

Article

Change Point Test for Length-Biased Lognormal Distribution under Random Right Censoring

Mei Li ¹, Wei Ning ² and Yubin Tian ^{3,*}¹ Faculty of Science, Kunming University of Science and Technology, Kunming 650500, China; biter_limei@163.com² Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403, USA; wning@bgsu.edu³ School of Mathematics and Statistics, Beijing Institute of Technology, Beijing 100081, China

* Correspondence: tianyb@bit.edu.cn

Abstract: The length-biased lognormal distribution is a length-biased version of lognormal distribution, which is developed to model the length-biased lifetime data from, for example, biological investigation, medical research, and engineering fields. Owing to the existence of censoring phenomena in lifetime data, we study the change-point-testing problem of length-biased lognormal distribution under random censoring in this paper. A procedure based on the modified information criterion is developed to detect changes in parameters of this distribution. Under the sufficient condition of the Fisher information matrix being positive definite, it is proven that the null asymptotic distribution of the test statistic follows a chi-square distribution. In order to evaluate the uncertainty of change point location estimation, a way of calculating the coverage probabilities and average lengths of confidence sets of change point location based on the profile likelihood and deviation function is proposed. The simulations are conducted, under the scenarios of uniform censoring and exponential censoring, to investigate the validity of the proposed method. And the results indicate that the proposed approach performs better in terms of test power, coverage probabilities, and average lengths of confidence sets compared to the method based on the likelihood ratio test. Subsequently, the proposed approach is applied to the analysis of survival data from heart transplant patients, and the results show that there are differences in the median survival time post-heart transplantation among patients of different ages.



Citation: Li, M.; Ning, W.; Tian, Y. Change Point Test for Length-Biased Lognormal Distribution under Random Right Censoring. *Mathematics* **2024**, *12*, 1760. <https://doi.org/10.3390/math12111760>

Academic Editors: Hongyan Dui, Keke Wang and Qianqian Zhao

Received: 22 April 2024

Revised: 21 May 2024

Accepted: 28 May 2024

Published: 5 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: change point test; length-biased lognormal distribution; modified information criterion; random right censoring

MSC: 62N03

1. Introduction

In fields such as medical research [1], survival analysis [2,3], and reliability studies [4], censored data often exist when the life model is a length-biased distribution. Typically, the type of censoring can be classified into left censoring, interval censoring, and right censoring, where the right censoring can further be categorized as Type I censoring, Type II censoring and random censoring, and random censoring is the most common type of right censoring [5,6]. It refers to situations where the total period of observation is fixed, but subjects enter into the study at different points in time. Some subjects experience the events of interest. Others do not, and some are lost to follow-ups. Others will still be alive at the end of the study period. In random censoring, the censored subjects do not all have the same censoring time [5]. In this case, if the censored data are deleted or the censored time is modeled as complete lifetime data, it may lead to the underestimation of important numerical characteristics such as the median or mean. Therefore, analyzing

right-censored data in a reasonable manner is of significant importance for exploring the underlying patterns. In this article, we focus on random right-censored data.

The length-biased lognormal distribution is a kind of length-biased distribution based on lognormal distribution. Its probability density function is defined as follows:

$$g(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}[\log x - (\mu + \sigma^2)]^2\right\}, \quad (1)$$

where $x > 0$, $-\infty < \mu < \infty$, $\sigma > 0$. For convenience, it is briefly denoted as $LBLN(\mu, \sigma)$. Because the length-biased lognormal distribution is important in describing the characteristics of lifetime distributions, it has been studied extensively. Sangsri and Akman considered it for product life modeling and deduced its corresponding reliability function [7,8]. Ratnaparkhi and Naik-Nimbalkar studied the estimation problems of length-biased lognormal distribution [9]. The existing literature primarily focuses on the statistical properties [7], parameter estimation [9], and practical applications of the length-biased lognormal distribution [8]. However, there is relatively limited research on the parameter change point test of this distribution under random censoring.

The problem of change point originates from the field of industrial quality control [10] and has become one of the hot topics in statistical research since it was proposed. The research contents mainly include two aspects. One is to test whether there is a change point, and the other is to detect the number and locations of change points if there are existing change points. Following over six decades of research and development, the study of the change point problem has been widely applied in epidemiology, biology, environmental science, and reliability engineering [11–13]. For censored data, there are a few studies that have focused on change point testing and estimation for linear transformation models and hazard function parameters based on right-censored data. For instance, Kosorok and Song [14] utilized a score test statistic to investigate the estimation and testing problems of change points for regression coefficients in a linear transformation model for right-censored survival data. Rabhi and Asgharian [15] studied the problem of change point estimation for the hazard function under biased sampling and right-censoring data. Chen et al. [16] proposed a detection procedure based on the empirical likelihood for the changes in the mean residual life functions with right-censored data.

While some of the literature has studied change point detection procedures for linear transformation models and interesting reliability parameters with right-censored data, research on change point testing for length-biased lognormal distribution with random right-censored data has been limited. In this change point detection problem, we can approach it as a model selection problem, where we aim to select the better option between the null hypothesis with no change and the alternative hypothesis with at least one change. Therefore, methods commonly used for model selection can be applied to change-point-testing problems.

Compared to the usual model selection problem, the change point problem introduces a special parameter—the change location. When the change point occurs near the middle of the process, there are no redundant parameters, making it easier to detect the change point. However, when it occurs near the beginning or the end of the sequence, the parameters of one segment of the data (the first part when it is near the beginning, the second part when it is near the end) become redundant. The traditional change point detection method is based on the likelihood ratio (LRT), which does not consider the contribution of the change location.

To address this limitation, we utilized the modified information criterion method, which considers the influence of change point positions on model complexity in the penalty term. This method was employed to detect the change point in the length-biased lognormal distribution with random right-censored data. Additionally, the asymptotic properties of the test statistics constructed based on the MIC with random right censoring have not been studied yet. This paper aims to investigate change point detection for the length-biased lognormal distribution under random right censoring.

The rest of this paper is organized as follows. In Section 2, a change-point-testing model for the parameters of the length-biased lognormal distribution is presented, which is based on random right censoring. A corresponding change point test method is introduced, and sufficient conditions for the positive definiteness of the Fisher information matrix under random right-censored data are provided, along with the asymptotic distribution of the test statistic. Then, a technique to calculate the coverage probabilities and average lengths of confidence sets of change point location based on profile likelihood function and deviation function is proposed. Simulations are carried out, with censoring rates of 10% and 30% under uniform and exponential censoring distributions, to indicate the performance of the detecting procedures in Section 3. In Section 4, we illustrate our method on a survival time dataset in heart transplant patients. Some discussion is provided in Section 5.

2. Methodology

2.1. Change Point Model

Suppose $T_1, T_2, \dots, T_n$ are independent and identically distributed (i.i.d.) survival times. A common feature of survival data is the presence of right censoring, which is the most common type of censoring. In the presence of censoring, we only observe $(U_1, \delta_1), (U_2, \delta_2), \dots, (U_n, \delta_n)$, where (U_i, δ_i) is obtained as follows,

$$(U_i, \delta_i) = \min\{(T_i, C_i), I(T_i \leq C_i)\}, i = 1, 2, \dots, n \quad (2)$$

where C_i are potential censoring times for n subjects, and they are often treated as random variables with distribution function $H(\eta)$ in statistical inference. And we assume that $\{T_1, T_2, \dots, T_n\}$ is independent of $\{C_1, C_2, \dots, C_n\}$.

If the independent random variables $T_1, T_2, \dots, T_n$ follow a length-biased lognormal distribution, that is $T_i \sim LBLN(\mu_i, \sigma_i)$, then the density of the length-biased lognormal distribution under random right-censored data pair (U_i, δ_i) is

$$p(u_i) = g(u_i)^{\delta_i} S_g(u_i)^{1-\delta_i}, \quad (3)$$

where

$$g(u_i) = \frac{1}{u_i \sigma \sqrt{2\pi}} \exp\left\{-\frac{[\log u_i - (\mu + \sigma^2)]^2}{2\sigma^2}\right\}, \quad (4)$$

$$S_g(u_i) = 1 - \Phi\left(\frac{\log u_i - (\mu + \sigma^2)}{\sigma}\right). \quad (5)$$

We are interested in testing the null hypothesis of no change in $\theta_i = (\mu_i, \sigma_i)$ of length-biased lognormal distribution with random right censoring

$$H_0 : \theta_1 = \theta_2 = \dots = \theta_n = \theta \quad (6)$$

against the following alternative hypothesis

$$H_1 : \theta_1 = \theta_2 = \dots = \theta_{k_1} \neq \theta_{k_1+1} = \dots = \theta_{k_2} \neq \theta_{k_2+1} \dots \neq \theta_{k_p+1} = \dots = \theta_n, \quad (7)$$

where $k_1, k_2, \dots, k_p$ are the locations of p changes in the sequence. Since the multiple change point problem can be simplified to a single change point problem by the binary segmentation method [17], we thus consider testing the null hypothesis (6) against the following alternative hypothesis:

$$H_1 : \theta_L = \theta_1 = \theta_2 = \dots = \theta_k \neq \theta_{k+1} = \dots = \theta_n = \theta_R, \quad (8)$$

where k is the unknown change point location, and θ_L and θ_R are the parameters before and after the change point, respectively.

Then, the likelihood function under the null hypothesis is

$$\begin{aligned} L_0(\theta) &= \prod_{i=1}^n g(u_i)^{\delta_i} S_g(u_i)^{1-\delta_i} \cdot \prod_{i=1}^n h(u_i)^{1-\delta_i} [1 - H(u_i)]^{\delta_i} \\ &= \prod_{i=1}^n \left\{ \frac{1}{u_i \sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma^2} [\log u_i - (\mu + \sigma^2)]^2 \right\} \right\}^{\delta_i} \\ &\quad \prod_{i=1}^n \left\{ 1 - \Phi \left[\frac{\log u_i - (\mu + \sigma^2)}{\sigma} \right] \right\}^{1-\delta_i} \prod_{i=1}^n h(u_i)^{1-\delta_i} [1 - H(u_i)]^{\delta_i}, \end{aligned} \quad (9)$$

where $h(\cdot)$ is the probability density function of the censored distribution, and $H(\cdot)$ is the probability distribution function of the censored distribution. The log-likelihood function after removing the constant term is

$$\begin{aligned} l_0(\theta) &= -\log \sigma \sum_{i=1}^n \delta_i - \frac{1}{2\sigma^2} \sum_{i=1}^n \delta_i \{ [\log u_i - (\mu + \sigma^2)]^2 \} \\ &\quad + \sum_{i=1}^n (1 - \delta_i) \log \left\{ 1 - \Phi \left[\frac{\log u_i - (\mu + \sigma^2)}{\sigma} \right] \right\}. \end{aligned} \quad (10)$$

Similarly, the log-likelihood function under the alternative hypothesis is

$$\begin{aligned} l_1(\theta_L, \theta_R) &= -\log \sigma_L \sum_{i=1}^k \delta_i - \frac{1}{2\sigma_L^2} \sum_{i=1}^k \delta_i \{ [\log u_i - (\mu_L + \sigma_L^2)]^2 \} \\ &\quad + \sum_{i=1}^k (1 - \delta_i) \log \left\{ 1 - \Phi \left[\frac{\log u_i - (\mu_L + \sigma_L^2)}{\sigma_L} \right] \right\} \\ &\quad - \log \sigma_R \sum_{i=k+1}^n \delta_i - \frac{1}{2\sigma_R^2} \sum_{i=k+1}^n \delta_i \{ [\log u_i - (\mu_R + \sigma_R^2)]^2 \} \\ &\quad + \sum_{i=k+1}^n (1 - \delta_i) \log \left\{ 1 - \Phi \left[\frac{\log u_i - (\mu_R + \sigma_R^2)}{\sigma_R} \right] \right\}. \end{aligned} \quad (11)$$

Because of the existence of nonlinear functions $\Phi(\cdot)$, we obtain $\hat{\theta}$, $\hat{\theta}_L$ and $\hat{\theta}_R$ by the numerical method.

2.2. Modified Information Criterion Procedure

The likelihood ratio test (LRT) procedure is one of the most popular methods for parametric change point analysis. Alternatively, a change point problem can be viewed as a model selection problem. That is, we choose a better model between the null hypothesis and the alternative hypothesis. The choice of a model under the null hypothesis corresponds to the situation of no change. Otherwise, it corresponds to the case that at least one change occurs. Hence, the information criterion, such as the Akaike information criterion (AIC) [18] and Schwarz information criterion (SIC) [19], can be used for the change point test. The literature is fruitful with studies which have been conducted in this direction. For instance, Chen and Gupta [20] developed a binary procedure combined with the SIC to search all of the possible variance change points in a sequence of independent Gaussian random variables. Chen and Gupta [21] provided the testing and estimation of a single change point in means and variances of a sequence of independent Gaussian normal random variables based on SIC and the unbiased version of the SIC and so on.

Model complexity is a very important factor in model selection based on the information criteria. The dimension of parameter space is usually used to measure model complexity. However, the method based on SIC does not consider the influence of the change location, which may cause redundancy when the change nears the beginning or the end of data. To solve this issue, Chen et al. [22] modified the traditional information

criterion by making the model complexity a function of the change point location, and denoted it as a modified information criterion (MIC). In this paper, we study the change point problem in parameters of the LBLN distribution with random right censoring based on MIC. Under the null hypothesis, the MIC is defined as

$$\text{MIC}(n) = -2 \log L_0(\hat{\theta}) + 2 \log(n), \quad (12)$$

where $\hat{\theta} = (\hat{\mu}, \hat{\sigma})$ is the MLE of $\theta = (\mu, \sigma)$ and n is the sample size. The associated MIC statistics under the alternative hypothesis is defined as

$$\text{MIC}(k) = -2 \log L_1(\hat{\theta}_L, \hat{\theta}_R) + \left\{ 4 + \left(\frac{2k}{n} - 1 \right)^2 \right\} \log(n), \quad (13)$$

where k is the unknown change point location, $\hat{\theta}_L$ and $\hat{\theta}_R$ are maximum likelihood estimators of θ_L and θ_R . For the penalty term $\{4 + (2k/n - 1)^2\} \log(n)$ in Equation (13), when $k \rightarrow 1$ or $k \rightarrow n$, it means that the change point positions appear at both ends, respectively, and the penalty term in $\text{MIC}(k)$ approaches $5 \log(n)$. When $k \rightarrow n/2$, the penalty term approaches $4 \log(n)$, and the complexity of the model is minimized. This is because when the change point position is close to both ends, there are not enough data for estimating the parameters, which may result in a large variance of parameter estimation. Thus, we should set a larger penalty in this case. When the change point is close to the middle of the data, there are enough data to estimate the parameters, and the variance of the parameter estimation is small, so setting a smaller penalty is sufficient. That is, when the suspected change points appear at both ends, stronger evidence is needed to prove the existence of such changes. Therefore, when k closes to 1 or n , a larger penalty needs to be set.

In order to ensure sufficient observations for parameter estimation, the range of k is $K = \{k | k_0 \leq k \leq n - k_0\}$, where $k_0 = 2[\log n]$. Based on the minimum information criterion, a model with the smallest MIC value will be considered the best one to fit the data. Then, the $\min_{k \in K} \text{MIC}(k)$ corresponds to the best model under H_1 . Further, we accept H_0 if

$$\text{MIC}(n) \leq \min_{k \in K} \text{MIC}(k),$$

that is, the model with no change is the best one. We reject H_0 if

$$\text{MIC}(n) > \min_{k \in K} \text{MIC}(k),$$

which indicates that the best model under H_1 is more appropriate to describe the data than the model under H_0 . It leads us to conclude that there exists at least one change in the data. Correspondingly, the location of change can be estimated as follows:

$$\hat{k} = \arg \min_{k \in K} \text{MIC}(k). \quad (14)$$

In order to make the conclusion more statistically convincing, we construct the test statistics-based MIC as follows:

$$M_n = \text{MIC}(n) - \min_{k \in K} \text{MIC}(k) + 2 \log(n). \quad (15)$$

Given critical value M under a certain significance level α , one can make the decision of accepting or rejecting H_0 by comparing M and M_n . To calculate the critical value of M_n , under certain conditions, the asymptotic distribution of the M_n can be obtained as shown in Theorem 1, and the following lemmas are required to prove this theorem.

Lemma 1. Under the random right-censoring data (U_i, δ_i) , if $(\delta_i A_{11} + (1 - \delta_i) B_{11})(\delta_i A_{22} + (1 - \delta_i) B_{22}) \neq (\delta_i A_{12} + (1 - \delta_i) B_{12})^2$, then the Fisher information matrix based on these data is positively definite, where

$$A(\theta) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, A_{ij} = E \left(-\frac{\partial^2 \log g(U_i)}{\partial \theta \partial \theta^T} \right), (i = j = 1, 2)$$

$$B(\theta) = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, B_{ij} = E \left(-\frac{\partial^2 \log S_g(U_i)}{\partial \theta \partial \theta^T} \right), (i = j = 1, 2).$$

Proof of Lemma 1. According to the definition of random right-censoring data pairs in Section 2.1, the density function under right-censoring data is $p(U_i; \theta)$, where

$$p(U_i; \theta) = g(U_i; \theta)^{\delta_i} S_g(U_i; \theta)^{(1-\delta_i)},$$

$$g(U_i; \theta) = \frac{1}{U_i \sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma^2} [\log U_i - (\mu + \sigma^2)]^2 \right\},$$

$$S_g(U_i; \theta) = 1 - \Phi \left(\frac{\log U_i - (\mu + \sigma^2)}{\sigma} \right).$$

Then, the logarithmic density function is

$$\log p(U_i; \theta) = \delta_i \log g(U_i; \theta) + (1 - \delta_i) \log S_g(U_i; \theta).$$

For convenience in writing, let $Y_i = [\log U_i - (\mu + \sigma^2)]/\sigma$. Therefore, under the random right-censoring data, the Fisher information matrix of the density function $p(U_i; \theta)$ based on the length-biased lognormal distribution is as follows:

$$I(\theta) = -E \begin{pmatrix} \frac{\partial^2 \log p(U_i; \theta)}{\partial \mu^2} & \frac{\partial^2 \log p(U_i; \theta)}{\partial \mu \partial \sigma} \\ \frac{\partial^2 \log p(U_i; \theta)}{\partial \sigma \partial \mu} & \frac{\partial^2 \log p(U_i; \theta)}{\partial \sigma^2} \end{pmatrix} = E \begin{pmatrix} -\frac{\partial^2 \log p(U_i; \theta)}{\partial \mu^2} & -\frac{\partial^2 \log p(U_i; \theta)}{\partial \mu \partial \sigma} \\ -\frac{\partial^2 \log p(U_i; \theta)}{\partial \sigma \partial \mu} & -\frac{\partial^2 \log p(U_i; \theta)}{\partial \sigma^2} \end{pmatrix},$$

where

$$-\frac{\partial^2 \log p(U_i; \theta)}{\partial \mu^2} = \frac{\delta_i}{\sigma^2} - \frac{1 - \delta_i}{\sigma^2} \left\{ \frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)} - \frac{\phi^2(Y_i)}{\{1 - \Phi(Y_i)\}^2} \right\},$$

$$-\frac{\partial^2 \log p(U_i; \theta)}{\partial \mu \partial \sigma} = \delta_i \left\{ \frac{2}{\sigma} + \frac{2}{\sigma^2} Y_i \right\} - (1 - \delta_i) \left\{ -\frac{1}{\sigma^2} \frac{\phi(Y_i)}{1 - \Phi(Y_i)} + \frac{1}{\sigma} \frac{Y_i \phi(Y_i) \left(2 + \frac{1}{\sigma} Y_i \right)}{1 - \Phi(Y_i)} \right.$$

$$\left. - \frac{1}{\sigma} \frac{\phi^2(Y_i) \left(2 + \frac{1}{\sigma} Y_i \right)}{\{1 - \Phi(Y_i)\}^2} \right\},$$

$$-\frac{\partial^2 \log p(U_i; \theta)}{\partial \sigma^2} = \delta_i \left\{ -\frac{1}{\sigma^2} + \frac{6}{\sigma} Y_i + \frac{3}{\sigma^2} Y_i^2 + 4 \right\} - (1 - \delta_i) \left\{ \frac{Y_i \phi(Y_i) \left(2 + \frac{1}{\sigma} Y_i \right)^2}{1 - \Phi(Y_i)} \right.$$

$$\left. - \frac{1}{\sigma^2} \frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)} - \frac{1}{\sigma} \frac{\phi(Y_i) \left(2 + \frac{1}{\sigma} Y_i \right)}{1 - \Phi(Y_i)} - \frac{\phi^2(Y_i) \left(2 + \frac{1}{\sigma} Y_i \right)^2}{\{1 - \Phi(Y_i)\}^2} \right\},$$

$$-\frac{\partial^2 \log p(U_i; \theta)}{\partial \sigma \partial \mu} = -\frac{\partial^2 \log p(U_i; \theta)}{\partial \mu \partial \sigma}.$$

Further, the expectations for the second derivative are

$$\begin{aligned}
E\left(-\frac{\partial^2 \log p(U_i; \boldsymbol{\theta})}{\partial \mu^2}\right) &= \frac{\delta_i}{\sigma^2} - \frac{1 - \delta_i}{\sigma^2} \left\{ E\left(\frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)}\right) - E\left(\frac{\phi^2(Y_i)}{\{1 - \Phi(Y_i)\}^2}\right) \right\}, \\
E\left(-\frac{\partial^2 \log p(U_i; \boldsymbol{\theta})}{\partial \mu \partial \sigma}\right) &= \delta_i E\left(\frac{2}{\sigma} + \frac{2}{\sigma^2} Y_i\right) - (1 - \delta_i) \left\{ -\frac{1}{\sigma^2} E\left(\frac{\phi(Y_i)}{1 - \Phi(Y_i)}\right) \right. \\
&\quad \left. + \frac{1}{\sigma} E\left(\frac{Y_i \phi(Y_i) \left(2 + \frac{1}{\sigma} Y_i\right)}{1 - \Phi(Y_i)}\right) - \frac{1}{\sigma} E\left(\frac{\phi^2(Y_i) \left(2 + \frac{1}{\sigma} Y_i\right)}{\{1 - \Phi(Y_i)\}^2}\right) \right\}, \\
E\left(-\frac{\partial^2 \log p(U_i; \boldsymbol{\theta})}{\partial \sigma^2}\right) &= \delta_i E\left(-\frac{1}{\sigma^2} + \frac{6}{\sigma} Y_i + \frac{3}{\sigma^2} Y_i^2 + 4\right) - (1 - \delta_i) \left\{ E\left(\frac{Y_i \phi(Y_i) \left(2 + \frac{1}{\sigma} Y_i\right)^2}{1 - \Phi(Y_i)}\right) \right. \\
&\quad \left. - \frac{1}{\sigma^2} E\left(\frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)}\right) - \frac{1}{\sigma} E\left(\frac{\phi(Y_i) \left(2 + \frac{1}{\sigma} Y_i\right)}{1 - \Phi(Y_i)}\right) \right. \\
&\quad \left. - E\left(\frac{\phi^2(Y_i) \left(2 + \frac{1}{\sigma} Y_i\right)^2}{\{1 - \Phi(Y_i)\}^2}\right) \right\}.
\end{aligned}$$

When calculating the above expectations, it is found that there is no analytical expression for some integrals. We use the Monte Carlo integration method to approximate. In order to reduce the number of these integrals and Monte Carlo approximations, the following expectations can be simplified according to the integration by parts method,

$$\begin{aligned}
E\left(\frac{\phi^2(Y_i)}{\{1 - \Phi(Y_i)\}^2}\right) &= 2E\left(\frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)}\right), \\
E\left(\frac{Y_i \phi^2(Y_i)}{\{1 - \Phi(Y_i)\}^2}\right) &= -E\left(\frac{\phi(Y_i)}{1 - \Phi(Y_i)}\right) + 2E\left(\frac{Y_i^2 \phi(Y_i)}{1 - \Phi(Y_i)}\right), \\
E\left(\frac{Y_i^2 \phi^2(Y_i)}{\{1 - \Phi(Y_i)\}^2}\right) &= -2E\left(\frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)}\right) + 2E\left(\frac{Y_i^3 \phi(Y_i)}{1 - \Phi(Y_i)}\right).
\end{aligned}$$

Then, the elements of the Fisher information matrix are

$$\begin{aligned}
E\left(-\frac{\partial^2 \log p(U_i; \boldsymbol{\theta})}{\partial \mu^2}\right) &= \frac{\delta_i}{\sigma^2} + \frac{1 - \delta_i}{\sigma^2} E\left(\frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)}\right), \\
E\left(-\frac{\partial^2 \log p(U_i; \boldsymbol{\theta})}{\partial \mu \partial \sigma}\right) &= \frac{2\delta_i}{\sigma} + \frac{1 - \delta_i}{\sigma} \left\{ 2E\left(\frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)}\right) + \frac{1}{\sigma} E\left(\frac{Y_i^2 \phi(Y_i)}{1 - \Phi(Y_i)}\right) \right\}, \\
E\left(-\frac{\partial^2 \log p(U_i; \boldsymbol{\theta})}{\partial \sigma^2}\right) &= \delta_i \left(\frac{2}{\sigma^2} + 4\right) - (1 - \delta_i) \left\{ \frac{2}{\sigma} E\left(\frac{\phi(Y_i)}{1 - \Phi(Y_i)}\right) + 4E\left(\frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)}\right) \right. \\
&\quad \left. - \frac{4}{\sigma} E\left(\frac{Y_i^2 \phi(Y_i)}{1 - \Phi(Y_i)}\right) - \frac{1}{\sigma^2} E\left(\frac{Y_i^3 \phi(Y_i)}{1 - \Phi(Y_i)}\right) \right\}.
\end{aligned}$$

Thus,

$$\begin{aligned}
I(\boldsymbol{\theta}) &= \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix} = \delta_i A(\boldsymbol{\theta}) + (1 - \delta_i) B(\boldsymbol{\theta}), \\
&= \delta_i \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} + (1 - \delta_i) \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \\
&= \begin{pmatrix} \delta_i A_{11} + (1 - \delta_i) B_{11} & \delta_i A_{12} + (1 - \delta_i) B_{12} \\ \delta_i A_{21} + (1 - \delta_i) B_{21} & \delta_i A_{22} + (1 - \delta_i) B_{22} \end{pmatrix}
\end{aligned}$$

where

$$A(\theta) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma^2} & \frac{2}{\sigma} \\ \frac{2}{\sigma} & \frac{2}{\sigma^2} + 4 \end{pmatrix}, B(\theta) = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

$$B_{11} = \frac{1}{\sigma^2} E \left(\frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)} \right),$$

$$B_{12} = B_{21} = \frac{2}{\sigma} E \left(\frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)} \right) + \frac{1}{\sigma^2} E \left(\frac{Y_i^2 \phi(Y_i)}{1 - \Phi(Y_i)} \right),$$

$$B_{22} = -\frac{2}{\sigma} E \left(\frac{\phi(Y_i)}{1 - \Phi(Y_i)} \right) - 4 E \left(\frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)} \right) + \frac{4}{\sigma} E \left(\frac{Y_i^2 \phi(Y_i)}{1 - \Phi(Y_i)} \right) + \frac{1}{\sigma^2} E \left(\frac{Y_i^3 \phi(Y_i)}{1 - \Phi(Y_i)} \right).$$

Since $(\delta_i A_{11} + (1 - \delta_i) B_{11})(\delta_i A_{22} + (1 - \delta_i) B_{22}) \neq (\delta_i A_{12} + (1 - \delta_i) B_{12})^2$, and $(\delta_i A_{11} + (1 - \delta_i) B_{11})(\delta_i A_{22} + (1 - \delta_i) B_{22}) - (\delta_i A_{12} + (1 - \delta_i) B_{12})^2 = I_{11} I_{22} - I_{12} I_{21} = |I(\theta)| \neq 0$, then the Fisher information matrix $I(\theta)$ is positively definite. $\square$

Theorem 1. Based on certain Wald conditions and regular conditions, and if $(\delta_i A_{11} + (1 - \delta_i) B_{11})(\delta_i A_{22} + (1 - \delta_i) B_{22}) \neq (\delta_i A_{12} + (1 - \delta_i) B_{12})^2$, then the asymptotic null distribution of the test statistic M_n for the length-biased lognormal distribution with random right censoring is

$$M_n \xrightarrow{D} \chi^2(2), \quad \text{as } n \rightarrow \infty,$$

where M_n is defined in Equation (15), and Wald conditions and regular conditions are listed in Appendix A.

Proof of Theorem 1. Lemma 3 in [22] indicates that k/n approaches $1/2$ under H_0 , which can be expressed as $\Delta_\epsilon = \{k : |k/n - 1/2| < \epsilon\}$. This lemma was proved under the Wald conditions and regularity conditions. Let $\theta_L = (\mu_L, \sigma_L)$, and $\theta_R = (\mu_R, \sigma_R)$, $\theta = (\mu, \sigma)$ is in a small neighborhood of the true value $\theta_0 = (\mu_0, \sigma_0)$, that is, $\mathcal{M}_\delta = \{\theta : |\theta - \theta_0| < \delta\}$, where $\delta > 0$. So for any $\epsilon > 0, \delta > 0$, it can be obtained that

$$\begin{aligned} M_n &= \text{MIC}(n) - \min_{k \in \Delta_\epsilon} \text{MIC}(k) + 2 \log n \\ &= \max_{k \in \Delta_\epsilon} \{ \text{MIC}(n) - \text{MIC}(k) \} + 2 \log n \\ &= \max_{k \in \Delta_\epsilon} \left\{ 2 \log L_1(\theta_L, \theta_R) - 2 \log L_0(\theta) - \left(\frac{2k}{n} - 1 \right)^2 \log n \right\} \\ &\leq 2 \max_{k \in \Delta_\epsilon} \left\{ \sup_{\theta_L \in \mathcal{M}_\delta} \sum_{i=1}^k \log p(U_i; \theta_L) + \sup_{\theta_R \in \mathcal{M}_\delta} \sum_{i=k+1}^n \log p(U_i; \theta_R) \right. \\ &\quad \left. - \sup_{\theta \in \mathcal{M}_\delta} \sum_{i=1}^n \log p(U_i; \theta) \right\} + o_p(1). \end{aligned} \quad (16)$$

where

$$\begin{aligned} p(U_i; \theta) &= g(U_i; \theta)^{\delta_i} S_g(U_i; \theta)^{(1-\delta_i)}, \\ g(U_i; \theta) &= \frac{1}{U_i \sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma^2} [\log U_i - (\mu + \sigma^2)]^2 \right\}, \\ S_g(U_i; \theta) &= 1 - \Phi \left(\frac{\log U_i - (\mu + \sigma^2)}{\sigma} \right). \end{aligned}$$

Let $\tilde{\theta}_L$, $\tilde{\theta}_R$, and $\tilde{\theta}$ be the maximum points of $\sum_{i=1}^k \log p(U_i; \theta)$, $\sum_{i=k+1}^n \log p(U_i; \theta)$, and $\sum_{i=1}^n \log p(U_i; \theta)$, respectively, in the interval $(\theta_0 - \delta, \theta_0 + \delta)$. Then, there must exist an

η such that $|\eta - \theta_0| < \delta$ when θ equals to $\tilde{\theta}_L$, $\tilde{\theta}_R$, or $\tilde{\theta}$. Furthermore, the Taylor expansion of $\log p(U_i; \theta)$ at $\theta = \theta_0 = (\mu_0, \sigma_0)$ can be obtained as follows:

$$\begin{aligned} \log p(U_i; \theta) - \log p(U_i; \theta_0) &= S(U_i; \theta)|_{\theta=\theta_0}(\theta - \theta_0)^T \\ &\quad + \frac{1}{2}(\theta - \theta_0)H(\theta_0)(\theta - \theta_0)^T + \frac{1}{6}M\delta^3 + o_p(1), \end{aligned}$$

where $(\theta - \theta_0) = (\mu - \mu_0, \sigma - \sigma_0)$, $S(U_i; \theta) = (\partial \log p(U_i; \theta)/\partial \mu, \partial \log p(U_i; \theta)/\partial \sigma)$ is the score function of $g(U_i; \theta)$. For convenience in writing, let $Y_i = \log U_i - (\mu + \sigma^2)/\sigma$, then the elements in $S(U_i; \theta)$ are

$$\begin{aligned} \frac{\partial \log p(U_i; \theta)}{\partial \mu} &= \frac{\delta_i}{\sigma} Y_i + \frac{1 - \delta_i}{\sigma} \frac{\phi(Y_i)}{1 - \Phi(Y_i)}, \\ \frac{\partial \log p(U_i; \theta)}{\partial \sigma} &= \delta_i \left\{ -\frac{1}{\sigma} + 2Y_i + \frac{1}{\sigma} Y_i^2 \right\} + (1 - \delta_i) \left\{ \left(2 + \frac{1}{\sigma} Y_i \right) \frac{\phi(Y_i)}{1 - \Phi(Y_i)} \right\}. \end{aligned}$$

$H(\theta)$ is the second-order derivative matrix of $p(U_i; \theta)$ with respect to each parameter, that is,

$$H(\theta) = \begin{pmatrix} \frac{\partial^2 \log p(U_i; \theta)}{\partial \mu^2} & \frac{\partial^2 \log p(U_i; \theta)}{\partial \mu \partial \sigma} \\ \frac{\partial^2 \log p(U_i; \theta)}{\partial \sigma \partial \mu} & \frac{\partial^2 \log p(U_i; \theta)}{\partial \sigma^2} \end{pmatrix},$$

and elements of $H(\theta)$ are represented as follows:

$$\begin{aligned} \frac{\partial^2 \log p(U_i; \theta)}{\partial \mu^2} &= -\frac{\delta_i}{\sigma^2} + \frac{1 - \delta_i}{\sigma^2} \left\{ \frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)} - \frac{\phi^2(Y_i)}{\{1 - \Phi(Y_i)\}^2} \right\}, \\ \frac{\partial^2 \log p(U_i; \theta)}{\partial \mu \partial \sigma} &= \delta_i \left\{ -\frac{2}{\sigma} - \frac{2}{\sigma^2} Y_i \right\} + (1 - \delta_i) \left\{ -\frac{1}{\sigma^2} \frac{\phi(Y_i)}{1 - \Phi(Y_i)} + \frac{1}{\sigma} \frac{Y_i \phi(Y_i) \left(2 + \frac{1}{\sigma} Y_i \right)}{1 - \Phi(Y_i)} \right. \\ &\quad \left. - \frac{1}{\sigma} \frac{\phi^2(Y_i) \left(2 + \frac{1}{\sigma} Y_i \right)}{\{1 - \Phi(Y_i)\}^2} \right\}, \\ \frac{\partial^2 \log p(U_i; \theta)}{\partial \sigma^2} &= \delta_i \left\{ \frac{1}{\sigma^2} - \frac{6}{\sigma} Y_i - \frac{3}{\sigma^2} Y_i^2 - 4 \right\} + (1 - \delta_i) \left\{ \frac{Y_i \phi(Y_i) \left(2 + \frac{1}{\sigma} Y_i \right)^2}{1 - \Phi(Y_i)} \right. \\ &\quad \left. - \frac{1}{\sigma^2} \frac{Y_i \phi(Y_i)}{1 - \Phi(Y_i)} - \frac{1}{\sigma} \frac{\phi(Y_i) \left(2 + \frac{1}{\sigma} Y_i \right)}{1 - \Phi(Y_i)} - \frac{\phi^2(Y_i) \left(2 + \frac{1}{\sigma} Y_i \right)^2}{\{1 - \Phi(Y_i)\}^2} \right\}, \\ \frac{\partial^2 \log p(U_i; \theta)}{\partial \sigma \partial \mu} &= \frac{\partial^2 \log p(U_i; \theta)}{\partial \mu \partial \sigma}. \end{aligned}$$

Then, Equation (16) can be further simplified as

$$\begin{aligned} M_n &\leq \max_{k \in \Delta_c} \left\{ \sup_{\theta_L \in \mathcal{M}_\delta} \left[2 \sum_{i=1}^k S(U_i; \theta_0)(\theta_L - \theta_0)^T + \sum_{i=1}^k (\theta_L - \theta_0)H(\theta_0)(\theta_L - \theta_0)^T \right] \right. \\ &\quad + \sup_{\theta_R \in \mathcal{M}_\delta} \left[2 \sum_{i=k+1}^n S(U_i; \theta_0)(\theta_R - \theta_0)^T + \sum_{i=k+1}^n (\theta_R - \theta_0)H(\theta_0)(\theta_R - \theta_0)^T \right] \\ &\quad \left. - \sup_{\theta \in \mathcal{M}_\delta} \left[2 \sum_{i=1}^n S(U_i; \theta_0)(\theta - \theta_0)^T + \sum_{i=1}^n (\theta - \theta_0)H(\theta_0)(\theta - \theta_0)^T \right] \right\} + o_p(1). \quad (17) \end{aligned}$$

$I(\theta)$ represents the Fisher information matrix, then $I(\theta) = -E(H(\theta))$. According to Lemma 1, we know that when $(\delta_i A_{11} + (1 - \delta_i) B_{11})(\delta_i A_{22} + (1 - \delta_i) B_{22}) \neq (\delta_i A_{12} + (1 -$

$\delta_i)B_{12})^2$, the Fisher information matrix is positive definite, and its inverse exists. Combining with the properties of quadratic functions, Equation (17) can be further expressed as

$$M_n \leq \max_{k \in \Delta_\epsilon} \left\{ \left(\frac{1}{\sqrt{k}} \sum_{i=1}^k S(U_i; \theta_0) \right) I^{-1}(\theta_0) \left(\frac{1}{\sqrt{k}} \sum_{i=1}^k S(U_i; \theta_0)^T \right) + \left(\frac{1}{\sqrt{n-k}} \sum_{i=k+1}^n S(U_i; \theta_0) \right) I^{-1}(\theta_0) \left(\frac{1}{\sqrt{n-k}} \sum_{i=k+1}^n S(U_i; \theta_0)^T \right) - \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n S(U_i; \theta_0) \right) I^{-1}(\theta_0) \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n S(U_i; \theta_0)^T \right) \right\} + o_p(1).$$

Let

$$Z_k = \sum_{i=1}^k I^{-\frac{1}{2}}(\theta_0) S(U_i; \theta_0), Z_n = \sum_{i=1}^n I^{-\frac{1}{2}}(\theta_0) S(U_i; \theta_0),$$

then

$$\begin{aligned} M_n &\leq \max_{k \in \Delta_\epsilon} \left\{ \frac{1}{\sqrt{k}} Z_k Z_k^T \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{n-k}} (Z_n - Z_k)(Z_n - Z_k)^T \frac{1}{\sqrt{n-k}} - \frac{1}{\sqrt{n}} Z_n Z_n^T \frac{1}{\sqrt{n}} \right\} + o_p(1) \\ &= \max_{k \in \Delta_\epsilon} \left\{ \frac{n(n-k) Z_k Z_k^T + nk(Z_n - Z_k)(Z_n - Z_k)^T - k(n-k) Z_n Z_n^T}{nk(n-k)} \right\} + o_p(1) \\ &= \max_{k \in \Delta_\epsilon} \left\{ \left[n \cdot \frac{k}{n} \left(1 - \frac{k}{n} \right) \right]^{-\frac{1}{2}} \left(Z_k - \frac{k}{n} Z_n \right) \left(Z_k - \frac{k}{n} Z_n \right)^T \left[n \cdot \frac{k}{n} \left(1 - \frac{k}{n} \right) \right]^{-\frac{1}{2}} \right\} \\ &\quad + o_p(1) \\ &\leq \max_{k \in \Delta_\epsilon} V(t) V^T(t) + o_p(1), \end{aligned} \quad (18)$$

where

$$\begin{aligned} V(t) &= \left(\frac{[nt]}{n} \left(1 - \frac{[nt]}{n} \right) \right)^{-\frac{1}{2}} n^{-\frac{1}{2}} \left\{ Z_{[nt]}^T + (nt - [nt]) S(U_{[nt]+1}; \theta_0) - \frac{[nt]}{n} Z_n^T \right\} \\ &= (V_1(t), V_2(t)), \end{aligned}$$

for $nt \in \Delta_\epsilon$.

Without loss of generality, for $t \in [\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon]$,

$$V_i(t) \xrightarrow{D} [t(1-t)]^{-\frac{1}{2}} B_0(t), \quad i = 1, 2, \quad (19)$$

where $\xrightarrow{D}$ means converges in distribution, $B_0(t)$ is the Brownian bridge. Then,

$$\sup_{nt \in \Delta_\epsilon} \left| B_0(t) - B_0\left(\frac{1}{2}\right) \right| \rightarrow 0$$

and

$$\left[\frac{1}{2} \left(1 - \frac{1}{2} \right) \right]^{-1} B_0^2\left(\frac{1}{2}\right) \sim \chi_1^2.$$

So

$$M_n \leq \sup_{|t-\frac{1}{2}| < \epsilon} V(t) V^T(t) + o_p(1) \rightarrow \sup_{|t-\frac{1}{2}| < \epsilon} \frac{1}{t(1-t)} B_0^2(t) + \frac{1}{t(1-t)} B_0^2(t), \quad (20)$$

And then there is $P(M_n \leq s) \geq P(\chi^2(2) \leq s - o_p(1))$. That is,

$$\lim_{n \rightarrow \infty} P(M_n \leq s) \geq P(\chi_2^2 \leq s).$$

And because $M_n \geq \text{MIC}(n) - \text{MIC}(n/2) + 2 \log n$,

$$\overline{\lim}_{n \rightarrow \infty} P(M_n \leq s) \leq P(\chi_2^2 \leq s).$$

Therefore, $M_n \xrightarrow{D} \chi^2(2)$ as $n \rightarrow \infty$. $\square$

2.3. Profile Likelihood Function and Deviance Function

The confidence distribution can be regarded as a distribution that depends on samples, which can be used to study the interval estimation and point estimation of the parameter of interest [23]. In particular, it can provide the confidence interval of the parameter of interest at any nominal level through the confidence curve. In the change point problem, the location of the change point is a discrete variable, and the uncertainty analysis of the estimated change location is a challenging task. Similar to reference [24], we establish the confidence curve of the change point location based on profile likelihood and deviation function to analyze the uncertainty of the change location estimation.

Assuming $X_1, X_2, \dots, X_k$ is a sample from the population density function, $f(x, \theta_L)$, $x_1, x_2, \dots, x_k$ is the corresponding sample observation. Meanwhile, $X_{k+1}, \dots, X_n$ comes from the population density function $f(x, \theta_R)$, and $x_{k+1}, \dots, x_n$ is the corresponding sample observation. Then, the log-likelihood function of $x_1, x_2, \dots, x_n$ is

$$l(k, \theta_L, \theta_R) = \sum_{i=1}^k \log f(x_i, \theta_L) + \sum_{i=k+1}^n \log f(x_i, \theta_R). \quad (21)$$

For a given change point k , the profile log-likelihood function can be obtained by maximizing the log-likelihood function (21), where the profile log-likelihood function is defined as follows:

$$l_{prof}(k) = \max_{\theta_L, \theta_R} l(k, \theta_L, \theta_R) = l(k, \hat{\theta}_L, \hat{\theta}_R), \quad (22)$$

where $\hat{\theta}_L, \hat{\theta}_R$ are MLEs of θ_L and θ_R for a given k . Then, $\hat{k}$ can be obtained by

$$\hat{k} = \arg l_{prof}(\hat{k}) = \arg \max_k l_{prof}(k). \quad (23)$$

The deviance function of $\hat{k}$ is given by

$$D(k, X) = 2\{l_{prof}(\hat{k}) - l_{prof}(k)\}, \quad (24)$$

where $X = (X_1, X_2, \dots, X_n)$. To construct k 's confidence curve based on deviance function, we consider the distribution of $D(k, X)$ at k , which is denoted as $R_k(t) = P_{k, \hat{\theta}_L, \hat{\theta}_R} \{D(k, X) \leq t\}$. However, the location of the change point is discrete; $R_k(t)$ does not satisfy Wilks' theorem. Therefore, we compute $R_k(t)$ through simulation. And the confidence curve of k is defined by

$$cc(k, x_{obs}) = P_{k, \hat{\theta}_L, \hat{\theta}_R} \{D(k, X) \leq D(k, x_{obs})\}, \quad (25)$$

and it can be obtained through the following simulation,

$$cc(k, x_{obs}) = \frac{1}{B} \sum_{j=1}^B I\{D(k, X_j^*) < D(k, x_{obs})\}, \quad (26)$$

where B is a large number and $B = 1000$ typically. X_j^* is a sample from $f(x, \hat{\theta}_L)$ and $f(x, \hat{\theta}_R)$ with a given k .

For comparison purposes, we estimate the location of the change point through (14) for computing the deviation function. The specific simulation steps are as follows,

- Step 1. Given sample size n and change point k_{true} , generating a group of random samples with change point k_{true} based on parameters θ_L and θ_R , record them as x_{obs} , and calculate the deviation $D(k, x_{obs})$ at each possible change point based on x_{obs} .
- Step 2. Compute $(\hat{k}, \hat{\theta}_L, \hat{\theta}_R)$ based on Step 1.
- Step 3. Generate $B = 1000$ random samples $X_j^* (j = 1, \dots, B)$, where $x_{j1}^*, \dots, x_{jk}^*$ are from $f(x, \hat{\theta}_L)$, $x_{j,k+1}^*, \dots, x_{jn}^*$ are from $f(x, \hat{\theta}_R)$. And compute $D(k, X_j^*)$. Then, the confidence curve $cc(k, x_{obs})$ under the possible change locations is the frequencies of $D(k, X_j^*) < D(k, x_{obs})$.
- Step 4. Repeat Step 1 to Step 3 for $N = 1000$ times to obtain N confidence curves about the change point k .
- Step 5. Given significance level α , the corresponding confidence set of each confidence curve is $K_{set} = \{k : cc(k, x_{obs}) \leq 1 - \alpha\}$.
- Step 6. Coverage probabilities of confidence sets. Based on the confidence set K_{set} obtained from each confidence curve, then the frequency of k_{true} in K_{set} is the coverage probability of the confidence set at the corresponding confidence level $1 - \alpha$:

$$cp = \frac{1}{N} \sum_{i=1}^N I\{k_{true} \in K_{set}\}.$$

- Step 7. Average size of confidence sets. Since the confidence set is the set of estimates of the change location, it is a set of discrete points. Thus, unlike the calculation method for continuous interval length, we define the number of elements in the confidence set as the size of the confidence set. Noting the number of elements in each confidence set K_{set} as $m_i (i = 1, \dots, N)$, then the average size of confidence sets is obtained by

$$l_{set} = \frac{1}{N} \sum_{i=1}^N m_i.$$

In conclusion, when cp is closer to $1 - \alpha$, the smaller l_{set} means the corresponding estimate method is better.

3. Simulation Study

Based on the no information characteristic of the uniform distribution and the memorylessness of the exponential distribution, these two types of distributions are commonly used as censoring distributions [16,25–28]. In this study, simulations are conducted in terms of Type I error, power, and coverage probabilities of confidence sets and average size of confidence sets based on uniform and exponential censoring distributions with various censoring rates, sample sizes and change locations.

Case 1: Firstly, the sample sizes are set as $n = \{50, 100, 150\}$, the true change locations are $k = \{n/4, n/2, 3n/4\}$ and the significance level is $\alpha = 0.05$. For a given change location k , we generate a set of survival time data $\{T_1, T_2, \dots, T_k\}$ from $LBLN(0, 1)$, $\{T_{k+1}, T_{k+2}, \dots, T_n\}$ from $LBLN(\mu_R, \sigma_R)$, where $\mu_R = \{-2, -1, 0, 1, 2\}$, $\sigma_R = \{0.5, 1.0, 1.5, 2.0\}$. The censoring time are from $U(0, m)$, where m determines the censoring rate of survival time observations. And it can be calculated by $P(T_i > C_i) = cr$, where cr is the censoring proportion given in advance. Let $cr = \{10\%, 30\%\}$, then $m = \{10e^{1.5}, 10/3e^{1.5}\}$.

Case 2: The sample size, change location, significance level, and parameters are the same as in Case 1. Different from Case 1, the censoring time observations are from $Exp(m)$, where m is determined by the censoring rate of survival time observations. Two different values of $m = \{0.025, 0.101\}$ are calculated by $P(T_i > C_i) = 10\%$ and $P(T_i > C_i) = 30\%$.

3.1. Critical Values and Probability of Type I Error

First of all, we obtain the critical values of the test and probability of Type I error through simulations. The probability of Type I error is approximated by the frequency of Type I errors in N repeated simulation experiments, given by $P_{errorI} = S_{nI}/N$, where S_{nI} represents the number of Type I errors in N repeated trials with a sample size of n . According to the Central Limit Theorem, S_{nI}/N approximately follows a normal distribution with mean of α and variance of $\alpha(1 - \alpha)/N$, that is,

$$\frac{S_{nI}}{N} \sim N\left(\alpha, \left[\sqrt{\frac{\alpha(1 - \alpha)}{N}}\right]^2\right), \quad (27)$$

then S_{nI}/N is likely to fall within $[\alpha - 3\sqrt{\alpha(1 - \alpha)/N}, \alpha + 3\sqrt{\alpha(1 - \alpha)/N}]$ with a high probability. This means that the values of S_{nI}/N fluctuate within this interval, with the length of the fluctuation interval primarily determined by the number of repeated simulation experiments N . In our simulation, $\alpha = 0.05$, $N = 1000$. Then, S_{nI}/N is highly likely to fluctuate within $[0.029, 0.071]$. And the specific results are shown in Table 1. From columns 7–10 in Table 1, it can be seen that the frequency of Type I error under both censoring rates almost all fluctuate within the range of $[0.029, 0.071]$ based on uniform and exponential censoring, which means the Type I error is effectively controlled.

Table 1. Critical values and the probability of Type I error.

<i>cr</i>	<i>n</i>	Critical Values				Frequency of Type I Error			
		Uniform Censoring		Exponential Censoring		Uniform Censoring		Exponential Censoring	
		LRT	MIC	LRT	MIC	LRT	MIC	LRT	MIC
10%	50	12.833	11.849	12.721	11.499	0.043	0.044	0.051	0.053
	100	12.952	11.644	13.282	12.195	0.065	0.061	0.062	0.055
	150	13.792	12.139	13.454	11.981	0.061	0.062	0.071	0.064
	200	19.839	17.790	15.305	13.361	0.045	0.051	0.050	0.048
	300	19.413	17.283	16.399	14.036	0.049	0.053	0.047	0.050
	400	21.132	19.067	15.850	13.406	0.048	0.049	0.048	0.053
30%	50	14.103	13.131	15.249	14.127	0.057	0.056	0.045	0.044
	100	16.057	14.486	15.600	14.720	0.045	0.041	0.059	0.054
	150	16.122	14.594	16.692	15.377	0.048	0.049	0.054	0.052
	200	19.232	17.951	14.800	13.103	0.054	0.051	0.053	0.054
	300	21.202	19.020	15.169	13.408	0.052	0.053	0.050	0.051
	400	21.871	19.655	16.126	13.543	0.047	0.051	0.045	0.053

3.2. Power Comparison

Tables 2 and 3 show the power of the likelihood ratio test method and the MIC criterion test method in the case of uniform and exponential censoring, respectively.

Taking the test results of uniform censored distribution as an example, we can obtain the following conclusions. First, excluding these combinations with equal powers, the test power calculated based on MIC is mostly higher than that based on the likelihood ratio under the same parameter settings. Second, the powers of the LRT-based method and MIC-based method increase with the increase in sample size. For example, when $(\mu_R, \sigma_R) = (1.0, 0.5)$, $cr = 10\%$, and $k = n/2$, the power of the LRT-based method increases from 0.514 ($n = 50$) to 0.910 ($n = 150$), the power of the MIC-based method increases from 0.547 ($n = 50$) to 0.913 ($n = 150$). Third, in most cases, the closer the change point occurs to the midpoint of the data series, the higher the power of the test. For example, the power of the MIC-based method is 0.587 at $k = 25$, while the power is 0.427 at $k = 12$ under $(\mu_R, \sigma_R) = (1, 1)$ and $cr = 10\%$. Fourth, when the change point position and parameter settings are same, the power of the test with a censoring rate of 10% is significantly higher than that of a test with a censoring rate of 30% in most cases. For instance, the power of

MIC-based method is 0.817 when $n = 100, k = 25, cr = 10\%$, $(\mu_R, \sigma_R) = (-2, 2)$. But the power is 0.539 when $cr = 30\%$. This may be because a larger censoring rate masks the true pattern of change and increases the difficulty of the change point test. Finally, when the sample size, change point position, and censoring rate are the same, the greater the parameter change and the higher the power of the change point test. For example, when the parameter increases from $(0, 1)$ to $(0, 1.5)$, the test power based on MIC is $(0.515, 0.769, 0.596)$ when $n = 50, cr = 10\%$, and increases to $(0.976, 0.999, 0.978)$ when the parameter increases to $(1, 1.5)$.

Table 2. The power of the change point test under uniform censoring ($\alpha = 0.05$).

n	k	cr	$(-2, 2)$		$(1, 0.5)$		$(2, 0.5)$		$(1, 1)$		$(0, 1.5)$		$(1, 1.5)$	
			LRT	MIC	LRT	MIC	LRT	MIC	LRT	MIC	LRT	MIC	LRT	MIC
50	12	10%	0.367	0.371	0.456	0.460	0.887	0.887	0.422	0.427	0.514	0.515	0.975	0.976
	25		0.624	0.677	0.514	0.547	0.950	0.961	0.587	0.629	0.732	0.769	0.997	0.999
	38		0.502	0.502	0.313	0.316	0.840	0.843	0.422	0.425	0.594	0.596	0.977	0.978
	12	30%	0.216	0.222	0.314	0.322	0.738	0.744	0.421	0.421	0.420	0.422	0.877	0.884
	25		0.365	0.410	0.328	0.347	0.900	0.915	0.525	0.555	0.622	0.670	0.976	0.985
	38		0.284	0.293	0.234	0.239	0.759	0.764	0.392	0.395	0.462	0.466	0.892	0.892
100	25	10%	0.800	0.817	0.713	0.730	0.986	0.988	0.809	0.803	0.911	0.916	1.000	1.000
	50		0.962	0.972	0.793	0.815	0.998	0.998	0.935	0.953	0.982	0.987	1.000	1.000
	75		0.847	0.858	0.657	0.643	0.984	0.986	0.814	0.815	0.928	0.931	1.000	1.000
	25	30%	0.519	0.539	0.473	0.500	0.981	0.980	0.753	0.766	0.823	0.834	1.000	1.000
	50		0.687	0.724	0.539	0.540	0.994	0.999	0.887	0.897	0.925	0.945	1.000	1.000
	75		0.575	0.555	0.382	0.391	0.986	0.986	0.738	0.756	0.704	0.801	0.999	0.999
150	38	10%	0.977	0.981	0.885	0.888	1.000	1.000	0.961	0.963	0.998	0.998	1.000	1.000
	75		0.996	0.996	0.910	0.913	1.000	1.000	0.995	0.996	0.998	0.999	1.000	1.000
	112		0.978	0.982	0.829	0.844	0.999	0.999	0.969	0.973	0.992	0.994	1.000	1.000
	38	30%	0.756	0.778	0.609	0.645	0.999	0.999	0.932	0.935	0.949	0.941	1.000	1.000
	75		0.911	0.925	0.650	0.666	1.000	1.000	0.985	0.992	0.995	0.998	1.000	1.000
	112		0.772	0.785	0.540	0.565	1.000	1.000	0.931	0.935	0.942	0.943	1.000	1.000

Table 3. The power of change point test under exponential censoring ($\alpha = 0.05$).

n	k	cr	$(-2, 2)$		$(1, 0.5)$		$(2, 0.5)$		$(1, 1)$		$(0, 1.5)$		$(1, 1.5)$	
			LRT	MIC	LRT	MIC	LRT	MIC	LRT	MIC	LRT	MIC	LRT	MIC
50	12	10%	0.372	0.373	0.490	0.486	0.871	0.876	0.455	0.447	0.520	0.524	0.977	0.975
	25		0.617	0.664	0.513	0.543	0.931	0.942	0.612	0.661	0.749	0.788	0.997	0.997
	38		0.527	0.516	0.312	0.325	0.816	0.815	0.445	0.448	0.599	0.596	0.982	0.981
	12	30%	0.242	0.253	0.319	0.307	0.742	0.739	0.404	0.408	0.422	0.415	0.903	0.898
	25		0.377	0.420	0.349	0.361	0.884	0.906	0.563	0.616	0.589	0.631	0.972	0.976
	38		0.292	0.285	0.225	0.229	0.755	0.766	0.417	0.417	0.457	0.455	0.878	0.876
100	25	10%	0.824	0.820	0.755	0.748	0.992	0.991	0.826	0.806	0.912	0.906	1.000	1.000
	50		0.964	0.977	0.782	0.794	0.999	0.999	0.920	0.943	0.990	0.994	1.000	1.000
	75		0.876	0.867	0.661	0.647	0.991	0.991	0.852	0.838	0.933	0.930	1.000	1.000
	25	30%	0.503	0.488	0.543	0.535	0.978	0.979	0.768	0.761	0.815	0.802	0.999	0.999
	50		0.733	0.754	0.547	0.548	0.998	0.998	0.891	0.913	0.916	0.939	1.000	1.000
	75		0.574	0.567	0.418	0.401	0.979	0.980	0.780	0.758	0.801	0.794	0.999	0.999
150	38	10%	0.978	0.975	0.898	0.886	1.000	1.000	0.970	0.966	0.993	0.990	1.000	1.000
	75		0.997	0.997	0.922	0.934	1.000	1.000	0.997	0.997	1.000	1.000	1.000	1.000
	112		0.983	0.983	0.808	0.798	0.999	0.999	0.965	0.962	0.996	0.996	1.000	1.000
	38	30%	0.789	0.777	0.602	0.584	1.000	1.000	0.926	0.920	0.953	0.952	1.000	1.000
	75		0.908	0.923	0.663	0.663	1.000	1.000	0.992	0.994	0.996	0.996	1.000	1.000
	112		0.817	0.803	0.567	0.542	0.998	0.998	0.933	0.926	0.952	0.945	1.000	1.000

3.3. Coverage Probabilities and Average Length of Confidence Sets

To evaluate the uncertainty of change point location estimation, some simulations are conducted to study the coverage probabilities and average lengths of confidence sets for change point location estimation under a uniform censored distribution with censoring rates of 10% and 30%, respectively. The specific results are shown in Tables 4 and 5.

Table 4. The coverage probability of confidence sets for change point location estimation under different parameter combinations.

n	(μ_R, σ_R)	$1 - \alpha$	$cr = 10\%$				$cr = 30\%$			
			$k = 12$		$k = 25$		$k = 12$		$k = 25$	
			LRT	MIC	LRT	MIC	LRT	MIC	LRT	MIC
50	(1.0, 1.5)	0.90	0.879	0.883	0.884	0.886	0.858	0.860	0.854	0.854
		0.95	0.932	0.934	0.936	0.936	0.917	0.917	0.914	0.914
		0.99	0.978	0.978	0.982	0.982	0.971	0.971	0.974	0.974
	(1.0, 2.0)	0.90	0.882	0.883	0.893	0.893	0.839	0.839	0.840	0.840
		0.95	0.933	0.933	0.933	0.933	0.902	0.902	0.911	0.911
		0.99	0.980	0.980	0.979	0.979	0.973	0.973	0.984	0.984
	(2.0, 2.0)	0.90	0.882	0.889	0.878	0.879	0.833	0.834	0.787	0.789
		0.95	0.930	0.931	0.928	0.928	0.901	0.900	0.836	0.834
		0.99	0.983	0.983	0.974	0.974	0.966	0.966	0.887	0.885
100	(1.0, 1.5)	0.90	0.880	0.880	0.870	0.870	0.910	0.910	0.870	0.870
		0.95	0.960	0.960	0.930	0.930	0.950	0.950	0.930	0.930
		0.99	0.990	0.990	0.980	0.980	0.990	0.990	0.960	0.960
	(1.0, 2.0)	0.90	0.883	0.883	0.860	0.860	0.930	0.930	0.920	0.910
		0.95	0.933	0.933	0.920	0.920	0.970	0.970	0.960	0.940
		0.99	0.984	0.984	0.980	0.980	0.990	0.990	0.980	0.980
	(2.0, 2.0)	0.90	0.880	0.880	0.850	0.860	0.820	0.820	0.850	0.850
		0.95	0.920	0.920	0.880	0.890	0.900	0.900	0.960	0.960
		0.99	0.990	0.990	0.970	0.980	0.970	0.970	0.970	0.970

From Table 4, we can observe that the coverage probability of confidence sets for change point location estimation with a censoring rate of 10% is higher than that corresponding to a censoring rate of 30%, and closer to the given nominal level.

From Table 5, it can be seen that under the same sample size, the same change point location, and the same parameter settings, the average length of confidence sets for change point location estimation corresponding to a censoring rate of 10% is smaller than that corresponding to a censoring rate of 30%. When the change point location is closer to the middle of the dataset, the average length of the confidence set is shorter. From the perspective of testing methods, except for the case of equality, the testing method based on the MIC criterion can provide a shorter confidence set. Under the same testing method and parameter settings, as the sample size increases, the average length of the confidence set becomes shorter. In summary, the smaller the censoring rate, the closer the coverage probability of the confidence set for the change point location estimation is to the nominal level, and the shorter the average length of the confidence sets.

Table 5. The average length of the confidence set for estimating the position of change points under different parameter combinations.

n	(μ_R, σ_R)	$1 - \alpha$	$cr = 10\%$				$cr = 30\%$			
			$k = 12$		$k = 25$		$k = 12$		$k = 25$	
			LRT	MIC	LRT	MIC	LRT	MIC	LRT	MIC
50	(1.0, 1.5)	0.90	1.617	1.606	1.258	1.256	2.728	2.712	2.194	2.194
		0.95	2.189	2.186	1.606	1.606	3.854	3.840	3.106	3.106
		0.99	4.068	4.062	2.832	2.832	6.842	6.818	5.392	5.392
	(1.0, 2.0)	0.90	1.892	1.886	1.842	1.840	3.196	3.173	3.102	3.102
		0.95	2.735	2.732	2.581	2.581	4.677	4.652	4.413	4.412
		0.99	4.884	4.881	4.436	4.437	8.218	8.198	7.437	7.436
	(2.0, 2.0)	0.90	1.345	1.338	1.279	1.277	2.157	2.151	1.911	1.899
		0.95	1.758	1.754	1.691	1.691	3.078	3.068	2.691	2.682
		0.99	3.207	3.207	2.999	2.999	5.598	5.593	4.767	4.755
100	(1.0, 1.5)	0.90	1.230	1.230	1.310	1.310	1.990	1.990	2.170	2.170
		0.95	1.680	1.680	1.690	1.690	2.770	2.770	3.070	3.070
		0.99	3.010	3.010	2.790	2.790	5.010	5.010	5.080	5.080
	(1.0, 2.0)	0.90	1.772	1.771	1.820	1.820	3.340	3.360	2.230	2.160
		0.95	2.459	2.459	2.420	2.420	4.650	4.650	3.400	3.290
		0.99	4.163	4.163	3.970	3.970	7.790	7.740	5.500	5.320
	(2.0, 2.0)	0.90	1.240	1.240	1.250	1.240	2.020	2.020	2.110	2.110
		0.95	1.680	1.680	1.600	1.610	2.780	2.770	2.940	2.940
		0.99	2.930	2.930	2.830	2.780	4.820	4.810	4.820	4.820

4. Application: Analysis of Survival Data for Heart Transplantation

This section applies the proposed testing method to the Stanford heart transplantation data from February 1980 [29]. The Stanford heart transplantation project ran from October 1967 to February 1980. In total, 184 patients underwent heart transplantation. One patient's survival time was 0, which is excluded from this study. Therefore, there are 183 samples, of which 71 patients have censored status, resulting in a censored rate of 38.798%, and their censored situation is random. Because only patients who had undergone heart transplantation and were alive at the beginning of data collection are included as observations, this dataset is a length-biased dataset. Next, there is some right censoring in the survival time of patients, meaning that some patients were still alive after the end of the observation period, and their exact survival time could not be determined. For such a randomly right-censored dataset, the proposed method was used to test whether there is a change point in the data.

Before testing, patients were divided into 43 groups according to their age. The median survival time of each group of patients was calculated as shown in Figure 1a. A Q-Q plot was generated to test whether the median survival time of patients follows a length-biased lognormal distribution as shown in Figure 1b. The results indicate that the patient survival time approximately follows a length-biased lognormal distribution. To make this conclusion more convincing, we performed a Kolmogorov–Smirnov test on the dataset. The test statistic was 0.210 and the corresponding p value was 0.306. Since the p value is much greater than 0.05, we accepted the null hypothesis, indicating that there is sufficient evidence to suggest that the dataset follows a length-biased lognormal distribution.

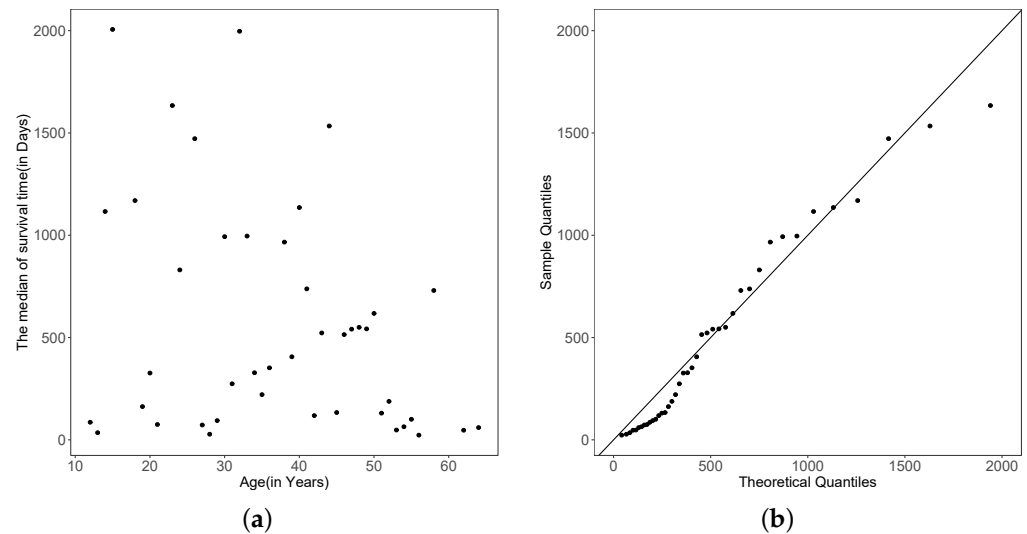


Figure 1. (a) Scatter plot of the median survival time. (b) Q-Q plot for length-biased lognormal distribution.

Combining the binary segmentation method, the LRT-based method and the MIC-based method are applied for this data. The values of the test statistics and corresponding p values based on this dataset are calculated and presented in Table 6.

Table 6. Change point test results for heart transplant survival time data.

No.	T_n	p_{LRT}	M_n	p_{MIC}	$\hat{k}$
1–43	16.303	0.001	15.032	0.001	34
1–34	8.093	0.013	8.045	0.012	15
16–34	17.726	0.005	17.066	0.005	30

In Table 6, p_{LRT} and p_{MIC} represent the p values for the two testing methods. From the p values, it can be observed that all p values are less than 0.05, which means that the null hypothesis for no change point should be rejected. Therefore, there are three change points in the dataset located at positions $\{15, 30, 34\}$, corresponding to ages 30, 46, and 50 in the data. This means that there are differences in the survival time among patients of different age groups after undergoing heart transplant surgery. The parameter estimation results before and after the change points are $(\hat{\mu}_1, \hat{\sigma}_1) = (1.807, 2.308)$, $(\hat{\mu}_2, \hat{\sigma}_2) = (5.424, 0.996)$, $(\hat{\mu}_3, \hat{\sigma}_3) = (6.346, 0.056)$, and $(\hat{\mu}_4, \hat{\sigma}_4) = (3.598, 0.945)$, respectively.

5. Conclusions And Discussions

Due to the frequent occurrence of random right censoring in life data, this study investigates the parameter change point test problem of length-biased lognormal distribution based on random right-censoring data. For the parameter change point model with random right censoring, a test statistic based on MIC is constructed, and the corresponding testing method is presented. Under the sufficient condition of the Fisher information matrix being positive definite, it is proven that the asymptotic distribution of the MIC-based test statistic is a chi-square distribution with two degrees of freedom under the null hypothesis.

To demonstrate the performance of the testing method, simulations are conducted to investigate the power of the change point test under uniform and exponential distributions. The simulation study considered various combinations of censoring rates and parameter settings. The simulation results indicate that the power of the method based on MIC is generally higher than that of the method based on LRT in most cases. Moreover, as the censored rate decreases, the testing power increases, which means that as the censoring rate decreases, the number of censored data decreases while the amount of complete lifetime data increases. Consequently, the data contain more valuable lifetime information, leading to an increase in the power of hypothesis testing.

For the purpose of assessing the uncertainty in estimation of change point location, the coverage probability and the average lengths of confidence sets are calculated. Simulation results indicate that as the censoring rate decreases, the coverage rate approaches the nominal level in most cases, and the average length of the confidence sets becomes shorter. From the perspective of testing methods, the method based on the MIC criterion yields shorter confidence sets. Under the same testing method and parameter settings, with an increase in sample size, the average length of the confidence sets decreases.

From the simulations, we observe that differences in sample size, change point location, and censoring rate all have an impact on the power of change point testing in random right-censored data. Therefore, factors such as testing methods, censored distribution, and censoring rate should be considered comprehensively when investigating the change point problem with censored data.

The unique contributions of the proposed method can be concluded as follows. In our study, the change point problem is viewed as a model selection problem, where we aim to select the better option between the null hypothesis of no change and the alternative hypothesis of at least one change. As a result, methods commonly used for model selection can be applied to change-point-testing problems. However, the change point problem involves a special parameter, namely, the change location. In this paper, we employ the MIC-based method, which takes into account the contribution of change point positions to model complexity in the penalty term. This method is utilized to detect change points in the length-biased lognormal distribution with randomly right-censored data.

The traditional change point detection method is based on the likelihood ratio. In the simulation section, we conduct a comprehensive comparison of the detection performance between the MIC-based method and the LRT-based method. The results indicate that the coverage probability and the average length of the confidence sets for change point estimation are comparable between the two methods. However, under the same parameter settings, the test powers based on the MIC method are generally higher than those based on the likelihood ratio.

There is still much work to be performed in the research of change point detection for the length-biased lognormal distribution in future. (i) In terms of statistical theory, this study primarily focuses on the asymptotic distribution of the test statistic. The estimation of the change point location is discrete, and asymptotic properties, such as the consistency of the estimator, are challenging issues that need to be addressed in future research. (ii) In terms of the method application, the method proposed in this paper can be used for change point detection in data with length-biased distribution characteristics. This will help to establish accurate models, make reasonable predictions of patients' remaining lifetimes, and analyze the effectiveness of treatment methods or medications for patients.

Author Contributions: Conceptualization, W.N. and Y.T.; Methodology: M.L.; writing—original draft preparation, M.L.; writing—review and editing, W.N. and Y.T.; supervision, Y.T.; funding acquisition, Y.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under Grant 12131001.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

LBLN	Length-biased lognormal distribution
LRT	Likelihood ratio test
AIC	Akaike information criterion
SIC	Schwarz information criterion
MIC	Modified information criterion
MLE	Maximum likelihood estimation

Appendix A

Wald Conditions and Regularity Conditions.

- W1. The distribution of X is either discrete for (μ, σ) or is absolutely continuous for (μ, σ) .
 W2. For sufficiently small δ and for sufficiently large ρ , the expected values are

$$E \left[\log \sup_{|\mu' - \mu| < \delta, |\sigma' - \sigma| < \delta} f(X; \mu', \sigma') \right]^2 < \infty,$$

$$E \left[\log \sup_{|\mu' - \mu| > \rho, |\sigma' - \sigma| > \rho} f(X; \mu', \sigma') \right]^2 < \infty.$$

- W3. The density function $f(x; \mu, \sigma)$ is continuous in (μ, σ) for all x .
 W4. The cumulative distribution function $F(x; \mu, \sigma)$ is identifiable.
 W5. $\lim_{\|\theta\| \rightarrow \infty} f(x; \theta) = 0$ for all x .
 W6. The parameter space Θ is a closed subset of the two-dimensional Cartesian space.
 W7. $f(x; \mu, \sigma, \delta) = \sup_{|\mu' - \mu| < \delta, |\sigma' - \sigma| < \delta} f(X; \mu', \sigma')$ is a measurable function of x for any fixed (μ, σ) and δ .
 R1. For each $\theta = (\mu, \sigma) \in \Theta$, the following derivatives exist for all x ,

$$\frac{\partial \log f(x; \mu, \sigma)}{\partial \theta}, \frac{\partial^2 \log f(x; \mu, \sigma)}{\partial \theta^2}, \frac{\partial^3 \log f(x; \mu, \sigma)}{\partial \theta^3}.$$

- R2. For θ in the neighborhood $N(\theta_0)$, there exist function $g(x)$ and $H(x)$ such that the following relations hold for all x ,

$$\left| \frac{\partial f(x, \theta)}{\partial \theta} \right| \leq g(x), \left| \frac{\partial^2 f(x, \theta)}{\partial \theta^2} \right| \leq g(x),$$

$$\left| \frac{\partial^2 \log f(x, \theta)}{\partial \theta^2} \right|^2 \leq H(x), \left| \frac{\partial^3 \log f(x, \theta)}{\partial \theta^3} \right| \leq H(x),$$

and

$$\int g(x) dx < \infty, \quad E_{\theta}(H(X)) < \infty.$$

- R3. For each $\theta \in \Theta$,

$$0 < E_{\theta} \left[\left(\frac{\partial \log f(X; \theta)}{\partial \theta} \right)^2 \right] < \infty, \quad E_{\theta} \left[\left| \frac{\partial \log f(X; \theta)}{\partial \theta} \right|^3 \right] < \infty$$

References

- Asgharian, M.; M'Lan, C.E.; Wolfson, D.B. Length-Biased Sampling with Right Censoring. *J. Am. Stat. Assoc.* **2002**, *97*, 201–209. [\[CrossRef\]](#)
- Asgharian, M.; Wolfson, D.B. Asymptotic behavior of the unconditional NPMLE of the length-biased survivor function from right censored prevalent cohort data. *Ann. Stat.* **2005**, *33*, 2109–2131. [\[CrossRef\]](#)
- Qin, J.; Ning, J.; Liu, H.; Shen, Y. Maximum Likelihood Estimations and EM Algorithms with Length-Biased Data. *J. Am. Stat. Assoc.* **2011**, *106*, 1434–1449. [\[CrossRef\]](#)
- Kvam, P. Length Bias in the Measurements of Carbon Nanotubes. *Technometrics* **2008**, *50*, 462–467. [\[CrossRef\]](#)
- Turkson, A.J.; Ayiah-Mensah, F.; Nimoh, V. Handling Censoring and Censored Data in Survival Analysis: A Standalone Systematic Literature Review. *Int. J. Math. Math. Sci.* **2021**, *2021*, 9307475. [\[CrossRef\]](#)
- Klein, J.P.; Moeschberger, M.L. *Survival Analysis: Techniques for Censored and Truncated Data*; Springer: New York, NY, USA, 2003; pp. 63–90. [\[CrossRef\]](#)
- Sansgiry, P.; Akman, O. Transformations of the lognormal distribution as a selection model. *Am. Stat.* **2000**, *54*, 307–309. [\[CrossRef\]](#)

8. Sansgiry, P.; Akman, O. Reliability estimation via length-biased transformation. *Commun. Stat.-Theory Methods* **2006**, *30*, 2473–2479. [\[CrossRef\]](#)
9. Ratnaparkhi, M.; Naik-Nimbalkar, U. The length-biased lognormal distribution and its application in the analysis of data from oil field exploration studies. *J. Mod. Appl. Stat. Meth.* **2012**, *11*, 255–260. [\[CrossRef\]](#)
10. Page, E.S. Continuous Insepection Schemes. *Biometrika* **1954**, *41*, 100–115. [\[CrossRef\]](#)
11. Cai, X.; Said, K.K.; Ning, W. Change-point Analysis with Bathtub Shape for the Exponential Distribution. *J. Appl. Stat.* **2016**, *43*, 2740–2750. [\[CrossRef\]](#)
12. Wang, P.; Tang, Y.; Bae, S.J.; He, Y. Bayesian Analysis of Two-Phase Degradation Data Based on Change-Point Wiener Process. *Reliab. Eng. Syst. Saf.* **2018**, *170*, 244–256. [\[CrossRef\]](#)
13. Cai, X.; Tian, Y.; Ning, W. Change-point Analysis of the Failure Mechanisms Based on Accelerated Life Tests. *Reliab. Eng. Syst. Saf.* **2019**, *188*, 515–522. [\[CrossRef\]](#)
14. Kosorok, M.R.; Song, R. Inference under Right Censoring for Transformation Models with a Change-point Based on a Covariate Threshold. *Ann. Stat.* **2007**, *35*, 957–989. [\[CrossRef\]](#)
15. Rabhi, Y.; Asgharian, M. Inference under Biased Sampling and Right Censoring for a Change Point in the Hazard Function. *Bernoulli* **2017**, *23*, 2720–2745. [\[CrossRef\]](#)
16. Chen, Y.J.; Ning, W.; Gupta, A.K. Empirical likelihood based detection procedure for change point in mean residual life functions under random censorship. *Pharm. Stat.* **2016**, *15*, 246–254. [\[CrossRef\]](#)
17. Vostrikova, L. Detecting Disorder in Multidimensional Random Processes. *Sov. Math. Dokl.* **1981**, *24*, 55–59.
18. Akaike, H.; Petrov, B.N.; Czaki, F. Information Theory and an Extension of the Maximum Likelihood Principle. In *Second International Symposium on Information Theory*; Springer: New York, NY, USA, 1973; pp. 267–281.
19. Schwarz, G.E. Estimating the Dimension of a Model. *Ann. Stat.* **1978**, *6*, 31–38. [\[CrossRef\]](#)
20. Chen, J.; Gupta, A.K. Testing and Locating Variance Change-points with Application to Stock Prices. *J. Am. Stat. Assoc.* **1997**, *92*, 739–747. [\[CrossRef\]](#)
21. Chen, J.; Gupta, A.K. Change Point Analysis of a Gaussian Model. *Stat. Pap.* **1999**, *40*, 323–333. [\[CrossRef\]](#)
22. Chen, J.; Gupta, A.K.; Pan, J. Information Criterion and Change Point Problem for Regular Models. *Sankhyā Ind. J. Stat.* **2006**, *68*, 252–282.
23. Singh, K.; Xie, M.; Strawderman, W.E. Confidence Distribution (CD): Distribution Estimator of a Parameter. In *Complex Datasets and Inverse Problems: Tomography, Networks and Beyond*; Lecture Notes-Monograph Series; IMS: Ann Arbor, MI, USA, 2007; Volume 54, pp. 132–150.
24. Cunen, C.; Hermansen, G.; Hjort, N.L. Confidence Distributions for Change-Points and Regime Shifts. *J. Stat. Plan. Infer.* **2017**, *195*, 14–34. [\[CrossRef\]](#)
25. Wan, F. Simulating Survival Data with Predefined Censoring Rates for Proportional Hazards Models. *Statist. Med.* **2017**, *36*, 838–854. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Deng, Y.; You, C.; Liu, Y.; Qin, J.; Zhou, X.H. Estimation of Incubation Period and Generation Time Based on Observed Length-Biased Epidemic Cohort with Censoring for COVID-19 Outbreak in Chin. *Biometrics* **2021**, *77*, 929–941. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Wan, F. Simulating Survival Data with Predefined Censoring Rates Under a Mixture of Non-Informative Right Censoring Schemes. *Commun. Stat. Simul. Comput.* **2022**, *51*, 3851–3867. [\[CrossRef\]](#)
28. Shahin, A.H.; Zhao, A.; Whitehead, A.C.; Alexander, D.C.; Jacob, J.; Barber, D. CenTime: Event-Conditional Modelling of Censoring in Survival Analysis. *Med. Image Anal.* **2024**, *91*, 103016.
29. Miller, R.; Halpern, J. Regression with Censored Data. *Biometrika* **1982**, *69*, 521–531. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.