

Article

Evidential Estimation of an Uncertain Mixed Exponential Distribution under Progressive Censoring

Kuang Zhou * and Yimin Shi

School of Mathematics and Statistics, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, China; ymshi@nwpu.edu.cn

* Correspondence: kzhoumath@nwpu.edu.cn

Received: 13 August 2020; Accepted: 29 September 2020; Published: 30 September 2020



Abstract: In this paper, the evidential estimation method for the parameters of the mixed exponential distribution is considered when a sample is obtained from Type-II progressively censored data. Different from the traditional statistical inference methods for censored data from mixture models, here we consider a very general form where there is some uncertain information about the sub-class labels of units. The partially specified label information, as well as the censored data are represented in a united frame by mass functions within the theory of belief functions. Following that, the evidential likelihood function is derived based on the completely observed failures and the uncertain information included in the data. Then, the optimization method using the evidential expectation maximization algorithm (E2M) is introduced. A general form of the maximal likelihood estimates (MLEs) in the sense of the evidential likelihood, named maximal evidential likelihood estimates (MELEs), can be obtained. Finally, some Monte Carlo simulations are conducted. The results show that the proposed estimation method can incorporate more information than traditional EM algorithms, and this confirms the interest in using uncertain labels for the censored data from finite mixture models.

Keywords: uncertain mixed exponential distribution; evidential likelihood; belief function theory; progressive censoring

1. Background

Mixture models are of great importance in many applied sciences such as survival analysis, pattern recognition, image analysis, economics, and so on [1,2]. In reliability analysis, there are only one population and one type of failure in the simple case of life distributions. However, in real applications, there may be more than one failure cause [3,4]. In this case, the response of the modeling process can be seen as from several distinct sub-populations, and finite mixture models can be used to represent the time of failures.

Suppose the life lengths of these n units are independent and identically distributed (i.i.d.) random variables with probability density function (pdf) $f(x; \theta)$ and cumulative distribution function (cdf) $F(x; \theta)$. Consider the mixture model of the form:

$$f(x; \theta) = \sum_{k=1}^m \pi_k f_k(x; \lambda_k), \quad (1)$$

where $\theta = \{\lambda_1, \dots, \lambda_m; \pi_1, \dots, \pi_m\}$, with $\pi_j \geq 0, j = 1, \dots, m$ and $\sum_j \pi_j = 1$. In this paper, we focus on the case where $f_k(x; \lambda_k)$ is the exponential density, although the model we develop can be

applied more generally. The mixed exponential distribution (MED) model with m components has its pdf and cdf respectively as:

$$f(x; \theta) = \sum_{k=1}^m \pi_k f_k(x; \lambda_k) = \sum_{k=1}^m \pi_k \lambda_k e^{-\lambda_k x} \quad (2)$$

and:

$$F(x; \theta) = 1 - \sum_{k=1}^m \pi_k S_k(x; \lambda_k) = 1 - \sum_{k=1}^m \pi_k e^{-\lambda_k x}, \quad (3)$$

where the parameters are $\pi = (\pi_1, \dots, \pi_m)$ and $\lambda = (\lambda_1, \dots, \lambda_m)$ with:

$$\sum_{k=1}^m \pi_k = 1, \quad 0 < \pi_k < 1, \quad (4)$$

and:

$$\lambda_k > 0, \quad k = 1, 2, \dots, m.$$

As we can see, the samples from mixed distributions are not precisely labeled by their origin sub-classes; thus, the heterogeneous dataset cannot be explicitly decomposed to homogeneous subgroups. Consequently, it brings about a barrier for estimating finite mixture models [5].

Consider a life-testing experiment where n units are placed under observation. For some reason, such as to save time and cost, we have to terminate the experiment before all items have failed. In real applications, the removal of units prior to failure is often pre-planned. Data obtained from such a type of experiments are called censored data. The most common censoring schemes are termed Type-I and Type-II censoring. In Type-I censoring, the experiment continues up to a prescribed time T . Any failures that occur after T cannot be observed. The endpoint of the experiment T is assumed to be independent of the failure times. In Type-II censoring, the experiment is terminated upon the N_{th} failure, where $N < n$ is prefixed. The experiments under these two kinds of test schemes have the drawback that they do not allow removal of samples at time points other than the termination of the experiment. However, the Type-II progressive censoring, which can be seen as a generalization of Type-II censoring, allows removing some fixed number of units at the time of the first observed $m - 1$ failures and removing all the remaining at the time of the m_{th} failure, at which time, the experiment terminates [6]. As a result, it is efficient in time and money [7] and has become very popular in the last few years [8–14].

The parameter estimation problems for different lifetime distributions, including the mixed exponential distribution and some other mixed distribution models, have been widely studied under different censoring schemes [15–17]. The most commonly adopted methods to obtain the estimations for mixture models are simply using the expectation maximization (EM) algorithm or Bayesian method. Lee and Scott [18] presented the EM algorithm for fitting multivariate Gaussian mixture models to data that are truncated, censored, or truncated and censored. In [19,20], the authors discussed the parameter estimation methods of the MED under Type-II progressively censored data and progressively hybrid Type-II censored data, respectively. Tahir et al. [21] studied the problem of estimating the parameters of a three-component mixture of exponential, Rayleigh and Burr Type-XII distributions using the Type-I right censoring scheme in the Bayesian framework. The maximum likelihood and Bayesian estimators of the parameters of a heterogeneous population represented by a finite mixture of two Pareto distributions were discussed in [22]. Feroze and Aslam [23] introduced the Bayesian approach for estimating the parameters of the two-component mixture of Weibull distribution under doubly censored samples. Feroze and Aslam [24] considered the Bayesian analysis of the three-component mixture of the Rayleigh distribution under doubly censored samples. All these works assume that we do not have any information about the label of the sub-class at all. However, in many applications, often, it is easy to get some imprecise and uncertain knowledge about the sub-class

labels. For instance, in medical surveillance databases, we can find partially labeled data provided by experts or from experience, that is, while not completely unlabeled, there is only uncertain information about class values [5]. The above-mentioned methods cannot deal with the possible available partial label information in the dataset from mixture models.

Zio [25] pointed out that uncertain information is a big challenge in reliability engineering. Specifically, there indeed exists many kinds of uncertain information in reliability analysis for censored data. First, we only know that the failure time of the censored unit belongs to the interval $[t^*, +\infty)$. Second, as mentioned, in the finite mixture models, there may be some uncertain information about the sub-class labels of the data. The theory of belief functions (also known as Dempster–Shafer theory (DST)) is appealing to represent data uncertainty. As an extension of probability theory, it has many advantages in dealing with uncertain information. Similar to the probability distribution over the discernment frame, in DST, the basic belief assignment (BBA) defined on the power set of the frame is used to represent the available information. Many scholars have studied how to measure the uncertainty of BBA [26–29]. Due to the effectiveness of DST in dealing with uncertain knowledge, it has already been widely used in many fields such as data classification/clustering [30,31], target recognition [32], decision making [33], fault diagnosis [34,35], complex networks [36,37], and so on.

In this paper, we try to handle the uncertain information in mixed distributions under Type-II progressive censoring using the theory of belief functions. Note that for the progressively censored data considered in this work, the incomplete observations are not influenced by the specific values taken by the random variables. Thus, the mechanism that causes complete failure data cases to become incompletely reported can be ignored. Under such a kind of coarsening at random (CAR) assumption, the statistical inference can proceed based on the so-called “face-value likelihood”, which measures the probability of incomplete observations by its marginal probability according to the underlying complete data distribution [38,39]. The EM algorithm is quite effective at maximizing the face-value likelihood [38], and it has been widely used for progressively censored data [7,40]. Different from the traditional estimation method using the EM algorithm, the proposed evidential parameter estimation model can take not only the uncertain censored observations, but also the prior partial information about class labels into consideration. The two kinds of uncertain information are modeled in a united frame by mass functions in belief function theory, and then, the evidential likelihood function is derived. The optimization method to get the optimal estimators, called maximal evidential likelihood estimates (MELEs), is derived based on the evidential-EM (E2M) algorithm [41]. Experimental results show that the proposed model can take advantage of the the partial information about the sub-class labels effectively and consequently improve the performance of the estimation model.

The rest of this paper is organized as follows. In Section 2, some basic concepts and the rationale of our method are briefly introduced. In Section 3, the evidential estimation model to get MELEs, which can maximize the evidential likelihood function, is presented in detail. In order to show the effectiveness of our approach, a real data application is discussed in Section 4, while some Monte Carlo simulations are conducted in Section 5. Finally, we conclude the paper in Section 6.

2. Preliminary Knowledge

Some necessary background knowledge related to this paper will be recalled in this section. The Type-II progressive censoring scheme is first introduced in Section 2.1. The basic knowledge about the theory of belief functions is then presented in Section 2.2. Finally, the evidential likelihood is recalled in Section 2.3.

2.1. Type-II Progressively Censoring Scheme

The Type-II progressive censoring scheme can be described as follows. Suppose n independent identical items are placed on a life-test. The integer $N (< n)$, which denotes the number of failures we would like to observe in the experiment, is fixed before the experiment. Let $R = (R_1, R_2, \dots, R_N)$,

where R_1, R_2, \dots, R_N are also fixed integers describing the progressive censoring scheme. They should satisfy:

$$R_1 + R_2 + \dots + R_N + N = n.$$

At the time of the first failure, say T_1 , R_1 of the remaining units are randomly removed. Similarly, at the time of the i_{th} failure, say T_i , R_i ($i = 1, 2, \dots, N$) of the remaining units are removed. At the time of the N_{th} failure, say T_N , the remaining $R_N = n - N - \sum_{i=1}^{N-1} R_i$ items are removed, and the experiment terminates. Therefore, in the presence of the Type-II progressive censoring scheme, the observed failures are $\{T_1, \dots, T_N\}$. For further details on this censoring scheme, the readers may refer to the excellent monograph of Balakrishnan and Aggarwala [6].

2.2. Theory of Belief Functions

The theory of belief functions is a mathematical theory that generalizes the theory of probabilities by giving up the additivity constraint. In this theory, justified degrees of support are assessed according to an evidential corpus, which is the set of all evidential pieces of evidence held by a source that justifies the degrees of support assigned to some subsets.

Let $\Omega = \{\omega_1, \omega_2, \dots, \omega_c\}$ be the finite domain of reference, called the discernment frame. The c elements in Ω are nonempty and mutually exclusive hypotheses related to a given problem. The belief functions are defined on the power set $2^\Omega = \{A : A \subseteq \Omega\}$. The function $m : 2^\Omega \rightarrow [0, 1]$ is said to be the basic belief assignment on 2^Ω , if it satisfies:

$$\sum_{A \subseteq \Omega} m(A) = 1. \quad (5)$$

Every $A \in 2^\Omega$ such that $m(A) > 0$ is called a focal element. The difference from probability models is that masses can be given to any subset of Ω instead of only to the atomic element of Ω . The credibility and plausibility functions are defined as in Equations (6) and (7), respectively:

$$Bel(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B), \quad \forall A \subseteq \Omega, \quad (6)$$

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B), \quad \forall A \subseteq \Omega. \quad (7)$$

Each quantity $Bel(A)$ measures the total support given to A , while $Pl(A)$ represents the potential amount of support to A . The function $pl : \Omega \rightarrow [0, 1]$ such that $pl(\omega) = Pl(\{\omega\})$ is called the contour function associated with m .

According to the type of the focal elements, we can define some particular mass functions. A categorical mass function is a normalized mass function that has a unique focal element A^* . This kind of mass function can be defined as:

$$m(A) = \begin{cases} 1 & \text{if } A = A^* \subset \Omega \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

A vacuous mass function is a particular categorical mass function focused on Ω . It is a special kind of categorical mass function with a unique focal element Ω . This type of mass function is defined as follows:

$$m(A) = \begin{cases} 1 & \text{if } A = \Omega \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

The vacuous mass function emphasizes the case of total ignorance.

A Bayesian mass function is a mass function for which all focal elements are elementary hypotheses, i.e., the focal elements are all singletons. It is defined as follows:

$$m(A) = \begin{cases} a \in [0, 1] & \text{if } |A| = 1 \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

As all focal elements are single points, this mass function is a probability distribution over frame Ω . Specifically, if a Bayesian mass function is categorical, it describes that there is no uncertainty at all, and we are completely sure about the state of the variable concerned.

If m_1 and m_2 are two independent mass functions defined on Ω , a new mass function can be formed by combining m_1 and m_2 using Dempster's rule directly:

$$(m_1 \oplus m_2)(A) = \frac{1}{1-k} \sum_{B \cap C = A} m_1(B)m_2(C), \quad (11)$$

where k is defined as

$$k = \sum_{B \cap C = \emptyset} m_1(B)m_2(C). \quad (12)$$

This describes the conflict between m_1 and m_2 . When $k = 1$, the two masses are completely in conflict, and they cannot be combined using Dempster's rule.

Suppose m_1 is a Bayesian mass function, and its corresponding contour function is a probability distribution function defined by $p_1(\omega) = m(\{\omega\})$. Assume m_2 is an arbitrary mass function with contour function $pl_2(\omega)$. The fused mass function of m_1 and m_2 by Dempster's rule yields a Bayesian mass function. Its corresponding contour function can be defined by [41]:

$$(p_1 \oplus pl_2)(\omega) = \frac{1}{1-k} p_1(\omega)pl_2(\omega), \quad (13)$$

where k is the conflict between m_1 and m_2 . It can be written as:

$$k = 1 - \sum_{\omega \in \Omega} p_1(\omega)pl_1(\omega). \quad (14)$$

The item $\sum_{\omega \in \Omega} p_1(\omega)pl_1(\omega)$ can be regarded as the mathematical expectation of pl_2 with respect to p_1 . If m_2 is categorical and such that $m_2(A) = 1$, then $p_1 \oplus pl_2$ is the probability distribution by conditioning p_1 with respect to A .

Let m^X and m^Y be two mass functions defined on Ω_X and Ω_Y , respectively, and Pl^X and Pl^Y are the associated plausibility functions. Pl^{XY} is the plausibility function defined on the product frame $\Omega_X \times \Omega_Y$. Variables X and Y are called cognitively independent if:

$$Pl^{XY}(A \times B) = Pl^X(A)Pl^Y(B), \forall A \subseteq \Omega_X, B \subseteq \Omega_Y. \quad (15)$$

If two variables are cognitively independent, the evidence on one variable does not affect the beliefs on other variables. It is clear that cognitive independence reduces to stochastic independence when m^X and m^Y are Bayesian.

2.3. Evidential Likelihood

Let X be a random vector with probability density function $p_X(\cdot; \theta)$, where θ is an unknown parameter. Assume that X is perfectly observed, and let x_0 be a realization of X . The complete likelihood function given x_0 would be defined as:

$$L(\theta; x_0) = p(x_0; \theta), \quad \forall \theta \in \Theta. \quad (16)$$

If x is not precisely observed, but we only know that $x \in A$ for $A \subseteq \Omega_X$, where Ω_X is the set from which X can take values. Then, the likelihood function given such imprecise data can be defined as [41]:

$$L(\theta; A) = \sum_{x \in A} p(x; \theta), \quad \forall \theta \in \Theta. \tag{17}$$

More generally, if the observation x is not only imprecise, but also uncertain, it can be described by a mass function m on Ω_X with focal elements A_1, \dots, A_r . The corresponding masses assigned to each focal element are $m(A_1), \dots, m(A_r)$. The likelihood function given such uncertain data can be extended as [41]:

$$L(\theta; m) = \sum_{i=1}^r m(A_i) L(\theta; A_i). \tag{18}$$

It can be rewritten as:

$$L(\theta; m) = \sum_{i=1}^r m(A_i) \sum_{x \in A_i} p(x; \theta) \tag{19a}$$

$$= \sum_{x \in \Omega_X} p(x; \theta) \sum_{A_i \ni x, A_i \subseteq \Omega_X} m(A_i) \tag{19b}$$

$$= \sum_{x \in \Omega_X} p(x; \theta) pl(x) \tag{19c}$$

$$= \mathbb{E}_\theta[pl(X)] \tag{19d}$$

$$\triangleq L(\theta; pl). \tag{19e}$$

From Equation (19c), we can see that $L(\theta; m)$ depends on m through its associated contour function pl . As a result, $L(\theta; m)$ can be denoted by $L(\theta; pl)$ indifferently. We can see that $L(\theta; pl)$ is the expectation of $pl(x)$ with respect to $p(x; \theta)$. It is often called the evidential likelihood function [42].

The evidential likelihood function $L(\theta; pl)$ can be seen as a special case of $L(\theta; A)$ and $L(\theta; x)$. If pl is associated with a categorical mass function m_A , then:

$$pl(x) = \begin{cases} 1, & x \in A \subseteq \Omega_X, \\ 0, & \text{otherwise.} \end{cases}$$

Equation (19c) equals the likelihood function given the imprecise data in Equation (17). If pl is associated with a certain mass function:

$$pl(x) = \begin{cases} 1, & X = x_0, \\ 0, & \text{otherwise.} \end{cases}$$

Equation (19c) degrades to the complete likelihood defined in Equation (16).

3. Maximum Evidential Likelihood Estimates for Uncertain Progressively Censored Data

The EM algorithm is usually adopted to obtain the maximal likelihood estimates (MLEs) for the datasets with missing information. However, it cannot deal with the partially labeled information included in the finite mixture model. In this section, we discuss the evidential likelihood for the uncertain progressively censored data from mixture distributions. The optimal statistical estimates that can be obtained by maximizing the evidential likelihood function (named MELEs) are derived, and the relation with the traditional EM model is also discussed.

3.1. Evidential Likelihood for Uncertain Progressively Censored Data

Assume that X_1, X_2, \dots, X_n are n independent variables that follow the mixture distribution in the form of Equation (1). Denote their probability density function by $f(x; \pi, \lambda)$, the probability distribution function by $F(x; \pi, \lambda)$ and the survival function by $s(x; \pi, \lambda)$. Suppose n units are placed on a life-test under the progressive censoring scheme.

In the Type-II progressive censoring scheme, we can denote the observed failure times by $T_i, i = 1, 2, \dots, N$, and use $T = (t_1, t_2, \dots, t_N)$ to describe the completely observed data. At the time of the i_{th} observed failure, we remove $R_i (i = 1, 2, \dots, N)$ items. Denote the failure times of these units by:

$$Z = \{z_{ij}, i = 1, 2, \dots, N; j = 1, 2, \dots, R_i\}.$$

It is easy to know that Z cannot be observed in the experiment.

Let $B_i^{(1)} = (b_{i1}^{(1)}, b_{i2}^{(1)}, \dots, b_{im}^{(1)})$. Each $b_{ij}^{(1)}$ in $B_i^{(1)}$ denotes whether the completely observed failure t_i comes from the j_{th} component of the mixture model. If $T_i \sim f_j(x; \theta)$, $b_{ij}^{(1)} = 1$. Otherwise, $b_{ij}^{(1)} = 0$. Similarly, we can define the indicator for the unobserved data z_{ij} . Let $B_{ij}^{(2)} = (b_{ij1}^{(2)}, b_{ij2}^{(2)}, \dots, b_{ijm}^{(2)})$, where $b_{ijk}^{(2)}$ represents whether the j_{th} removed unit at the time t_i comes from the k_{th} component in the mixture model. Let $B^{(1)} = \{B_i^{(1)}, i = 1, 2, \dots, N\}$ and $B^{(2)} = \{B_{ij}^{(2)}, i = 1, 2, \dots, N, j = 1, 2, \dots, R_i\}$.

The complete variables in this life-test can be denoted by $W = (X, B)$, where $X = (T, Z)$ and $B = (B^{(1)}, B^{(2)})$. It is clear that the observed data in this model are T , and the hidden variables are Z and B . Let $\theta = (\pi, \lambda)$, then the probability function of W under the Type-II progressive censoring scheme is:

$$p(w; \theta) = \prod_{i=1}^N \left[\prod_{k=1}^m (\pi_k f_k(t_i; \lambda_k))^{b_{ik}^{(1)}} \cdot \prod_{j=1}^{R_i} \prod_{k=1}^m (\pi_k f_k(z_{ij}; \lambda_k))^{b_{ijk}^{(2)}} \right]. \tag{20}$$

The likelihood function based on the complete data is:

$$L(\theta; w) = p(w; \theta), \tag{21}$$

As mentioned, in this life-test, the observed data are $T = \{t_i, i = 1, 2, \dots, N\}$. The likelihood function based on the observed data is:

$$L(\theta) = C \prod_{i=1}^N [f(t_i; \theta) (s(t_i; \theta))^{R_i}], \tag{22}$$

where C is a constant number. The maximum likelihood estimates can be obtained by maximizing Equation (22). The EM algorithm is usually adopted to get the MLEs for the progressively censored data [7]. As we can see, this likelihood function is based on the observed data, but it ignores some possible uncertain information of the partial labels. In fact, there are two kinds of uncertain information in this case. One is the failure time of the removed units during the experiment, for which we only know their censored time. The other is the sub-class label of each unit. Here, we use the mass functions in belief function theory to model these two types of uncertain information in a united frame. In order to establish a united frame for the data, we can describe all the data obtained in the life-test using different kinds of mass functions.

If X is a completely observed data and we assume that $X = t^*$, we can model this information with a certain mass function. Its contour function can be defined as:

$$pl_1^{(X)}(x) = \begin{cases} 1, & x = t^*, \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

If X is the censored data at time t^* , we only know that $x \in [t^*, +\infty)$. We can model this uncertain information using the following contour function:

$$pl_2^{(X)}(x) = \begin{cases} 1, & x \in [t^*, +\infty), \\ 0, & \text{otherwise.} \end{cases} \tag{24}$$

As mentioned before, in real applications, some uncertain information could be obtained from the experts or from historical data. The plausibility is adopted to express the uncertain information about the partial class labels of units. For the completely observed data t_i , the plausibility about the proposition that the unit with failure t_i comes from the k_{th} component is $pl_{t_i,k}^{(B)}$. For the censored data z_{ij} , we can use $pl_{z_{ij},k}^{(B)}$.

Suppose that the failure time X and the sub-class label B are cognitively independent. Therefore, the contour function of the complete data $w_i = (x_i, b_i)$ can be defined by:

$$pl(w_i) = \prod_{k=1}^m \left(pl_{x_i,k}^{(B)} \right)^{b_{ik}} \cdot pl^{(X)}(x_i). \tag{25}$$

where $pl_{x_i,k}^{(B)}$ is the plausibility that x_i comes from the k_{th} sub-group and $pl^{(X)}(x_i)$ describes the information about the failures. For the completely observed data and censored data, $pl^{(X)}(x_i)$ can be defined as Equations (23) and (24), respectively.

Considering the progressive censoring scheme, given the data $w = (x, b)$, the contour function of x can be defined as:

$$pl(w) = \prod_{i=1}^n pl^{(X)}(x_i) pl^{(B)}(b_i) \tag{26a}$$

$$= \prod_{i=1}^n \left[\prod_{k=1}^m \left(pl_{x_i,k}^{(B)} \right)^{b_{ik}} p^{(X)}(x_i) \right] \tag{26b}$$

$$= \prod_{i=1}^N \left[\prod_{k=1}^m \left(pl_{t_i,k}^{(B)} \right)^{b_{ik}^{(1)}} pl_1^{(X)}(t_i) \cdot \prod_{j=1}^{R_i} \prod_{k=1}^m \left(pl_{z_{ij},k}^{(B)} \right)^{b_{ijk}^{(2)}} pl_2^{(X)}(z_{ij}) \right]. \tag{26c}$$

The evidential likelihood for the progressively censored data can be derived as:

$$L(\theta; pl) = \prod_{i=1}^n \mathbb{E}_\alpha [pl(W_i)] \tag{27a}$$

$$= \prod_{i=1}^n \mathbb{E}_\alpha [pl^{(X)}(X_i) pl^{(B)}(B_i)] \tag{27b}$$

$$= \prod_{i=1}^n \sum_{k=1}^m \pi_k \mathbb{E}_\alpha [pl^{(X)}(X_i) pl^{(B)}(B_i) | b_{ik} = 1] \tag{27c}$$

$$= \prod_{i=1}^n \sum_{k=1}^m \pi_k pl_{x_i,k}^{(B)} \mathbb{E}_\alpha [pl^{(X)}(X_i) | b_{ik} = 1] \tag{27d}$$

$$= \prod_{i=1}^N \sum_{k=1}^m \pi_k pl_{t_i,k}^{(B)} \mathbb{E}_\alpha [pl_1^{(X)}(t_i) | b_{ik}^{(1)} = 1] \tag{27e}$$

$$\cdot \prod_{i=1}^N \prod_{j=1}^{R_i} \sum_{k=1}^m \pi_k pl_{z_{ij},k}^{(B)} \mathbb{E}_\alpha [pl_2^{(X)}(z_{ij}) | b_{ijk}^{(2)} = 1] \tag{27e}$$

$$= \prod_{i=1}^N \left[\sum_{k=1}^m \pi_k pl_{t_i,k}^{(B)} f_k(t_i; \lambda_k) \cdot \prod_{j=1}^{R_i} \sum_{k=1}^m \pi_k pl_{z_{ij},k}^{(B)} s_k(t_i; \lambda_k) \right], \tag{27f}$$

where $\theta = (\pi, \lambda)$ denotes the unknown parameters.

Corollary 1. *If there is no information about the sub-class label of each unit, we can use the vacuous mass function on Ω_B , i.e., $pl_{x_i,k}^{(B)} = 1$. Therefore, for the observed data and the censored data, we can get:*

$$\begin{aligned} pl_{t_i,k}^{(B)} &= 1, i = 1, 2, \dots, N; k = 1, \dots, m, \\ pl_{z_{ij},k}^{(B)} &= 1, i = 1, 2, \dots, N; j = 1, 2, \dots, R_i; k = 1, \dots, m. \end{aligned} \tag{28}$$

In this case, the evidential likelihood function defined in Equation (27f) can be rewritten as:

$$L(\theta; pl) = \prod_{i=1}^N \left[\left(\sum_{k=1}^m \pi_k f_k(t_i; \lambda_k) \right) \cdot \left(\sum_{k=1}^m \pi_k s_k(t_i; \lambda_k) \right)^{R_i} \right]. \tag{29}$$

This is the same as Equation (22).

From Corollary 1, we can see that the traditional likelihood can be regarded as a special case of the evidential likelihood for the progressively censored data when there is no prior information about the sub-class label of each unit.

3.2. The Optimal Estimates under the Evidential Likelihood

The E2M algorithm can be evoked to derive the MELEs for the uncertain progressively censored data from mixture models. Given the initial parameter value of $\theta = \theta^{(0)}$, the E-step and M-step are repeated alternatively.

- E-step: Derive the Q function, which can be seen as the expectation of $\log L(\theta; w)$ with respect to $p(w|pl; \theta^{(q)})$.

From Equation (13), the probability mass function $p(\cdot|pl; \theta^{(q)})$ can be expressed as:

$$p(w|pl; \theta^{(q)}) = \frac{p_1(w; \theta^{(q)}) p l_2(w)}{\sum_{w' \in \Omega} p_1(w'; \theta^{(q)}) p l_2(w')} = \frac{p(w; \theta^{(q)}) p l(w)}{L(\theta^{(q)}; pl)}. \tag{30}$$

The probability function $p(\cdot|pl; \theta^{(q)})$ can be seen as the Dempster combination of $p(w; \theta^{(q)})$ and $pl(w)$. The former describes the random uncertainty due to the underlying population, while the latter reflects the epistemic uncertainty due to the partial knowledge.

According to Equation (30), the Q function at the q_{th} iteration is:

$$Q(\theta, \theta^{(q)}) = \mathbb{E}[\log L(\theta; w)|pl; \theta^{(q)}] \tag{31a}$$

$$= \frac{\sum_{w \in \Omega_W} \log(L(\theta; w)) p(w; \theta^{(q)}) p l(w)}{L(\theta^{(q)}; pl)}, \tag{31b}$$

where $L(\theta; w)$ is the likelihood function based on the complete data (see Equation (21)).

- M-step: This is the same as the traditional EM algorithm. In this step, the value of θ can be obtained by maximizing $Q(\theta, \theta^{(q)})$.

Theorem 1. $L(\theta^{(q)}; pl), q = 0, 1, 2, \dots$ of the evidential likelihood function obtained by the above-mentioned E2M algorithm is not decreasing, i.e., it verifies:

$$L(\theta^{(q+1)}; pl) \geq L(\theta^{(q)}; pl). \tag{32}$$

Proof. From Equation (30), we know:

$$p(w|pl; \theta) = \frac{p(w; \theta)pl(w)}{L(\theta; pl)} = \frac{L(\theta; w)pl(w)}{L(\theta; pl)}. \tag{33}$$

From the above equation, we can get:

$$\frac{L(\theta; w)}{L(\theta; pl)} = \frac{p(w|pl; \theta)}{pl(w)} \triangleq g(w|p; \theta). \tag{34}$$

Then, we can have:

$$\log L(\theta; pl) = \log L(\theta; w) - \log g(w|pl; \theta). \tag{35}$$

Taking the expectation of both sides of the above equation with respect to $p(w|pl; \theta^{(q)})$, then it is easy to get:

$$\log L(\theta; pl) = \mathbb{E} \left[\log L(\theta; w) | pl; \theta^{(q)} \right] - \mathbb{E} \left[\log g(x|pl; \theta) | pl; \theta^{(q)} \right]. \tag{36}$$

Let

$$H(\theta, \theta^{(q)}) = \mathbb{E} \left[\log g(x|pl; \theta) | pl; \theta^{(q)} \right], \tag{37}$$

and then, we have:

$$\begin{aligned} & \log L(\theta^{(q+1)}; pl) - \log L(\theta^{(q)}; pl) \\ &= \left(Q(\theta^{(q+1)}, \theta^{(q)}) - Q(\theta^{(q)}, \theta^{(q)}) \right) - \left(H(\theta^{(q+1)}, \theta^{(q)}) - H(\theta^{(q)}, \theta^{(q)}) \right). \end{aligned} \tag{38}$$

It is easy to see that $Q(\theta^{(q+1)}, \theta^{(q)}) \geq Q(\theta^{(q)}, \theta^{(q)})$ since in the M-step, the value of θ is updated by maximizing $Q(\theta, \theta^{(q)})$ with respect to θ . For any θ ,

$$\begin{aligned} & H(\theta, \theta^{(q)}) - H(\theta^{(q)}, \theta^{(q)}) \\ &= \mathbb{E} \left[\log \frac{g(w|pl; \theta)}{g(w|pl; \theta^{(q)})} \middle| pl; \theta^{(q)} \right] \end{aligned} \tag{39a}$$

$$\leq \log \mathbb{E} \left[\frac{g(w|pl; \theta)}{g(w|pl; \theta^{(q)})} \middle| pl; \theta^{(q)} \right] \tag{39b}$$

$$\leq \log \int \frac{g(w|pl; \theta)}{g(w|pl; \theta^{(q)})} pl(w) g(w|pl; \theta^{(q)}) dw \tag{39c}$$

$$\leq \log \int g(w|pl; \theta) pl(w) dw \tag{39d}$$

$$\leq \log \int p(w|pl; \theta) dw = 0, \tag{39e}$$

where the inequality (39b) is obtained from Jensen’s inequality.

As we can see, the first difference in Equation (38) is nonnegative, while the second one is no larger than zero. Thus, it is easy to get $\log L(\theta^{(q+1)}; pl) \geq \log L(\theta^{(q)}; pl)$, and consequently, we have the inequality (32). □

Corollary 2. When the contour function $pl(w)$ degrades to the categorical mass function, the probability function $p(w|pl; \theta^{(q)})$ degrades to the conditional probability function $P(\cdot | y, \theta^{(q)})$ in EM.

3.3. The Estimation Methodology for a Mixed Exponential Distribution

The likelihood function based on the complete data w for Type-II progressively censored data from MED is:

$$L(\pi, \lambda; w) = \prod_{i=1}^N \left[\prod_{k=1}^m \left(\pi_k \lambda_k e^{-\lambda_k t_i} \right)^{b_{ik}^{(1)}} \cdot \prod_{j=1}^{R_i} \prod_{k=1}^m \left(\pi_k \lambda_k e^{-\lambda_k z_{ij}} \right)^{b_{ijk}^{(2)}} \right]. \tag{40}$$

The log-likelihood is:

$$\log L(\pi, \lambda; w) = \sum_{i=1}^N \left[\sum_{k=1}^m b_{ik}^{(1)} (\log \pi_k + \log \lambda_k - \lambda_k t_i) + \sum_{j=1}^{R_i} \sum_{k=1}^m b_{ijk}^{(2)} (\log \pi_k + \log \lambda_k - \lambda_k z_{ij}) \right]. \tag{41}$$

As before, let $\theta = (\pi, \lambda)$ denote the unknown parameters of the MED model. Given the initial parameter $\theta^{(0)} = (\pi^{(0)}, \lambda^{(0)})$, the alternative iterations of the E2M algorithm are described in the following.

- E-step: Let the parameters at the q_{th} iteration be $\theta^{(q)} = (\pi^{(q)}, \lambda^{(q)})$. The expectation of $\log L(\pi, \lambda; w)$ with respect to $p(w|pl; \theta^{(q)})$ (Q function) can be defined as:

$$Q(\theta, \theta^{(q)}) = \mathbb{E}[\log L(\pi, \lambda; w) | pl; \theta^{(q)}] \tag{42a}$$

$$= \sum_{i=1}^N \left\{ \sum_{k=1}^m (\log \pi_k + \log \lambda_k - \lambda_k t_i) \mathbb{E}[b_{ik}^{(1)} | pl; \theta^{(q)}] + \sum_{j=1}^{R_i} \sum_{k=1}^m [(\log \pi_k + \log \lambda_k) \mathbb{E}[b_{ijk}^{(2)} | pl; \theta^{(q)}] - \lambda_k \mathbb{E}[b_{ijk}^{(2)} z_{ij} | pl; \theta^{(q)}]] \right\}. \tag{42b}$$

From Equation (13), we know:

$$p(x_i, b_{ik}^{(1)} = 1 | pl; \theta^{(q)}) = \frac{\pi_k^{(q)} f_k(t_i; \lambda_k^{(q)}) pl_{t_i,k}^{(B)} pl^{(X)}(t_i)}{\sum_{l=1}^m \pi_l^{(q)} \int_{x_i} f_l(x_i; \lambda_l^{(q)}) pl_{x_i,l}^{(B)} pl^{(X)}(x_i) dx_i} \tag{43a}$$

$$= \frac{\pi_k^{(q)} f_k(t_i; \lambda_k^{(q)}) pl_{t_i,k}^{(B)} pl^{(X)}(t_i)}{\sum_{l=1}^m \pi_l^{(q)} f_l(t_i; \lambda_l^{(q)}) pl_{t_i,l}^{(B)}}. \tag{43b}$$

As $b_{ik}^{(1)}$ is a Boolean variable, it is easy to get:

$$\mathbb{E}[b_{ik}^{(1)} | pl; \theta^{(q)}] = P(b_{ik}^{(1)} = 1 | pl; \theta^{(q)}) = \frac{\pi_k^{(q)} f_k(t_i; \lambda_k^{(q)}) pl_{t_i,k}^{(B)}}{\sum_{l=1}^m \pi_l^{(q)} f_l(t_i; \lambda_l^{(q)}) pl_{t_i,l}^{(B)}} \triangleq a_{ik}^{(q)}. \tag{44}$$

Similarity, we can get:

$$p(z_{ij}, b_{ijk}^{(2)} = 1 | pl; \theta^{(q)}) = \frac{\pi_k^{(q)} f_k(z_{ij}; \lambda_k^{(q)}) pl_{z_{ij},k}^{(B)} pl^{(X)}(z_{ij})}{\sum_{l=1}^m \pi_l^{(q)} s_l(z_{ij}; \lambda_l^{(q)}) pl_{z_{ij},l}^{(B)}}. \tag{45}$$

$$\mathbb{E}[b_{ijk}^{(2)} | pl; \theta^{(q)}] = P(b_{ijk}^{(2)} = 1 | pl; \theta^{(q)}) = \frac{\pi_k^{(q)} s_k(z_{ij}; \lambda_k^{(q)}) pl_{z_{ij},k}^{(B)}}{\sum_{l=1}^m \pi_l^{(q)} s_l(z_{ij}; \lambda_l^{(q)}) pl_{z_{ij},l}^{(B)}} \triangleq e_{ijk}^{(q)}. \tag{46}$$

For the expectation of $b_{ijk}^{(2)} z_{ij}$, we can get:

$$\mathbb{E}[b_{ijk}^{(2)} z_{ij} | pl; \theta^{(q)}] = \mathbb{E}[z_{ij} | b_{ijk}^{(2)} = 1, pl; \theta^{(q)}] P(b_{ijk}^{(2)} = 1 | pl; \theta^{(q)}), \tag{47}$$

where $P(b_{ijk}^{(2)} = 1 | pl; \theta^{(q)})$ can be obtained by Equation (46). From Equations (45) and (46), we can get:

$$p(z_{ij}|b_{ijk}^{(2)} = 1, pl; \theta^{(q)}) = \frac{p(z_{ij}, b_{ijk}^{(2)} = 1|pl; \theta^{(q)})}{P(b_{ijk}^{(2)} = 1|pl; \theta^{(q)})} = \frac{f_k(z_{ij}; \lambda_k^{(q)})p_l^{(X)}(z_{ij})}{s_k(t_i; \lambda_k^{(q)})}. \tag{48}$$

As $z_{ij} \in [t_i, +\infty)$, we can get:

$$\mathbb{E}[z_{ij}|b_{ijk}^{(2)} = 1, pl; \theta^{(q)}] = \int_{t_i}^{\infty} z_{ij} \frac{f_k(z_{ij}; \lambda_k^{(q)})}{s_k(t_i; \lambda_k^{(q)})} dz_{ij} \triangleq \xi_{ik}^{(q)}. \tag{49}$$

Therefore,

$$\mathbb{E}[b_{ijk}^{(2)} z_{ij}|pl; \theta^{(q)}] = \xi_{ik}^{(q)} e_{ijk}^{(q)}. \tag{50}$$

Substituting Equations (44), (46), and (50) into Equation (42b), we can get the Q function in the q th iteration:

$$Q(\theta, \theta^{(q)}) = \sum_{i=1}^N \left[\sum_{k=1}^m (\log \pi_k + \log \lambda_k - \lambda_k t_i) a_{ik}^{(q)} + \sum_{j=1}^{R_i} \sum_{k=1}^m [(\log \pi_k + \log \lambda_k) e_{ijk}^{(q)} - \lambda_k \xi_{ik}^{(q)} e_{ijk}^{(q)}] \right], \tag{51}$$

where $\theta = (\pi, \lambda)$, $\pi = (\pi_1, \dots, \pi_m)$, $\lambda = (\lambda_1, \dots, \lambda_m)$, and $\sum_{k=1}^m \pi_k = 1$.

Corollary 3. *If there is no uncertain information in the experiment, i.e., $pl_{x_i,k} = 1$, the conditional expectations defined in Equations (44), (46), and (50) are the same as those in the conventional EM algorithm:*

$$\begin{aligned} \mathbb{E}[b_{ik}^{(1)}|pl; \theta^{(q)}] &= \frac{\pi_k^{(q)} f_k(t_i; \lambda_k^{(q)})}{\sum_{l=1}^m \pi_l^{(q)} f_l(t_i; \lambda_l^{(q)})}, \\ \mathbb{E}[b_{ijk}^{(2)}|pl; \theta^{(q)}] &= \frac{\pi_k^{(q)} s_k(t_i; \lambda_k^{(q)})}{\sum_{l=1}^m \pi_l^{(q)} s_l(t_i; \lambda_l^{(q)})}, \\ \mathbb{E}[b_{ijk}^{(2)} z_{ij}|pl; \theta^{(q)}] &= \frac{\pi_k^{(q)} \int_{t_i}^{\infty} z_{ij} f_k(z_{ij}; \lambda_k^{(q)}) dz_{ij}}{\sum_{l=1}^m \pi_l^{(q)} s_l(t_i; \lambda_l^{(q)})}. \end{aligned}$$

- M-step: In order to get the updated estimates, we have to maximize the $Q(\theta, \theta^{(q)})$ function with respect to parameters π and λ . The Lagrangian method can be adopted to consider the constraint $\sum_{k=1}^m \pi_k = 1$. Let:

$$L(\theta, \eta) = Q(\theta, \theta^{(q)}) + \eta \left(\sum_{k=1}^m \pi_k - 1 \right).$$

We can get:

$$\frac{\partial L(\theta, \eta)}{\partial \pi_k} = \frac{A_k^{(q)}}{\pi_k} + \eta = 0, \tag{52}$$

$$\frac{\partial L(\theta, \eta)}{\partial \lambda_k} = \frac{A_k^{(q)}}{\lambda_k} - C_k^{(q)} - \sum_{i=1}^N t_i a_{ik}^{(q)} = 0, \tag{53}$$

$$\frac{\partial L(\theta, \eta)}{\partial \eta} = \sum_{k=1}^m \pi_k - 1 = 0, \tag{54}$$

where

$$A_k^{(q)} = \sum_{i=1}^N a_{ik}^{(q)} + \sum_{i=1}^N \sum_{j=1}^{R_i} e_{ijk}^{(q)}, \tag{55}$$

$$C_k^{(q)} = \sum_{i=1}^N \sum_{j=1}^{R_i} \tilde{c}_{ik}^{(q)} e_{ijk}^{(q)}. \tag{56}$$

From Equations (52) and (54), we know:

$$\eta = - \sum_{k=1}^m A_k^{(q)}.$$

Thus, it is easy to get:

$$\pi_k^{(q+1)} = \frac{A_k^{(q)}}{\sum_{k=1}^m A_k^{(q)}}, k = 1, \dots, m, \tag{57}$$

$$\lambda_k^{(q+1)} = \frac{A_k^{(q)}}{C_k^{(q)} + \sum_{i=1}^N t_i a_{ik}^{(q)}}, k = 1, \dots, m. \tag{58}$$

where $A_k^{(q)}$ and $C_k^{(q)}$ are defined in Equations (55) and (56), respectively.

The E-step and M-step will be repeated iteratively until:

$$\sqrt{\sum_{i=1}^m (\pi_i^{(q+1)} - \pi_i^{(q)})^2} \leq \epsilon \tag{59}$$

and:

$$\sqrt{\sum_{i=1}^m (\lambda_i^{(q+1)} - \lambda_i^{(q)})^2} \leq \epsilon, \tag{60}$$

where ϵ is a given small positive constant. Suppose that the algorithm stops at the l_{th} step, then $\theta^{(l)} = (\pi^{(l)}, \lambda^{(l)})$ are the MELEs for the parameters. The whole process of the estimation is summarized in Algorithm 1.

Algorithm 1 : The evidential estimation approach for uncertain MED under progressive censoring.

Input: The observed failures (t_1, \dots, t_N) and the censoring scheme (R_1, \dots, R_N) .

Initialization:

- (1) Let $q = 0$.
- (2) Initialize the parameters $\theta^{(q)} = (\pi^{(q)}, \lambda^{(q)})$.

repeat

- (1) E-step: Calculate the Q function based on the current parameters $Q(\theta, \theta^{(q)})$.
- (2) M-step: Maximize $Q(\theta, \theta^{(q)})$ with respect to θ and update the parameters $\theta^{(q+1)}$.
- (3) $q = q + 1$.

until Both Equations (59) and (60) hold true.

Output: The MELEs of the parameters.

4. Data Analysis

In this section, a real dataset of the air conditioning system of an airplane is considered to illustrate the practical application of the proposed evidential estimation model. The group of failure times of the system is given as follows:

1, 3, 5, 7, 11, 11, 11, 12, 14, 14, 14, 16, 16, 20, 21, 23, 42, 47, 52, 62, 71, 71, 87, 90, 95, 120, 120, 225, 246, 261.

This dataset was analyzed by different authors using various models [43,44]. In [19], the authors found that the MED model with two components ($m = 2$) provided a good fit to the given failure data. The fitted probability density function can be obtained by the maximal likelihood estimations:

$$\hat{f}(x) = \hat{\pi}_1 \hat{\lambda}_1 e^{-\hat{\lambda}_1 x} + \hat{\pi}_2 \hat{\lambda}_2 e^{-\hat{\lambda}_2 x}, x \geq 0,$$

with $\hat{\pi}_1 = 0.3459, \hat{\pi}_2 = 0.6541, \hat{\lambda}_1 = 0.0647, \hat{\lambda}_2 = 0.0121$. If we do not have any information about the sub-class labels of each sample, the MLEs of the above MED can be obtained using the EM algorithm [19]:

$$\hat{\pi}_1^{MLE} = 0.1136, \hat{\pi}_2^{MLE} = 0.8864, \hat{\lambda}_1^{MLE} = 0.1091, \hat{\lambda}_2^{MLE} = 0.0161.$$

Suppose we have some uncertain information about the sub-class labels. This type of knowledge can be obtained by experts or from historical data. Here, we can use the evidential *c*-means (ECM) clustering algorithm [31] to obtain the contour function for each sample. From the given failures, it is easy to see that they can be divided into two groups. The credal partition found by ECM can provide us with the plausibility of the data belonging to the sub-groups. In ECM, let $C = 2$, and the other parameters are set as the default. The plausibility matrix describing the partial information about the sub-class label of each sample is illustrated in Table 1.

The maximum likelihood estimations in the sense of evidential likelihood (denoted by MELE) that can integrate the uncertain label information can be obtained:

$$\hat{\pi}_1^{MELE} = 0.3433, \hat{\pi}_2^{MELE} = 0.6567, \hat{\lambda}_1^{MELE} = 0.0651, \hat{\lambda}_2^{MELE} = 0.0121.$$

We can see that the estimation results by the proposed estimation model are closer to the estimations by the complete data.

Table 1. The plausibility matrix for the data.

ID	$pl_{x_i,1}^{(B)}$	$pl_{x_i,2}^{(B)}$	ID	$pl_{x_i,1}^{(B)}$	$pl_{x_i,2}^{(B)}$
1	0.9971	0.0080	16	0.9976	0.0076
2	0.9980	0.0057	17	0.9775	0.0858
3	0.9987	0.0038	18	0.9687	0.1264
4	0.9992	0.0022	19	0.9590	0.1770
5	0.9999	0.0003	20	0.9388	0.3094
6	0.9999	0.0003	21	0.9247	0.4584
7	0.9999	0.0003	22	0.9247	0.4584
8	1.0000	0.0001	23	0.9266	0.7394
9	1.0000	0.0001	24	0.9319	0.7860
10	1.0000	0.0001	25	0.9435	0.8549
11	1.0000	0.0001	26	1.0000	1.0000
12	0.9998	0.0006	27	1.0000	1.0000
13	0.9998	0.0006	28	0.0001	1.0000
14	0.9989	0.0035	29	0.0189	0.9930
15	0.9985	0.0047	30	0.0474	0.9813

As we know, in the traditional EM algorithm, the solution depends on the initial values adopted. In order to show the influences of the initial values for the proposed estimation model, we will try with different initialization strategies. For parameter π , we evoke the estimation model with:

$$\pi_1^{(0)} = |0.5 + N(0, 0.1)|, \pi_2^{(0)} = 1 - \pi_1^{(0)},$$

while for λ , we set:

$$\lambda_1^{(0)} = |0.2 + N(0, 0.1)|, \lambda_2^{(0)} = |0.8 + N(0, 0.1)|,$$

where $|x|$ is the absolute value of x and $N(0, 0.1)$ denotes a random number from the normal distribution with $\mu = 0, \sigma^2 = 0.1$. The E2M algorithm and EM algorithm are both repeated 100 times. The mean value and the standard variances of the estimations are reported, which is illustrated in Table 2.

Table 2. The mean and standard variance (sd) values of the estimations with different initializations. MELE, maximal evidential likelihood estimate.

Parameter	MLE		MELE	
	Mean	sd	Mean	sd
λ_1	0.0898	0.0101	0.0598	0.0083
λ_2	0.0177	0.0305	0.0277	0.0045
π_1	0.1931	0.1094	0.0931	0.0182
π_2	0.8069	0.1094	0.4069	0.0154

It is easy to see that the standard variation values of MELEs are smaller than those of MLEs. We can conclude that the use of the uncertain information can reduce the influence of the initializations in the estimation model. Many initialization methods have been studied for the EM algorithm, which is worth learning and a reference for the proposed estimation model. However, in this paper, we do not focus on the proper initialization method and leave it for our future research.

5. Simulations

Monte Carlo simulations are conducted in this section to show the performance of the proposed method. Suppose that X_1, X_2, \dots, X_n are n independent variables from the mixed exponential distribution in Equation (2). Here, we consider the model with two mixture components ($m = 2$) under the Type-II progressive censoring scheme. The parameters to be estimated are $\pi_1, \pi_2, \lambda_1, \lambda_2$ with $\pi_1 + \pi_2 = 1$.

To simulate uncertainty on class labels, the original generated data are processed as follows. For the i_{th} observed failure, an error probability q_i is drawn randomly from a beta distribution with mean ρ and standard deviation σ . In real applications, q_i can be obtained by experts or from historical data. With probability q_i , the label of z_i is replaced by a completely value of \tilde{z}_i (with a uniform distribution over all the class labels). The contour function on class labels can be defined as:

$$pl_{x_i,k}^{(B)} = P(b_{ik} = 1 | b_{ik}^*) = \begin{cases} \frac{q_i}{m}, & b_{ik}^* = 0, \\ \frac{q_i}{m} + 1 - q_i, & b_{ik}^* = 1, \end{cases} \quad i = 1, \dots, n; k = 1, \dots, m. \quad (61)$$

If there is no uncertainty, i.e., $pl_{x_i,k} = 1$, the evidential likelihood degrades to the traditional likelihood. Let $\pi_1 = 0.3, \pi_2 = 0.7; \lambda_1 = 0.6, \lambda_2 = 0.1$. Set the initial value of the parameters as:

$$\pi_1^{(0)} = 0.5, \pi_2^{(0)} = 0.5, \lambda_1^{(0)} = 1, \lambda_2^{(0)} = 0.5.$$

In this experiment, we set the number of units placed on the life-test to be $n = 20, 100, 200, 400$. We consider different censoring proportions for a given sample size. Let $N = n \times (0.1, 0.2, 0.3)$. For the

parameters to model the uncertain label information, we set $\rho = 0.2, \sigma = 0.2$. Three censoring schemes are considered:

- Scheme 1: $R_1 = \dots, R_{N-1} = 0, R_N = n - N$.
- Scheme 2: $R_1 = n - N, R_2 = \dots, R_N = 0$.
- Scheme 3: $R_1 = R_2 = \dots = R_{N-1} = 1, R_N = n - 2 * N + 1$.

Given the sample size n and the censoring scheme, the E2M algorithm will be repeated $s = 100$ times. The bias and mean squared error (MSE) values of the estimators are reported. The results are illustrated in Tables 3–8.

From the tables, we can see that, with the increasing size of n and N , for the estimators obtained both by the proposed evidential estimation model and by the EM algorithm, the values of the bias and MSE decrease in most cases. This is consistent with our common sense about the maximum likelihood estimations. For different censoring schemes, it is obvious that the estimation efficiency is highly related to the proportion of the censored samples. When the number of units placed in the life-test is fixed, with the increasing number of censored data, the values of MSE increase.

Overall, in most cases, the biases and MSEs of the MELEs by the proposed evidential estimation model are smaller than those of MLEs by the EM algorithm, which can be attributed to the use of the uncertain prior information about the sub-class labels. When the proportion of the censored data is large, the MELEs obtained by the proposed estimation model are significantly better.

Table 3. The statistical properties of the estimators of λ_1 and λ_2 with Scheme 1. $\hat{\lambda}_i^{MELE}$ and $\hat{\lambda}_i^{MLE}$ are the MELE by the proposed estimation method and the MLE by the traditional EM algorithm, respectively.

n	N	Bias				MSE			
		$\hat{\lambda}_1^{MLE}$	$\hat{\lambda}_1^{MELE}$	$\hat{\lambda}_2^{MLE}$	$\hat{\lambda}_2^{MELE}$	$\hat{\lambda}_1^{MLE}$	$\hat{\lambda}_1^{MELE}$	$\hat{\lambda}_2^{MLE}$	$\hat{\lambda}_2^{MELE}$
20	2	0.3940	0.2279	0.3193	0.1421	5.4890	3.6382	0.0990	0.0808
20	4	0.9357	0.7783	0.2508	0.1373	4.4376	2.6257	0.0933	0.0764
20	6	0.1281	1.0112	0.1934	0.1189	5.9467	2.6358	0.0858	0.0683
100	10	0.7666	0.2661	0.2825	0.1249	3.1009	0.0790	0.0970	0.0642
100	20	0.4059	0.2158	0.2504	0.1509	1.6919	0.3333	0.0823	0.0652
100	30	0.2601	0.1422	0.1860	0.1427	0.6354	0.1728	0.0865	0.0657
200	20	0.6175	0.2911	0.2775	0.1167	7.9779	0.7749	0.1008	0.0645
200	40	0.2238	0.2470	0.2451	0.1639	1.0183	0.5144	0.0747	0.0604
200	60	0.4149	0.1986	0.1992	0.1454	3.9797	0.3526	0.0810	0.0665
400	40	0.9731	0.3065	0.2739	0.1268	7.4599	0.0686	0.0936	0.0635
400	80	0.7262	0.3163	0.2461	0.1242	9.7900	0.0716	0.0820	0.0540
400	120	0.2421	0.1978	0.2037	0.1219	0.3790	0.2534	0.0768	0.0455

Table 4. The statistical properties of the estimators of π_1 and π_2 with Scheme 1. $\hat{\pi}_i^{MELE}$ and $\hat{\pi}_i^{MLE}$ are the MELE by the proposed estimation method and the MLE by the traditional EM algorithm, respectively.

n	N	Bias				MSE			
		$\hat{\pi}_1^{MLE}$	$\hat{\pi}_1^{MELE}$	$\hat{\pi}_2^{MLE}$	$\hat{\pi}_2^{MELE}$	$\hat{\pi}_1^{MLE}$	$\hat{\pi}_1^{MELE}$	$\hat{\pi}_2^{MLE}$	$\hat{\pi}_2^{MELE}$
20	2	0.0508	0.0892	0.0579	0.1193	0.3963	0.2967	0.1314	0.0596
20	4	0.1200	0.0392	0.0627	0.0508	0.1669	0.1140	0.1275	0.0752
20	6	0.1178	0.0739	0.0811	0.0066	0.1644	0.1364	0.1272	0.0869
100	10	0.1316	0.0589	0.0751	0.0825	0.1578	0.1116	0.1339	0.0546
100	20	0.1471	0.0810	0.0491	0.0504	0.1643	0.1215	0.1122	0.0649
100	30	0.1695	0.1146	0.0573	0.0140	0.1795	0.1416	0.1180	0.0861
200	20	0.1420	0.0718	0.0833	0.0775	0.1596	0.1150	0.1402	0.0561
200	40	0.1503	0.0917	0.0361	0.0451	0.1674	0.1281	0.1014	0.0626
200	60	0.1744	0.1190	0.0546	0.0008	0.1820	0.1430	0.1121	0.0819
400	40	0.1529	0.0801	0.0732	0.0739	0.1685	0.1189	0.1304	0.0563
400	80	0.1629	0.0946	0.0558	0.0461	0.1751	0.1260	0.1140	0.0649
400	120	0.1799	0.1170	0.0581	0.0037	0.1860	0.1405	0.1097	0.0794

Table 5. The statistical properties of the estimators of λ_1 and λ_2 with Scheme 2. $\hat{\lambda}_i^{MELE}$ and $\hat{\lambda}_i^{MLE}$ are the MELE by the proposed estimation method and the MLE by the traditional EM algorithm, respectively.

<i>n</i>	<i>N</i>	Bias				MSE			
		$\hat{\lambda}_1^{MELE}$	$\hat{\lambda}_1^{MLE}$	$\hat{\lambda}_2^{MELE}$	$\hat{\lambda}_2^{MLE}$	$\hat{\lambda}_1^{MELE}$	$\hat{\lambda}_1^{MLE}$	$\hat{\lambda}_2^{MELE}$	$\hat{\lambda}_2^{MLE}$
20	2	9.1642	10.1750	0.1960	0.1354	4.8531	9.1105	0.0942	0.1050
20	4	12.2537	13.2526	0.2083	0.1614	5.6519	5.0538	0.0905	0.0959
20	6	4.0073	4.3486	0.2546	0.1650	1.3179	4.2556	0.1318	0.0993
100	10	12.9948	9.4179	0.1790	0.1272	7.8862	9.7824	0.0925	0.1078
100	20	1.2818	1.3895	0.1991	0.1259	11.6120	3.5432	0.1077	0.1003
100	30	0.4817	0.7400	0.2286	0.1477	1.9537	5.4708	0.1082	0.0890
200	20	1.9440	0.9701	0.2530	0.1425	2.9915	1.5640	0.1163	0.1022
200	40	0.3792	0.5135	0.2234	0.1697	1.6746	4.1021	0.0834	0.0781
200	60	0.2999	1.2155	0.2286	0.1742	1.1539	1.5056	0.0962	0.0792
400	40	2.6339	1.3680	0.1942	0.1198	1.1806	1.7962	0.0848	0.0982
400	80	4.7986	0.6004	0.2031	0.1395	1.3783	0.5436	0.0840	0.0810
400	120	0.2702	0.0976	0.2199	0.1410	0.6702	0.4962	0.0924	0.0876

Table 6. The statistical properties of the estimators of π_1 and π_2 with Scheme 2. $\hat{\pi}_i^{MELE}$ and $\hat{\pi}_i^{MLE}$ are the MELE by the proposed estimation method and the MLE by the traditional EM algorithm, respectively.

<i>n</i>	<i>N</i>	Bias				MSE			
		$\hat{\pi}_1^{MELE}$	$\hat{\pi}_1^{MLE}$	$\hat{\pi}_2^{MELE}$	$\hat{\pi}_2^{MLE}$	$\hat{\pi}_1^{MELE}$	$\hat{\pi}_1^{MLE}$	$\hat{\pi}_2^{MELE}$	$\hat{\pi}_2^{MLE}$
20	2	0.0308	0.0205	0.0646	0.0040	0.2471	0.2576	0.1382	0.1103
20	4	0.0765	0.0688	0.0386	0.0083	0.2212	0.1841	0.1215	0.1036
20	6	0.1701	0.1613	0.0350	0.0546	0.1786	0.1867	0.1235	0.1256
100	10	0.1830	0.1791	0.0728	0.0210	0.1831	0.1849	0.1425	0.1148
100	20	0.1936	0.1878	0.0741	0.0009	0.1916	0.1972	0.1369	0.1215
100	30	0.1968	0.1897	0.0523	0.0286	0.1915	0.1977	0.1186	0.1129
200	20	0.1976	0.1832	0.0575	0.0530	0.1893	0.1999	0.1330	0.1140
200	40	0.1974	0.1926	0.0303	0.0234	0.1941	0.1976	0.1028	0.0929
200	60	0.1979	0.1934	0.0258	0.0286	0.1943	0.1986	0.1030	0.1023
400	40	0.2011	0.1929	0.0802	0.0058	0.1945	0.2008	0.1365	0.1017
400	80	0.1965	0.1905	0.0605	0.0031	0.1926	0.1978	0.1149	0.0983
400	120	0.2001	0.1931	0.0590	0.0199	0.1956	0.2020	0.1195	0.0981

Table 7. The statistical properties of the estimators of λ_1 and λ_2 with Scheme 3. $\hat{\lambda}_i^{MELE}$ and $\hat{\lambda}_i^{MLE}$ are the MELE by the proposed estimation method and the MLE by the traditional EM algorithm, respectively.

<i>n</i>	<i>N</i>	Bias				MSE			
		$\hat{\lambda}_1^{MELE}$	$\hat{\lambda}_1^{MLE}$	$\hat{\lambda}_2^{MELE}$	$\hat{\lambda}_2^{MLE}$	$\hat{\lambda}_1^{MELE}$	$\hat{\lambda}_1^{MLE}$	$\hat{\lambda}_2^{MELE}$	$\hat{\lambda}_2^{MLE}$
20	2	7.9040	0.1744	0.1278	0.1974	8.9819	0.2180	0.1060	0.0689
20	4	3.0177	5.2017	0.1306	0.1556	7.6359	1.3141	0.1031	0.0694
20	6	1.4218	0.5737	0.1603	0.1936	3.1824	1.1476	0.0872	0.0804
100	10	1.7309	0.0356	0.1230	0.1750	4.4582	1.5949	0.0954	0.0651
100	20	0.9920	0.1063	0.1410	0.1343	3.2922	1.3056	0.0891	0.0678
100	30	0.5160	0.4788	0.1254	0.1470	1.6549	2.8739	0.0843	0.0732
200	20	0.8511	0.2909	0.1285	0.1770	4.9667	0.0717	0.0902	0.0638
200	40	0.4303	0.3190	0.1322	0.1393	1.5101	0.0742	0.0873	0.0649
200	60	0.6432	0.4140	0.1272	0.1505	4.1079	2.7520	0.0851	0.0778
400	40	0.4359	0.3001	0.1124	0.1753	1.2242	0.0669	0.0971	0.0630
400	80	0.0547	0.3164	0.1377	0.0423	0.9854	0.1682	0.0896	0.0628
400	120	0.0770	0.0161	0.1391	0.0752	0.9764	0.0856	0.0829	0.0743

Table 8. The statistical properties of the estimators of π_1 and π_2 with Scheme 3. $\hat{\pi}_i^{\text{MELE}}$ and $\hat{\pi}_i^{\text{MLE}}$ are the MELE by the proposed estimation method and the MLE by the traditional EM algorithm, respectively.

n	N	Bias				MSE			
		$\hat{\pi}_1^{\text{MLE}}$	$\hat{\pi}_1^{\text{MELE}}$	$\hat{\pi}_2^{\text{MLE}}$	$\hat{\pi}_2^{\text{MELE}}$	$\hat{\pi}_1^{\text{MLE}}$	$\hat{\pi}_1^{\text{MELE}}$	$\hat{\pi}_2^{\text{MLE}}$	$\hat{\pi}_2^{\text{MELE}}$
20	2	0.0438	0.0149	0.0722	0.0974	0.2629	0.2063	0.1415	0.0554
20	4	0.0771	0.0028	0.0694	0.0556	0.1643	0.1263	0.1377	0.0686
20	6	0.1432	0.0668	0.0397	0.0064	0.1712	0.1335	0.1140	0.0967
100	10	0.1445	0.0685	0.0770	0.0750	0.1617	0.1120	0.1335	0.0576
100	20	0.1526	0.0947	0.0590	0.0343	0.1692	0.1294	0.1214	0.0722
100	30	0.1806	0.1440	0.0746	0.0530	0.1861	0.1644	0.1221	0.1055
200	20	0.1463	0.0718	0.0715	0.0770	0.1624	0.1153	0.1273	0.0557
200	40	0.1685	0.1000	0.0678	0.0393	0.1793	0.1290	0.1225	0.0677
200	60	0.1771	0.1463	0.0728	0.0495	0.1834	0.1615	0.1220	0.1078
400	40	0.1511	0.0763	0.0876	0.0753	0.1675	0.1182	0.1384	0.0555
400	80	0.1644	0.0977	0.0623	0.0423	0.1782	0.1289	0.1223	0.0652
400	120	0.1718	0.1415	0.0609	0.0248	0.1794	0.1557	0.1156	0.0961

6. Conclusions

In this paper, a method for estimating the parameters of a mixed distribution model under Type-II progressive censoring was introduced. The main advantage of the proposed formalism is that it can combine uncertainty captured by the imperfect observation process due to censoring and the partial knowledge about the sub-class information in the mixture model. These two types of uncertain information can be represented in a united frame by the use of belief functions, and then, the evidential likelihood is derived, which can be seen as a generalized likelihood criterion. It can be maximized using the evidential EM algorithm. As an illustration, the method was applied to the MED model. A real dataset was considered, and some Monte Carlo simulations were carried out to demonstrate the performance of the resulting estimates. The results show that the estimations by the proposed method gain better estimation efficiency, which can be attributed to the effective integration of the uncertain information in the model.

More generally, the evidential estimation approach introduced in this paper is applicable to any other mixed distributions under different censoring schemes, where data uncertainty in the parametric statistical model arises from the imperfect observation process and reliability engineering applications such as the life data with multiple failure causes.

Author Contributions: The authors contributed equally to the presented mathematical framework, the programming and the writing of the paper. Both authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (No.61701409), the Natural Science Basic Research Plan in Shaanxi Province of China (No.2018JQ6005), the Aero Science Foundation of China (No.20182053023), and the Fundamental Research Funds for the Central Universities of China (No.310201911cx041).

Acknowledgments: The authors would like to express deep thanks to the editors and referees for their helpful comments and suggestions which significantly improved the paper.

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

References

1. Figueiredo, M.A.T.; Jain, A.K. Unsupervised learning of finite mixture models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 381–396. [[CrossRef](#)]
2. Xiang, S.; Yao, W.; Yang, G. An overview of semiparametric extensions of finite mixture models. *Stat. Sci.* **2019**, *34*, 391–404. [[CrossRef](#)]
3. He, L.; Lu, Z.; Li, X. Failure-mode importance measures in structural system with multiple failure modes and its estimation using copula. *Reliab. Eng. Syst. Saf.* **2018**, *174*, 53–59. [[CrossRef](#)]

4. Liao, Y.; Xiang, Y.; Keedy, E. Reliability analysis of flow meters with multiple failure modes in the process industry. In Proceedings of the 2020 Annual Reliability and Maintainability Symposium (RAMS), Palm Springs, CA, USA, 27–30 January 2020; pp. 1–6.
5. Lee, Y. Expectation-Maximization Algorithms for Learning a Finite Mixture of Univariate Survival Time Distributions from Partially Specified Class Values. Ph.D. Thesis, Iowa State University, Ames, IA, USA, 2013.
6. Balakrishnan, N.; Aggarwala, R. *Progressive Censoring: Theory, Methods, and Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2000.
7. Ng, H.K.T.; Chan, P.S.; Balakrishnan, N. Estimation of parameters from progressively censored data using EM algorithm. *Comput. Stat. Data Anal.* **2002**, *39*, 371–386. [[CrossRef](#)]
8. Balakrishnan, N.; Kannan, N.; Lin, C.T.; Ng, H.T. Point and interval estimation for Gaussian distribution, based on progressively Type-II censored samples. *IEEE Trans. Reliab.* **2003**, *52*, 90–95. [[CrossRef](#)]
9. Basak, I.; Balakrishnan, N. Robust estimation under progressive censoring. *Comput. Stat. Data Anal.* **2003**, *44*, 349–376. [[CrossRef](#)]
10. Abo-Eleneen, Z.A. The entropy of progressively censored samples. *Entropy* **2011**, *13*, 437–449. [[CrossRef](#)]
11. Kotb, M.S.; Raqab, M.Z. Statistical inference for modified Weibull distribution based on progressively type-II censored data. *Math. Comput. Simul.* **2019**, *162*, 233–248. [[CrossRef](#)]
12. Bai, X.; Shi, Y.; Liu, Y.; Liu, B. Reliability inference of stress–strength model for the truncated proportional hazard rate distribution under progressively Type-II censored samples. *Appl. Math. Model.* **2019**, *65*, 377–389. [[CrossRef](#)]
13. Zhang, Z.; Gui, W. Statistical inference of reliability of Generalized Rayleigh distribution under progressively type-II censoring. *J. Comput. Appl. Math.* **2019**, *361*, 295–312. [[CrossRef](#)]
14. Feroze, N.; Aslam, M.; Khan, I.H.; Khan, M.H. Bayesian reliability estimation for the Topp–Leone distribution under progressively type-II censored samples. *Soft Comput.* **2020**, 1–22. [[CrossRef](#)]
15. Cho, Y.; Sun, H.; Lee, K. An estimation of the entropy for a Rayleigh distribution based on doubly-generalized Type-II hybrid censored samples. *Entropy* **2014**, *16*, 3655–3669. [[CrossRef](#)]
16. Algarni, A.; Almarashi, A.M.; Okasha, H.; Ng, H.K.T. E-Bayesian Estimation of Chen Distribution Based on Type-I Censoring Scheme. *Entropy* **2020**, *22*, 636. [[CrossRef](#)]
17. Cho, Y.; Sun, H.; Lee, K. Estimating the entropy of a Weibull distribution under generalized progressive hybrid censoring. *Entropy* **2015**, *17*, 102–122. [[CrossRef](#)]
18. Lee, G.; Scott, C. EM algorithms for multivariate Gaussian mixture models with truncated and censored data. *Comput. Stat. Data Anal.* **2012**, *56*, 2816–2829. [[CrossRef](#)]
19. Tian, Y.; Tian, M.; Zhu, Q. Estimating a finite mixed exponential distribution under progressively type-II censored data. *Commun. Stat. Theory Methods* **2013**, *43*, 3762–3776. [[CrossRef](#)]
20. Tian, Y.; Zhu, Q.; Tian, M. Estimation for mixed exponential distributions under type-II progressively hybrid censored samples. *Comput. Stat. Data Anal.* **2015**, *89*, 85–96. [[CrossRef](#)]
21. Tahir, M.; Aslam, M.; Hussain, Z.; Abbas, N. On the finite mixture of exponential, Rayleigh and Burr Type-XII distributions: Estimation of parameters in Bayesian framework. *Electron. J. Appl. Stat. Anal.* **2017**, *10*, 271–293.
22. Abushal, T.A.; AL-Zaydi, A.M. Inference on Constant-Partially Accelerated Life Tests for Mixture of Pareto Distributions under Progressive Type-II Censoring. *Open J. Stat.* **2017**, *7*, 323. [[CrossRef](#)]
23. Feroze, N.; Aslam, M. Approximate Bayesian analysis of doubly censored samples from mixture of two Weibull distributions. *Commun. Stat. Theory Methods* **2019**, *48*, 2862–2878. [[CrossRef](#)]
24. Feroze, N.; Aslam, M. Three-Component Mixture of Rayleigh Model Under Doubly Censored Samples: A Bayesian Look. *Commun. Math. Stat.* **2019**, *7*, 417–443. [[CrossRef](#)]
25. Zio, E. Reliability engineering: Old problems and new challenges. *Reliab. Eng. Syst. Saf.* **2009**, *94*, 125–141. [[CrossRef](#)]
26. Deng, Y. Deng entropy. *Chaos Solitons Fractals* **2016**, *91*, 549–553. [[CrossRef](#)]
27. Tang, Y.; Zhou, D.; Xu, S.; He, Z. A weighted belief entropy-based uncertainty measure for multi-sensor data fusion. *Sensors* **2017**, *17*, 928. [[CrossRef](#)] [[PubMed](#)]
28. Huang, Z.; Yang, L.; Jiang, W. Uncertainty measurement with belief entropy on the interference effect in the quantum-like Bayesian networks. *Appl. Math. Comput.* **2019**, *347*, 417–428. [[CrossRef](#)]
29. Wu, D.; Tang, Y. An improved failure mode and effects analysis method based on uncertainty measure in the evidence theory. *Qual. Reliab. Eng. Int.* **2020**. [[CrossRef](#)]

30. Liu, Z.g.; Liu, Y.; Dezert, J.; Cuzzolin, F. Evidence combination based on credal belief redistribution for pattern classification. *IEEE Trans. Fuzzy Syst.* **2019**, *28*, 618–631. [[CrossRef](#)]
31. Masson, M.H.; Denoeux, T. ECM: An evidential version of the fuzzy *c*-means algorithm. *Pattern Recognit.* **2008**, *41*, 1384–1397. [[CrossRef](#)]
32. Pan, L.; Deng, Y. An association coefficient of a belief function and its application in a target recognition system. *Int. J. Intell. Syst.* **2020**, *35*, 85–104. [[CrossRef](#)]
33. Qin, M.; Tang, Y.; Wen, J. An Improved Total Uncertainty Measure in the Evidence Theory and Its Application in Decision Making. *Entropy* **2020**, *22*, 487. [[CrossRef](#)]
34. Zhang, H.; Deng, Y. Weighted belief function of sensor data fusion in engine fault diagnosis. *Soft Comput.* **2020**, *24*, 2329–2339. [[CrossRef](#)]
35. Xiao, F.; Cao, Z.; Jolfaei, A. A novel conflict measurement in decision making and its application in fault diagnosis. *IEEE Trans. Fuzzy Syst.* **2020**. [[CrossRef](#)]
36. Zhou, K.; Martin, A.; Pan, Q.; Liu, Z. Median evidential *c*-means algorithm and its application to community detection. *Knowl. Based Syst.* **2015**, *74*, 69–88. [[CrossRef](#)]
37. Lepskiy, A.; Meshcheryakova, N. Belief functions for the importance assessment in multiplex networks. In Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems 2020, Lisbon, Portugal, 15–19 June 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 283–296.
38. Jaeger, M. The AI&M procedure for learning from incomplete data. In Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence, Cambridge, MA, USA, 13–16 July 2006; pp. 225–232.
39. Couso, I.; Dubois, D. A general framework for maximizing likelihood under incomplete data. *Int. J. Approx. Reason.* **2018**, *93*, 238–260. [[CrossRef](#)]
40. Zhang, M.; Ye, Z.; Xie, M. A stochastic em algorithm for progressively censored data analysis. *Qual. Reliab. Eng. Int.* **2014**, *30*, 711–722. [[CrossRef](#)]
41. Denoeux, T. Maximum likelihood estimation from uncertain data in the belief function framework. *Knowl. Data Eng. IEEE Trans.* **2013**, *25*, 119–130. [[CrossRef](#)]
42. Denoeux, T.; Kanjanatarakul, O.; Sriboonchitta, S. A new evidential *k*-nearest neighbor rule based on contextual discounting with partially supervised learning. *Int. J. Approx. Reason.* **2019**, *113*, 287–302. [[CrossRef](#)]
43. Gupta, R.D.; Kundu, D. Exponentiated exponential family: An alternative to gamma and Weibull distributions. *Biom. J. J. Math. Methods Biosci.* **2001**, *43*, 117–130. [[CrossRef](#)]
44. Mokhtari, E.B.; Rad, A.H.; Yousefzadeh, F. Inference for Weibull distribution based on progressively Type-II hybrid censored data. *J. Stat. Plan. Inference* **2011**, *141*, 2824–2838. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).