distribution, *International Journal of Applied Mathematics & Statistics*, **57**, 79–95.

Singh U, Singh SK, and Singh RK (2014). A comparative study of traditional estimation methods and maximum product spacings method in generalized inverted exponential distribution, *Journal of Statistics Applications & Probability*, **3**, 153–169.

Xia ZP, Yu JY, Cheng LD, Liu LF, and Wang WM (2009). Study on the breaking strength of jute fibres using modified Weibull distribution, *Composites Part A: Applied Science and Manufacturing*, **40**, 54–59.

Received January 21, 2023; Revised March 4, 2023; Accepted March 4, 2023