

Article

Different Statistical Inference Algorithms for the New Pareto Distribution Based on Type-II Progressively Censored Competing Risk Data with Applications

Essam A. Ahmed ^{1,2}, Tariq S. Alshammari ³ and Mohamed S. Eliwa ^{4,5,*} ¹ College of Business Administration, Taibah University, Al-Madina 41411, Saudi Arabia² Department of Mathematics, Sohag University, Sohag 82524, Egypt³ Department of Mathematics, College of Science, University of Ha'il, Hail, Saudi Arabia⁴ Department of Statistics and Operations Research, College of Science, Qassim University, Saudi Arabia⁵ Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt

* Correspondence: m.eliwa@qu.edu.sa

Abstract: In this research, the statistical inference of unknown lifetime parameters is proposed in the presence of independent competing risks using a progressive Type-II censored dataset. The lifetime distribution associated with a failure mode is assumed to follow the new Pareto distribution, with consideration given to two distinct competing failure reasons. Maximum likelihood estimators (MLEs) for the unknown model parameters, as well as reliability and hazard functions, are derived, noting that they are not expressible in closed form. The Newton–Raphson, expectation maximization (EM), and stochastic expectation maximization (SEM) methods are employed to generate maximum likelihood (ML) estimations. Approximate confidence intervals for the unknown parameters, reliability, and hazard rate functions are constructed using the normal approximation of the MLEs and the normal approximation of the log-transformed MLEs. Additionally, the missing information principle is utilized to derive the closed form of the Fisher information matrix, which, in turn, is used with the delta approach to calculate confidence intervals for reliability and hazards. Bayes estimators are derived under both symmetric and asymmetric loss functions, with informative and non-informative priors considered, including independent gamma distributions for informative priors. The Monte Carlo Markov Chain sampling approach is employed to obtain the highest posterior density credible intervals and Bayesian point estimates for unknown parameters and reliability characteristics. A Monte Carlo simulation is conducted to assess the effectiveness of the proposed techniques, with the performances of the Bayes and maximum likelihood estimations examined using average values and mean squared errors as benchmarks. Interval estimations are compared in terms of average lengths and coverage probabilities. Real datasets are considered and examined for each topic to provide illustrative examples.

Keywords: progressive type-II censored data; competing risk model; expectation maximization algorithm; stochastic expectation maximization algorithm; Bayes theorem; Metropolis–Hastings algorithm; observed Fisher information matrix; computer simulation; statistics and numerical data

MSC: 62F10; 62F12; 62F15; 62F40



Citation: Ahmed, E.A.; Alshammari, T.S.; Eliwa, M.S. Different Statistical Inference Algorithms for the New Pareto Distribution Based on Type-II Progressively Censored Competing Risk Data with Applications. *Mathematics* **2024**, *12*, 2136. <https://doi.org/10.3390/math12132136>

Academic Editor: Diana Mindrila

Received: 16 May 2024

Revised: 1 July 2024

Accepted: 4 July 2024

Published: 7 July 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/>).

1. Introduction

Lifetime studies offer a valuable approach across engineering sciences, medical fields, and economics for exploring the survival distribution of entities or individuals. Analyzing data from such studies necessitates the consideration of the anticipated lifetime distribution, with typical distributions including the exponential, Weibull, Burr, Pareto, and Gamma distributions. Among these, the Pareto distribution family is widely acknowledged in the literature for its ability to model data with heavy tails. One significant outcome of

advancements in Pareto modeling is the spotlight it casts on the advantageous properties for measuring income inequality. The Pareto distribution has found applications in various disciplines such as actuarial science, economics, finance, life testing, reliability, survival analysis, and engineering, owing to its tractable nature as a lifespan model. However, it becomes evident that modeling data with heavy-tailed distributions necessitates considerations beyond the capabilities of the Pareto distribution alone. Hence, there arises a need for alternatives. Several distributions based on the Pareto framework have been proposed to address this need. The new Pareto (NP) distribution, introduced by [1], is one such recently discovered distribution that merits further investigation due to its significance across multiple domains. The probability density function (PDF) and cumulative distribution function (CDF) of the NP distribution, denoted $NP(\alpha, \beta)$, are as follows:

$$f(x, \alpha, \beta) = \frac{2\alpha\beta^\alpha x^{\alpha-1}}{(x^\alpha + \beta^\alpha)^2}, x \geq \beta, \alpha > 0, \beta > 0, \quad (1)$$

and

$$F(x, \alpha, \beta) = 1 - \frac{2\beta^\alpha}{x^\alpha + \beta^\alpha}, x \geq \beta, \alpha > 0, \beta > 0, \quad (2)$$

where α and β denote the shape and scale parameters of the distribution, respectively. The survival function (SF) for the NP distribution is derived from Equation (2) as follows:

$$S(x, \alpha, \beta) = \frac{2\beta^\alpha}{x^\alpha + \beta^\alpha}, x \geq \beta, \alpha > 0, \beta > 0. \quad (3)$$

The hazard rate function (HRF) can be formulated as

$$H(x, \alpha, \beta) = \frac{f(x, \alpha, \beta)}{1 - F(x, \alpha, \beta)} = \frac{\alpha x^{\alpha-1}}{(x^\alpha + \beta^\alpha)}, x \geq \beta, \alpha > 0, \beta > 0. \quad (4)$$

To highlight certain unique characteristics of this function, its first derivative with respect to x can be expressed as follows:

$$H'(x, \alpha, \beta) = \frac{\alpha(\alpha - 1)x^{\alpha-2}(x^\alpha + \beta^\alpha) - \alpha^2 x^{2\alpha-2}}{(x^\alpha + \beta^\alpha)^2} = \frac{\alpha x^{\alpha-2} \Delta(x)}{(x^\alpha + \beta^\alpha)^2}. \quad (5)$$

As mentioned in some of the literature, ref. [2] reported that the NP hazard rate function can be unimodal shaped or decreasing depending on its parameters. For $x > \beta > 0$, it becomes evident that $\Delta(x) = (\alpha - 1)\beta^\alpha - x^\alpha$ and $\Delta(\beta) = \beta^\alpha(\alpha - 2) < 0$ if $\alpha < 2$. Consequently, for $\alpha < 2$, $H'(x, \alpha, \beta) < 0$, indicating that $H(x, \alpha, \beta)$ is decreasing concerning x . In the case of $\alpha > 2$, the function $H(x, \alpha, \beta)$ exhibits a distinct mode at $x = x_0$, where $H(x, \alpha, \beta)$ increases for all $x < x_0$ and decreases for all $x > x_0$, with $x_0 = \beta\{\alpha - 1\}^{1/\alpha}$. Figure 1 shows how the behavior of $H(x, \alpha, \beta)$ is changed when the significant increasing of the parameters α and β happened, as we can see this effect very clearly from the graph. It is clear to us from this figure that the hazard rate of the NP distribution may exhibit either an upside-down bathtub (unimodal) shape or a decreasing trend, contingent upon the values of various parameters.

The NP distribution is often employed for modeling a wide range of real-world datasets, such as insurance, reliability, engineering, and economics. For example, ref. [1] delved into the mathematical properties of the NP distribution, demonstrating its utility in modeling income and reliability data through the analysis of seven real datasets. Ref. [3] used the new Pareto distribution to explain the income of the upper-class group using the Malaysian household incomes dataset from 2012–2019. Further insights into this distribution were uncovered in [2], where the authors derived simpler formulas for relevant risk measures and inequality indices. Ref. [4] conducted parameter estimation for the NP distribution using two distinct methodologies: classical estimation and Bayesian inference,

with the latter utilizing importance sampling. Additionally, ref. [5] explored parameter estimation across various phases of progressively censored samples derived from the NP distribution and made predictions regarding failure times for removed units using a Bayesian approach and Markov chain Monte Carlo techniques. Furthermore, ref. [6] applied this distribution to data on rainfall and COVID-19. Ref. [7] utilized Type-II censored data and full samples to estimate NP distribution parameters employing the mixed Gibbs sampling method. They used the NP distribution to model the data of the air conditioning system in an aircraft, as well as some data for mechanical units. Recently, ref. [8] applied the NP distribution to some different data, including the net worth of the affluent in Singapore and China, the daily increment of NASDAQ-100, and the daily new-case rate of COVID-19.

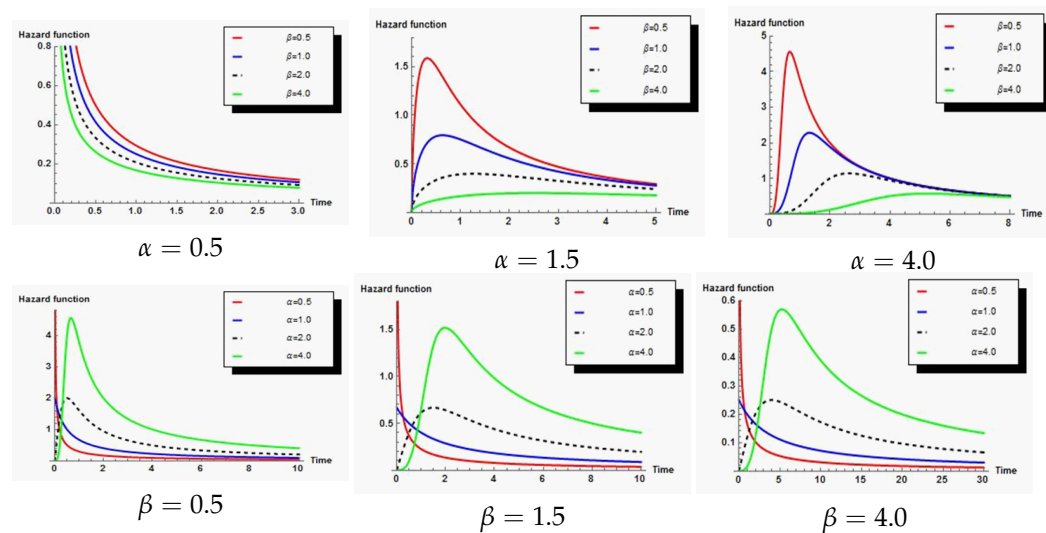


Figure 1. The effects of parameters α and β on the hazard function of $NP(\alpha, \beta)$.

Recognizing the presence or absence of competing risks is pivotal when considering the distribution of units' lifetimes in research. In realistic testing scenarios, it is well understood that failures often stem from multiple factors, competing within the lifespan continuum. This phenomenon is referred to as a competing risk model in the statistical literature and finds widespread application across various scientific domains, including economics, medical sciences, electronic engineering, and social sciences. In the medical field, for instance, reducing mortality rates from specific diseases is a key objective of public health initiatives. However, assessing the impact on overall mortality and life expectancy is complicated by the presence of competing causes of death. Individuals face risks from various ailments like heart disease, diabetes, hypertension, cancer, AIDS, and tuberculosis, all vying for their longevity, even though death is typically attributed to a single cause. Similarly, in assessing the reliability of car tires, factors such as sidewall damage, punctures, or tread wear can lead to failure, highlighting the competition among different failure modes. In industrial reliability contexts, the focus on specific system components prone to failure may overlook potential risks posed by other components, thus introducing a competing risk problem. Competition among failure causes exists in each scenario, with only one mode typically leading to failure. Therefore, the precise inference of each failure mode in the presence of others becomes necessary. Moreover, in medical survival analysis and industrial experimentation, the item of interest often becomes inadvertently lost or removed before failure, leading to data censorship as a result of these investigations.

Due to the challenge of obtaining comprehensive lifespan data for every component in life experiments, researchers often categorize control systems into two types: Type-I and Type-II. Type-I censorship occurs when the experimental duration is fixed, and the number of failures becomes random. Conversely, Type-II censorship arises when the experimental time is unpredictable, but the number of failures is predetermined. Both types entail that any remaining items cannot be removed before the test concludes. Therefore, a progressive

Type-II censoring approach, introduced in [9], aims to streamline resources and reduce costs by implementing multiple censoring stages according to a predefined scheme. The practical implementation of progressive Type-II censorship involves initiating the experiment with n units to be tested. The experiment concludes at the failure time of the m th (where $m < n$) unit. Upon the occurrence of the first failure, R_1 units are randomly selected from the remaining $n - 1$ units, and the failure time, $X_{1:m:n}$, is recorded. Subsequently, R_2 units are randomly removed after noting the second failure time, $X_{2:m:n}$, and this process continues until the m th product fails. At this point, all surviving units are removed, and the failure time, $X_{m:m:n}$, is noted. Here, R_m represents the number of units removed for the m th time. Thus, $X_{1:m:n}, X_{2:m:n}, \dots, X_{m:m:n}$ constitute the progressive Type-II censored sample. It is essential to note that the total number of units, n , is the sum of $R_1, R_2, \dots, R_m$ and m , with $R_1, R_2, \dots, R_m$ being predefined constants. Further insights into censorship-related studies can be found in [9], which serves as a valuable resource in this area. Building upon the progressive Type-II censored sample drawn from an NP distribution, ref. [6] investigated issues related to estimating unknown characteristics and predicting the failure periods of the eliminated units.

In recent years, there has been a growing interest among researchers in statistical inference, particularly concerning the presence of competing risks and various forms of censoring in available data. A multitude of recent studies have addressed this topic, spanning different censoring schemes and statistical models. For instance, ref. [10] explored competing risk analysis under Type-II hybrid censoring for the two-parameter exponential distribution. Ref. [11] investigated statistical inference for competing hazard models using progressive interval censored Weibull data. Ref. [12] analyzed a competing risk model with Weibull distributions under a unified hybrid censoring scheme, deriving maximum likelihood estimators and approximate confidence intervals for distributional parameters. Similarly, ref. [13] focused on statistical inference for competing risk models with inverted exponentiated distributions under a unified hybrid censoring scheme. Additionally, researchers have examined various Bayesian and maximum likelihood estimation techniques under different censoring scenarios and model assumptions. Notably, there is a noticeable gap in the literature regarding the application of the new Pareto lifespan distribution within competing risk models under any form of censorship. Hence, our study aimed to address this gap by studying the maximum likelihood and Bayesian estimation of the unknown parameters, reliability, and hazard rate functions of the NP distribution under progressively censored competing risks models. Moreover, we introduce innovative methodologies, such as employing expectation maximization (EM) and stochastic EM (SEM) algorithms for deriving maximum likelihood estimators for the NP distribution, a novel approach not previously explored. To achieve these objectives, we employ a variety of methods, including the Newton–Raphson method, EM, SEM algorithms, and Bayesian estimation techniques, utilizing different loss functions and prior distributions. Furthermore, we conducted a Monte Carlo simulation analysis to evaluate the efficacy of the proposed estimators using metrics such as absolute bias, mean squared error, average width, and coverage probability. Finally, we illustrate our findings using a real-world dataset, emphasizing the practical relevance of our research.

This paper's structure is organized as follows: Section 2 provides descriptions of the models under the competing risks scenario. In Section 3, maximum likelihood estimators (MLEs) and corresponding approximate confidence intervals (ACIs) for the NP distribution's unknown parameters in the progressively censored competing risk setting are derived using the Newton–Raphson (NR) method. The existence and uniqueness of MLEs are also discussed in this section. Section 4 applies the EM and SEM approximation methods to obtain the MLEs of the unknown parameters. Additionally, a Fisher information matrix is constructed to facilitate approximate interval estimations in this section. Section 5 utilizes the Markov chain Monte Carlo (MCMC) technique to generate Bayes estimates for the unknown parameters, reliability, and hazard functions, assuming independent gamma priors for the unknown parameters. HPD credible intervals for the unknown parameters

are also established based on MCMC samples. Section 6 presents a Monte Carlo simulation investigation to compare the suggested estimates in terms of average bias, mean squared error (MSE), average width (AW), and coverage probability (CP). In Section 7, real-world datasets and a simulation example are examined for illustrative purposes. Finally, Section 8 concludes the paper with some closing remarks.

2. Model and Data Overview

Consider a life testing experiment commencing with a set of $n \in N$ identical units, where the failure times of these units are denoted by $X_1, X_2, \dots, X_n$. To simplify, let us assume that there are only two distinct causes of failure. Thus, for each unit $i = 1, \dots, n$, we have $X_i = \min\{X_{i1}, X_{i2}\}$, where X_{ik} , ($k = 1, 2$) represents the latent failure time of the i th unit under the k th cause of failure. We make the assumptions that the latent failure times X_{i1} and X_{i2} are statistically independent and the pairs (X_{i1}, X_{i2}) are identically and independently distributed (i.i.d).

As is clear from Figure 1, the effect of the β parameter is not significant on the hazard function, unlike the α parameter. Therefore, here, we assume that the failure times follow the NP distribution, with a common scale parameter β and distinct shape parameters (α_k , $k = 1, 2$). The CDF $F_k(x)$ and the PDF of the j th failure cause of a random variable X_{ij} are provided in [1] as follows:

$$f_k(x, \Theta) = \frac{2\alpha_k \beta^{\alpha_k} x^{\alpha_k-1}}{(x^{\alpha_k} + \beta^{\alpha_k})^2}, \text{ and } F_k(x, \Theta) = 1 - \frac{2\beta^{\alpha_k}}{x^{\alpha_k} + \beta^{\alpha_k}}, x \geq \beta, \alpha_k > 0, \beta > 0. \quad (6)$$

Given the observation $X_i = \min\{X_{i1}, X_{i2}\}$, where X_{i1} and X_{i2} represent two distinct failure times for a test unit, we solely consider the smaller of the two, indicating the overall failure time. Subsequently, the CDF of this overall failure time is readily derived as follows:

$$\begin{aligned} F(x, \Theta) &= 1 - (1 - F_1(x, \Theta))(1 - F_2(x, \Theta)) = 1 - \left(\frac{2\beta^{\alpha_1}}{x^{\alpha_1} + \beta^{\alpha_1}} \right) \left(\frac{2\beta^{\alpha_2}}{x^{\alpha_2} + \beta^{\alpha_2}} \right) \\ &= 1 - 4 \prod_{j=1}^2 \frac{\beta^{\alpha_j}}{(x^{\alpha_j} + \beta^{\alpha_j})}, x \geq \beta, \alpha_k > 0, \beta > 0. \end{aligned} \quad (7)$$

Consequently, the PDF can be represented as

$$\begin{aligned} f(x, \Theta) &= \frac{4\beta^{(\alpha_1+\alpha_2)} \left[(\alpha_1 + \alpha_2)x^{(\alpha_1+\alpha_2)} + \alpha_2\beta^{\alpha_1}x^{\alpha_2} + \alpha_1\beta^{\alpha_2}x^{\alpha_1} \right]}{x(x^{\alpha_1} + \beta^{\alpha_1})^2(x^{\alpha_2} + \beta^{\alpha_2})^2} \\ &= 4\beta^{(\alpha_1+\alpha_2)} \sum_{k=1}^2 \frac{\alpha_k x^{\alpha_k-1}}{(x^{\alpha_k} + \beta^{\alpha_k})^2(x^{\alpha_{3-k}} + \beta^{\alpha_{3-k}})} \\ &= \sum_{k=1}^2 f_k(x, \Theta) \bar{F}(x, \Theta), x \geq \beta, \alpha_k > 0, \beta > 0. \end{aligned} \quad (8)$$

Subsequently, the survival and hazard functions manifest in the subsequent form:

$$S(t, \Theta) = 4 \prod_{j=1}^2 \frac{\beta^{\alpha_j}}{(t^{\alpha_j} + \beta^{\alpha_j})}, t \geq \beta, \alpha_k > 0, \beta > 0, \quad (9)$$

and

$$H(t, \Theta) = \sum_{k=1}^2 \frac{\alpha_k x^{\alpha_k-1}}{(x^{\alpha_k} + \beta^{\alpha_k})}, t \geq \beta, \alpha_k > 0, \beta > 0. \quad (10)$$

Let

$$I(\delta_i = k) = \begin{cases} 1, & \delta_i = k \\ 0, & \text{otherwise} \end{cases}, i = 1, 2, \dots, m,$$

then

$$m_k = \sum_{i=1}^m I(\delta_i = k),$$

signifies the total count of units that failed because of cause k ($k = 1, 2$) and $m_1 + m_2 = m$. In exploiting the independence of the latent failure times X_{i1} and X_{i2} for $i = 1, 2, \dots, n$, the relative risk rate attributed to a specific cause (let us say, cause 1) can be derived as follows:

$$\begin{aligned} p &= P(X_1 < X_2) = \int_{\beta}^{\infty} F_1(x, \alpha_1, \beta) f_2(x, \alpha_2, \beta) dx \\ &= \int_{\beta}^{\infty} \frac{2\alpha_2 \beta^{\alpha_2} x^{\alpha_2-1}}{(x^{\alpha_2} + \beta^{\alpha_2})^2} \left(1 - \frac{2\beta^{\alpha_1}}{x^{\alpha_1} + \beta^{\alpha_1}}\right) dx \\ &= 1 - \int_{\beta}^{\infty} \frac{4\alpha_2 \beta^{\alpha_1+\alpha_2} x^{\alpha_2-1}}{(x^{\alpha_2} + \beta^{\alpha_2})^2 (x^{\alpha_1} + \beta^{\alpha_1})} dx. \end{aligned} \quad (11)$$

A numerical approach is necessary to solve the integral on the right side of Equation (11) since it lacks an analytical solution. Once $P(X_1 < X_2)$ is determined, $P(X_2 < X_1)$ can be computed using the relationship $P(X_2 < X_1) = 1 - P(X_1 < X_2)$. Hence, if $m_1 \sim \text{Binomial}(m, p)$, then $m_2 \sim \text{Binomial}(m, 1 - p)$. In progressive Type-II censoring, the total number of failed individuals m and the predefined censored scheme $(R_1, R_2, \dots, R_m)$ are specified in advance, where $R_m = n - m - \sum_{i=1}^m R_i$. Consequently, the observed set of progressive Type-II censored data with competing risks can be expressed as

$$(X_{1:m:n}, \delta_1, R_1), (X_{j:m:n}, \delta_j, R_j), \dots, (X_m, \delta_m, R_m).$$

To simplify the notation, we denote $X_{i:m:n}$ as X_i . Then, the likelihood equation for the progressive Type-II censored data, based on the competing risk model, is expressed as

$$L = C(R) \prod_{i=1}^m [f_1(x_i) \bar{F}_2(x_i)]^{I(\delta_i=1)} [f_2(x_i) (\bar{F}_1(x_i))]^{I(\delta_i=2)} [\bar{F}_1(x_i) \bar{F}_2(x_i)]^{R_i}, \quad (12)$$

where

$$C(R) = n(n - R_1 - 1)(n - R_1 - R_2 - m + 1), \text{ and } \bar{F}_k(x_i) = 1 - F_k(x_i), k = 1, 2. \quad (13)$$

3. Estimation of Maximum Likelihood Using Different Algorithms

A widely employed and highly regarded statistical estimation technique is MLE. MLE stands out as a preferred and efficient choice for parametric estimation procedures, owing to its consistency, efficiency, and asymptotic normality properties. This section is dedicated to utilizing maximum likelihood to derive both point and interval estimates for the unknown parameters α_1 , α_2 , and β , as well as for the reliability $S(t, \alpha_1, \alpha_2, \beta)$ and hazard $H(t, \alpha_1, \alpha_2, \beta)$ functions, within the framework of progressive Type-II censoring with competing risk data.

3.1. MLEs via Newton–Raphson Algorithm

A widely utilized numerical technique for determining MLEs is the NR algorithm. In this context, we outline the NR algorithm for computing the MLEs of α_1 , α_2 , β , $S(t)$, and $H(t)$. Let $x_{1:m:n} < x_{2:m:n} < \dots < x_{m:m:n}$ denote the order statistics of Type-II progressively censored data with competing risks from independent NP distributions $NP(\alpha_1, \beta)$ and $NP(\alpha_2, \beta)$, respectively, and a pre-fixed censoring scheme $(R_1, R_2, \dots, R_m)$. Henceforth, we will use $\mathbf{x} = (x_1, x_2, \dots, x_m)$ in lieu of $(x_{1:m:n}, x_{2:m:n}, \dots, x_{m:m:n})$. The likelihood function utilizing Equations (6) and (12) is presented as follows:

$$L \propto \prod_{i=1}^m \prod_{k=1}^2 \left\{ \frac{\alpha_k \beta^{\alpha_k} \beta^{\alpha_{3-k}} x_i^{\alpha_k-1}}{(x_i^{\alpha_k} + \beta^{\alpha_k})^2 (x_i^{\alpha_{3-k}} + \beta^{\alpha_{3-k}})} \right\}^{I(\delta_i=k)} \prod_{i=1}^m \prod_{k=1}^2 \left\{ \frac{\beta^{\alpha_k}}{(x_i^{\alpha_k} + \beta^{\alpha_k})} \right\}^{R_i}.$$

In disregarding the additive constant, the log-likelihood function is expressed as

$$\begin{aligned}\mathcal{L} = \ln L = & \sum_{k=1}^2 m_k \log \alpha_k + n \log \beta \sum_{k=1}^2 \alpha_k + \sum_{k=1}^2 \sum_{i=1}^m (\alpha_k - 1) I(\delta_i = k) \log x_i \\ & - \sum_{k=1}^2 \sum_{i=1}^m \Delta_k(\delta_i, R_i) \log(x_i^{\alpha_k} + \beta^{\alpha_k}),\end{aligned}\quad (14)$$

where

$$\Delta_k(\delta_i, R_i) = I(\delta_i = k) + R_i + 1, i = 1, 2, \dots; \text{ and } k = 1, 2. \quad (15)$$

Taking the derivative of Equation (15) with respect to α_k and β , where $k = 1, 2$, yields the partial derivatives of the likelihood as follows:

$$\frac{\partial \mathcal{L}}{\partial \alpha_k} \propto \frac{m_k}{\alpha_k} + n \log \beta + \sum_{i=1}^m I(\delta_i = k) \log x_i - \sum_{i=1}^m \Delta_k(\delta_i, R_i) \left[\frac{x_i^{\alpha_k} \log x_i + \beta^{\alpha_k} \log \beta}{(x_i^{\alpha_k} + \beta^{\alpha_k})} \right], \quad (16)$$

and

$$\frac{\partial \mathcal{L}}{\partial \beta} \propto \sum_{k=1}^2 \left\{ \frac{n \alpha_k}{\beta} - \sum_{i=1}^m \frac{\Delta_k(\delta_i, R_i) \alpha_k \beta^{\alpha_k - 1}}{(x_i^{\alpha_k} + \beta^{\alpha_k})} \right\}. \quad (17)$$

Concurrently solving the intricate nonlinear equations $\frac{\partial \mathcal{L}}{\partial \alpha_k} = 0$ for $k = 1, 2$ and $\frac{\partial \mathcal{L}}{\partial \beta} = 0$ allows for the determination of the MLEs of α_k and β . However, obtaining closed forms for Equations (16) and (17) proves challenging. Hence, various numerical techniques, such as NR, are employed to compute the MLEs for the unknown parameters. Notably, given $x \geq \beta$, the MLE for β is straightforwardly derived as $\beta = x_1$. Furthermore, it is evident that the MLE for α_k can be obtained as a fixed-point solution of the following equation:

$$\Phi(\alpha_k) = 1/\alpha_k, k = 1, 2, \quad (18)$$

where

$$\Phi(\alpha_k) = m_k^{-1} \left\{ n \log x_1 + \sum_{i=1}^m I(\delta_i = k) \log x_i - \sum_{i=1}^m \frac{\Delta_k(\delta_i, R_i) (x_i^{\alpha_k} \log x_i + x_1^{\alpha_k} \log x_1)}{(x_i^{\alpha_k} + x_1^{\alpha_k})} \right\}, k = 1, 2. \quad (19)$$

The prevalent approach for numerically solving Equation (19) is through the straightforward iterative method, $\Phi(\alpha_k^{(j)}) = \alpha_k^{(j)}$, where (j) is the value obtained in the j th iteration. The solutions to the likelihood equations are denoted as $\hat{\alpha}_k$ ($k = 1, 2$) and $\hat{\beta}$. The following steps illustrate the specific processes involved in this iteration method:

Step 1: Initial values for $\Theta = (\alpha_1, \alpha_2, \beta)$ should be given with $j = 0$; that is, $\Theta^{(0)} = (\alpha_1^{(0)}, \alpha_2^{(0)}, \beta^{(0)})$.

Step 2: In the j th iteration, calculate $\left(\frac{\partial \mathcal{L}}{\partial \alpha_1}, \frac{\partial \mathcal{L}}{\partial \alpha_2}, \frac{\partial \mathcal{L}}{\partial \beta} \right)^T \Big|_{\alpha_1=\alpha_1^{(j)}, \alpha_2=\alpha_2^{(j)}, \beta=\beta^{(j)}}$ and $I = I(\alpha_1^{(j)}, \alpha_2^{(j)}, \beta^{(j)})$, where

$$I = I(\alpha_1^{(j)}, \alpha_2^{(j)}, \beta^{(j)}) = \begin{bmatrix} -\mathcal{L}_{11} & -\mathcal{L}_{12} & -\mathcal{L}_{13} \\ -\mathcal{L}_{21} & -\mathcal{L}_{22} & -\mathcal{L}_{23} \\ -\mathcal{L}_{31} & -\mathcal{L}_{32} & -\mathcal{L}_{33} \end{bmatrix}_{\alpha_1=\alpha_1^{(j)}, \alpha_2=\alpha_2^{(j)}, \beta=\beta^{(j)}}. \quad (20)$$

The observed information matrix of the parameters α_1 , α_2 , and β is denoted as I . The elements of the matrix I are as follows:

$$\mathcal{L}_{kk} = \frac{\partial^2 \mathcal{L}}{\partial \alpha_k^2} \propto -\frac{m_k}{\alpha_k^2} - \sum_{i=1}^m \left[\frac{\Delta_k(\delta_i, R_i) \beta^{\alpha_k} x_i^{\alpha_k} \log^2(x_i/\beta)}{(x_i^{\alpha_k} + \beta^{\alpha_k})^2} \right], k = 1, 2, \quad (21)$$

$$\mathcal{L}_{12} = \mathcal{L}_{21} = \frac{\partial^2 \mathcal{L}}{\partial \alpha_k \partial \alpha_{3-k}} = \frac{\partial^2 \mathcal{L}}{\partial \alpha_{3-k} \partial \alpha_k} = 0 \quad (22)$$

$$\mathcal{L}_{k3} = \mathcal{L}_{3k} = \frac{\partial^2 \mathcal{L}}{\partial \alpha_k \partial \beta} = \frac{\partial^2 \mathcal{L}}{\partial \beta \partial \alpha_k} = \frac{n}{\beta} - \sum_{i=1}^m \frac{\Delta_k(\delta_i, R_i) \beta^{\alpha_k-1} [x_i^{\alpha_k} (1 - \alpha_k \log(x_i/\beta)) + \beta^{\alpha_k}]}{(x_i^{\alpha_k} + \beta^{\alpha_k})^2}, \quad (23)$$

and

$$\begin{aligned} \mathcal{L}_{33} = \frac{\partial^2 \mathcal{L}}{\partial \beta^2} &\propto -\frac{n}{\beta^2} \sum_{k=1}^2 \alpha_k - 2 \sum_{k=1}^2 \sum_{i=1}^m \frac{\alpha_k \beta^{\alpha_k-2} I(\delta_i = k) [\beta^{\alpha_k} + (\alpha_k - 1) x_i^{\alpha_k}]}{(x_i^{\alpha_k} + \beta^{\alpha_k})^2} \\ &- \sum_{k=1}^2 \sum_{i=1}^m \frac{[I(\delta_i = 3 - k) + R_i] \alpha_k \beta^{\alpha_k-2} [\beta^{\alpha_k} + (\alpha_k - 1) x_i^{\alpha_k}]}{(x_i^{\alpha_k} + \beta^{\alpha_k})^2}, k = 1, 2. \end{aligned} \quad (24)$$

Step 3: Set

$$(\alpha_1^{(j+1)}, \alpha_2^{(j+1)}, \beta^{(j+1)})^T = A \times \left(\frac{\partial \mathcal{L}}{\partial \alpha_1}, \frac{\partial \mathcal{L}}{\partial \alpha_2}, \frac{\partial \mathcal{L}}{\partial \beta} \right)^T \Big|_{\alpha_1=\alpha_1^{(j)}, \alpha_2=\alpha_2^{(j)}, \beta=\beta^{(j)}},$$

where $A = (\alpha_1^{(j)}, \alpha_2^{(j)}, \beta^{(j)})^T + I^{-1}(\alpha_1^{(j)}, \alpha_2^{(j)}, \beta^{(j)})$, $(\alpha_1, \alpha_2, \beta)^T$ is the transpose of vector $(\alpha_1, \alpha_2, \beta)$, and $I^{-1}(\alpha_1^{(j)}, \alpha_2^{(j)}, \beta^{(j)})$ represents the inverse of the matrix $I(\alpha_1^{(j)}, \alpha_2^{(j)}, \beta^{(j)})$.

Step 4: In setting $j = j + 1$, the MLEs of the parameters (denoted by $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}$) can be obtained by repeating Steps 2 and 3 until $\left| (\alpha_1^{(j+1)}, \alpha_2^{(j+1)}, \beta^{(j+1)})^T - (\alpha_1^{(j)}, \alpha_2^{(j)}, \beta^{(j)})^T \right| < \varepsilon$, where ε is a threshold value that is fixed in advance.

Utilizing the acquired point estimators and the invariance property of MLEs, we can derive the estimates of the reliability and hazard functions from Equations (9) and (10), which are expressed as follows:

$$\hat{S}(t) = 4\beta^{\hat{\alpha}_1 + \hat{\alpha}_2} \left[\prod_{k=1}^2 (t^{\hat{\alpha}_k} + \hat{\beta}^{\hat{\alpha}_k}) \right]^{-1}, \text{ and } \hat{H}(t) = \sum_{k=1}^2 \frac{\hat{\alpha}_k x^{\hat{\alpha}_k-1}}{(t^{\hat{\alpha}_k} + \hat{\beta}^{\hat{\alpha}_k})}, t > \beta > 0. \quad (25)$$

3.2. Existence and Uniqueness of MLEs

The existence and uniqueness of the MLEs are fundamental aspects to consider in statistical inference.

Theorem 1. *The maximum likelihood estimator of α_k , where $k = 1, 2$, exists and is unique for the lifetimes of objects subject to competing risks and following the NP distribution with parameters (α_1, β) and (α_2, β) when $m_k > 0$ for $k = 1, 2$.*

Proof. Since $x \geq \beta$, we determine the MLE $\hat{\beta} = x_1$, and the MLE $\hat{\alpha}_k$ of α_k , $k = 1, 2$, can be reported from the solution of the following equation:

$$\frac{\partial \mathcal{L}}{\partial \alpha_k} \propto \frac{m_k}{\alpha_k} + n \log x_1 + \sum_{i=1}^m I(\delta_i = k) \log x_i - \sum_{i=1}^m \Delta_k(\delta_i, R_i) \left[\frac{x_i^{\alpha_k} \log x_i + x_1^{\alpha_k} \log x_1}{(x_i^{\alpha_k} + x_1^{\alpha_k})} \right], k = 1, 2. \quad (26)$$

The left-hand side of Equation (26) is a continuous function. As α_k approaches 0, the left-hand side of Equation (26) tends to infinity, and as $\alpha_k \rightarrow \infty$, the left-hand side of Equation (26) tends to ϕ , where ϕ represents

$$\begin{aligned} \phi &= \lim_{\alpha_k \rightarrow \infty} \frac{\partial \mathcal{L}}{\partial \alpha_k} = n \log x_1 - \sum_{i=1}^m I(\delta_i = k) \log x_i - \sum_{i=1}^m [I(\delta_i = 3 - k) + R_i] \log x_i \\ &= n \log x_1 - \sum_{i=1}^m \log x_i = -\sum_{i=2}^m \log x_i < 0. \end{aligned} \quad (27)$$

Hence, the solution of Equation (26) exists. Additionally, we find that the second derivative of α_k ($k = 1, 2$) takes the following form:

$$\frac{\partial^2 \mathcal{L}}{\partial \alpha_k^2} \propto -\frac{m_k}{\alpha_k^2} - \sum_{i=1}^m \left[\frac{\Delta_k(\delta_i, R_i) \beta^{\alpha_k} x_i^{\alpha_k} \log^2(x_i/\beta)}{(x_i^{\alpha_k} + \beta^{\alpha_k})^2} \right] < 0. \quad (28)$$

As Equation (28) is always negative, Equation (26) has a unique solution, and this solution represents the maximum likelihood estimator (MLE) of α_k ($k = 1, 2$). Consequently, we deduce that $\frac{\partial \mathcal{L}}{\partial \alpha_k}$ is a continuous function on $(0, \infty)$, and it monotonically decreases from ∞ to negative values. This demonstrates the existence and uniqueness of the MLE of α_k ($k = 1, 2$). $\square$

3.3. Approximate Confidence Intervals

In this section, we derive the approximate confidence intervals (ACIs) for the unknown parameters α_1 , α_2 , and β , as well as for the reliability and hazard functions, utilizing the asymptotic normality of the MLEs. This approach is based on the asymptotic properties of MLEs. Based on the regularity conditions, the MLEs $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta})$ are approximately normally distributed with mean $(\alpha_1, \alpha_2, \beta)$ and variance–covariance matrix $I^{-1}(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta})$. Or equivalently,

$$(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}) - (\alpha_1, \alpha_2, \beta) \sim N\left(0, I^{-1}(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta})\right), \quad (29)$$

where from (20),

$$I^{-1}(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}) = \begin{bmatrix} -\mathcal{L}_{11} & -\mathcal{L}_{12} & -\mathcal{L}_{13} \\ -\mathcal{L}_{21} & -\mathcal{L}_{22} & -\mathcal{L}_{23} \\ -\mathcal{L}_{31} & -\mathcal{L}_{32} & -\mathcal{L}_{33} \end{bmatrix}_{\Theta=\hat{\Theta}}^{-1} = \begin{bmatrix} \text{Var}(\hat{\alpha}_1) & \text{Cov}(\hat{\alpha}_1, \hat{\alpha}_2) & \text{Cov}(\hat{\alpha}_1, \hat{\beta}) \\ \text{Cov}(\hat{\alpha}_2, \hat{\alpha}_1) & \text{Var}(\hat{\alpha}_2) & \text{Cov}(\hat{\alpha}_2, \hat{\beta}) \\ \text{Cov}(\hat{\beta}, \hat{\alpha}_1) & \text{Cov}(\hat{\beta}, \hat{\alpha}_2) & \text{Var}(\hat{\beta}) \end{bmatrix}, \quad (30)$$

is the inverse of the observed Fisher information matrix with $\Theta = (\alpha_1, \alpha_2, \beta)$. Furthermore, the expression $\mathcal{L}_{ij} = \frac{\partial^2 \mathcal{L}}{\partial \Theta_i \partial \Theta_j}$, for $i, j = 1, 2, 3$, is given by Equations (21)–(24). Thus, the $100(1 - \gamma)\%$ ACIs for α_1 , α_2 , and β are given by

$$\left(\hat{\alpha}_1 \pm Z_{\gamma/2} \sqrt{\text{Var}(\hat{\alpha}_1)}\right), \left(\hat{\alpha}_2 \pm Z_{\gamma/2} \sqrt{\text{Var}(\hat{\alpha}_2)}\right), \text{ and } \left(\hat{\beta} \pm Z_{\gamma/2} \sqrt{\text{Var}(\hat{\beta})}\right), \quad (31)$$

where $Z_{\gamma/2}$ is the upper $(\gamma/2)$ -th point of the standard normal distribution $N(0, 1)$. Alternatively, the ACIs for $S(t)$ and the hazard function $H(t)$ can be constructed using the asymptotic normality of the MLE. To derive these intervals, we employ the delta method (Greene [14]) to estimate the variances of their estimators. The delta method is a statistical technique used to approximate the probability distribution of a function of an asymptotically normal estimator by employing the Taylor series approximation. In utilizing this method, the approximate variances of $\hat{S}(t)$ and $\hat{H}(t)$ are determined through the following steps:

Step 1: Let Ω_1 and Ω_2 be two quantities that take the following forms:

$$\Omega_1 = \left(\frac{\partial S(t)}{\partial \alpha_1}, \frac{\partial S(t)}{\partial \alpha_2}, \frac{\partial S(t)}{\partial \beta}\right) \text{ and } \Omega_2 = \left(\frac{\partial H(t)}{\partial \alpha_1}, \frac{\partial H(t)}{\partial \alpha_2}, \frac{\partial H(t)}{\partial \beta}\right),$$

where from Equations (9) and (10),

$$\begin{aligned} \frac{\partial S(t)}{\partial \alpha_k} &= \frac{4\beta^{\alpha_k + \alpha_{3-k}} t^{\alpha_k} \{\log t - \log \beta\}}{(t^{\alpha_k} + \beta^{\alpha_k})^2 (t^{\alpha_{3-k}} + \beta^{\alpha_{3-k}})}, k = 1, 2, \\ \frac{\partial S(t)}{\partial \beta} &= -\frac{4\beta^{\alpha_1 + \alpha_2 - 1} \{(\alpha_1 + \alpha_2) t^{\alpha_1 + \alpha_2} + \alpha_1 \beta^{\alpha_2} t^{\alpha_1} + \alpha_2 \beta^{\alpha_1} t^{\alpha_2}\}}{(t^{\alpha_1} + \beta^{\alpha_1})^2 (t^{\alpha_2} + \beta^{\alpha_2})^2}, \end{aligned}$$

$$\frac{\partial H(t)}{\partial \alpha_k} = \frac{t^{\alpha_k-1} \{t^{\alpha_k} + \beta^{\alpha_k} (1 + \alpha_k (\log t - \log \beta))\}}{(t^{\alpha_k} + \beta^{\alpha_k})^2}, k = 1, 2,$$

and

$$\frac{\partial H(t)}{\partial \beta} = -\frac{1}{\beta t} \sum_{k=1}^2 \frac{t^{\alpha_k} \alpha_k^2 \beta^{\alpha_k}}{(t^{\alpha_k} + \beta^{\alpha_k})^2}, k = 1, 2.$$

Step 2: Using the following formulas, determine the approximate variances of $S(t)$ and $H(t)$:

$$\text{Var}(\hat{S}(t)) \simeq [\Omega_1 \text{Var}(\hat{\Theta}) \Omega_1^T]_{\Theta=\hat{\Theta}}, \text{ and } \text{Var}(\hat{H}(t)) \simeq [\Omega_2 \text{Var}(\hat{\Theta}) \Omega_2^T]_{\Theta=\hat{\Theta}}, \quad (32)$$

where $\text{Var}(\hat{\Theta})$ is obtained from (30) for $\hat{\Theta} = (\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta})$.

Step 3: Determine the $100(1 - \gamma)\%$ asymptotic confidence intervals for $S(t)$ and $H(t)$ using the following formula:

$$\left(\hat{S}(t) \pm Z_{\gamma/2} \sqrt{\text{Var}(\hat{S}(t))} \right), \text{ and } \left(\hat{H}(t) \pm Z_{\gamma/2} \sqrt{\text{Var}(\hat{H}(t))} \right) \quad (33)$$

3.4. Log-Normal Approximation Confidence Intervals (LACIs)

In instances where the lower bound of the asymptotic confidence intervals may fall below 0, conflicting with the prerequisite $\Theta > 0$, the issue can be circumvented through log transformation and the delta method. We have $\ln \hat{\Theta} \sim N(\ln \Theta, \text{Var}(\ln \hat{\Theta}))$, where $\text{Var}(\ln \hat{\Theta}) = \frac{\text{Var}(\hat{\Theta})}{\hat{\Theta}^2}$. The log-transformed MLEs of $\Theta = (\alpha_1, \alpha_2, \beta)$ obtain their asymptotic confidence intervals at $100(1 - \gamma)\%$ as follows:

$$\left[\hat{\Theta} \exp \left(-Z_{(\gamma/2)} \sqrt{\frac{\text{Var}(\hat{\Theta})}{\hat{\Theta}^2}} \right), \hat{\Theta} \exp \left(Z_{(\gamma/2)} \sqrt{\frac{\text{Var}(\hat{\Theta})}{\hat{\Theta}^2}} \right) \right], \hat{\Theta} = (\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}). \quad (34)$$

Similarly, the $100(1 - \gamma)\%$ LACIs of $S(t)$ and $H(t)$ can be listed as

$$\left[\hat{S}(t) \exp \left(-Z_{(\gamma/2)} \sqrt{\frac{\text{Var}(\hat{S}(t))}{\hat{S}(t)^2}} \right), \hat{S}(t) \exp \left(Z_{(\gamma/2)} \sqrt{\frac{\text{Var}(\hat{S}(t))}{\hat{S}(t)^2}} \right) \right], \quad (35)$$

and

$$\left[\hat{H}(t) \exp \left(-Z_{(\gamma/2)} \sqrt{\frac{\text{Var}(\hat{H}(t))}{\hat{H}(t)^2}} \right), \hat{H}(t) \exp \left(Z_{(\gamma/2)} \sqrt{\frac{\text{Var}(\hat{H}(t))}{\hat{H}(t)^2}} \right) \right]. \quad (36)$$

4. Expectation Maximization (EM) Algorithm

Given that the ML estimators are not straightforwardly derived and that initial values can impact the maximum likelihood estimates obtained from the NR method, we resort to the EM algorithm to compute the estimations of the unknown parameters $(\alpha_1, \alpha_2, \beta)$, alongside the reliability function $S(t)$ and hazard function $H(t)$ of the NP models. Ref. [15] introduced a generic iterative approach for computing the MLEs of unknown parameters when observed and censored data are available, known as the EM algorithm. This method offers several advantages, such as its capability to address complex problems, the assured increase in log-likelihood with each iteration, the straightforward yet meticulous computations, and the elimination of the necessity for second- and higher-order derivatives. The EM algorithm comprises two key steps: the expectation step (E-step) and the maximization step (M-step). In the E-step, the conditional expectation of the missing data is determined based on the observed data and the current parameter estimates, and the pseudo-log-likelihood function is calculated. Subsequently, in the M-step, the likelihood function under the observed, and censored data are maximized. This method is particularly useful for analyzing partially or censored datasets, such as those encountered in the Type-II

progressive censoring with competing risks sample model. Thus, we employed the EM algorithm to obtain the MLEs of $\alpha_1, \alpha_2, \beta, S(t)$, and $H(t)$.

4.1. Point Estimation via EM Algorithm

Consider the observed sample denoted as $X = (x_{1:m:n}, x_{2:m:n}, \dots, x_{m:m:n})$ and the censored data as $Z = (Z_1, Z_2, \dots, Z_m)$, where $Z_i = (z_{i1}, z_{i2}, \dots, z_{iR_i})$ for $i = 1, 2, \dots, m$. Here, the censored data represent the missing data. It is important to note that the complete sample comprises both the observed sample and the censored data. Thus, the complete sample is represented as $W = (X, Z)$. For the parameter set $\Theta = (\alpha_1, \alpha_2, \beta)$, the likelihood function of the complete sample data W is expressed as

$$L_c(\Theta) \propto \prod_{i=1}^m \prod_{k=1}^2 \left\{ [f_k(x_{i:m:n}) \bar{F}_{3-k}(x_{i:m:n})] \prod_{j=1}^{R_i} [f_k(z_{ij}) \bar{F}_{3-k}(z_{ij})] \right\}^{I(\delta_i=k)}. \quad (37)$$

Hence, the log-likelihood function for α_1, α_2 , and β derived from a complete sample is given by

$$\begin{aligned} \mathcal{L}_c = & \sum_{k=1}^2 n_k \ln \alpha_k + n \ln \beta \sum_{k=1}^2 \alpha_k + \sum_{k=1}^2 (\alpha_k - 1) \left\{ \sum_{i=1}^m I(\delta_i = k) \left(\log x_i + \sum_{j=1}^{R_i} \log z_{ij} \right) \right\} \\ & - \sum_{k=1}^2 \left\{ \sum_{i=1}^m \Psi(\delta_i) \left[\log(x_i^{\alpha_k} + \beta^{\alpha_k}) + \sum_{j=1}^{R_i} \log(z_{ij}^{\alpha_k} + \beta^{\alpha_k}) \right] \right\}, \end{aligned} \quad (38)$$

where $n_k = \sum_{i=1}^m I(\delta_i = k)(R_i + 1)$ represents the total number of failures for each cause in the complete dataset with size n , $n = n_1 + n_2$, and $\Psi(d) = 2I(d = k) + I(d = 3 - k) = I(d = k) + 1$. After obtaining the log-likelihood function in Equation (38) with respect to α and β and setting the normal equations to zero, the maximum likelihood estimators (MLEs) of the parameters α_1, α_2 , and β for the entire sample W can be determined as follows:

$$\begin{aligned} \frac{\partial \mathcal{L}_c}{\partial \alpha_k} = & \frac{n_k}{\alpha_k} + n \ln \beta + \sum_{i=1}^m I(\delta_i = k) \left[\log x_i + \sum_{j=1}^{R_i} \log z_{ij} \right] \\ & - \sum_{i=1}^m \Psi(\delta_i) \left[\frac{(x_i^{\alpha_k} \log x_i + \beta^{\alpha_k} \log \beta)}{(x_i^{\alpha_k} + \beta^{\alpha_k})} + \sum_{j=1}^{R_i} \frac{z_{ij}^{\alpha_k} \log z_{ij} + \beta^{\alpha_k} \log \beta}{(z_{ij}^{\alpha_k} + \beta^{\alpha_k})} \right] = 0, k = 1, 2, \end{aligned} \quad (39)$$

and

$$\frac{\partial \mathcal{L}_c}{\partial \beta} = \frac{n \sum_{k=1}^2 \alpha_k}{\beta} - \sum_{k=1}^2 \sum_{i=1}^m \Psi(\delta_i) \left[\frac{\alpha_k \beta^{\alpha_k-1}}{x_i^{\alpha_k} + \beta^{\alpha_k}} + \sum_{j=1}^{R_i} \frac{\alpha_k \beta^{\alpha_k-1}}{z_{ij}^{\alpha_k} + \beta^{\alpha_k}} \right] = 0. \quad (40)$$

The pseudo-log-likelihood function changes during the E-step to become

$$\begin{aligned} & \frac{n_k}{\alpha_k} + n \ln \beta + \sum_{i=1}^m I(\delta_i = k) \log x_i + \sum_{i=1}^m I(\delta_i = k) \sum_{j=1}^{R_i} E[\log z_{ij} | z_{ij} > x_i] \\ & - \sum_{i=1}^m \Psi(\delta_i) \left[\frac{x_i^{\alpha_k} \log x_i + \beta^{\alpha_k} \log \beta}{x_i^{\alpha_k} + \beta^{\alpha_k}} + E \left(\frac{z_{ij}^{\alpha_k} \log z_{ij} + \beta^{\alpha_k} \log \beta}{z_{ij}^{\alpha_k} + \beta^{\alpha_k}} | z_{ij} > x_i \right) \right] = 0, k = 1, 2, \end{aligned} \quad (41)$$

and

$$\frac{n \sum_{k=1}^2 \alpha_k}{\beta} - \sum_{k=1}^2 \sum_{i=1}^m \Psi(\delta_i) \left[\frac{\alpha_k \beta^{\alpha_k-1}}{x_i^{\alpha_k} + \beta^{\alpha_k}} + \sum_{j=1}^{R_i} E \left[\frac{\alpha_k \beta^{\alpha_k-1}}{z_{ij}^{\alpha_k} + \beta^{\alpha_k}} | z_{ij} > x_i \right] \right] = 0. \quad (42)$$

In the following steps, we need to derive the following result.

Theorem 2. Given $X_1 = x_1, \dots, X_i = x_i$, the conditional distribution of z_{ij} , $j = 1, \dots, R_i$, which has the left-truncated NP distribution at x_i , is given by

$$f_{Z|X}(z_{ij}|X_i = x_i = x_i, \alpha_1, \alpha_2, \beta) = \frac{f_Z(z_{ij}|\alpha_1, \alpha_2, \beta)}{1 - F_X(x_i|\alpha_1, \alpha_2, \beta)}, z_{ij} > x_i, \quad (43)$$

where $F_X(x_i; \alpha_1, \alpha_2, \beta)$ and $f_Z(z_{ij}; \alpha_1, \alpha_2, \beta)$ can be found in Equations (7) and (8), respectively.

Proof. The proof is comprehensible and can be explored in detail in [16]. Hence, based on Equation (43), the conditional expectations described in Equations (39) and (42) are generated, as illustrated below:

$$\begin{aligned} E_1(x_i, \alpha_1, \alpha_2, \beta) &= E[\log z_{ij}|z_{ij} > x_i] \\ &= \frac{4\beta^{(\alpha_1+\alpha_2)}}{1 - F_X(x_i; \alpha_1, \alpha_2, \beta)} \sum_{k=1}^2 \int_{x_i}^{\infty} \frac{\alpha_k y^{\alpha_k-1} \log y}{(y^{\alpha_k} + \beta^{\alpha_k})^2 (y^{\alpha_{3-k}} + \beta^{\alpha_{3-k}})} dy, \end{aligned} \quad (44)$$

$$\begin{aligned} E_2(x_i, \alpha_1, \alpha_2, \beta) &= E\left[\frac{z_{ij}^{\alpha_k} \log z_{ij} + \beta^{\alpha_k} \log \beta}{z_{ij}^{\alpha_k} + \beta^{\alpha_k}} | z_{ij} > x_i\right] \\ &= \frac{4\beta^{(\alpha_1+\alpha_2)}}{1 - F_X(x_i; \alpha_1, \alpha_2, \beta)} \sum_{k=1}^2 \int_{x_i}^{\infty} \frac{\alpha_k y^{\alpha_k-1} \log(y^{\alpha_k} \log y + \beta^{\alpha_k} \log \beta)}{(y^{\alpha_k} + \beta^{\alpha_k})^3 (y^{\alpha_{3-k}} + \beta^{\alpha_{3-k}})} dy, \end{aligned} \quad (45)$$

and

$$\begin{aligned} E_3(x_i, \alpha_1, \alpha_2, \beta) &= E\left[\frac{\alpha_k \beta^{\alpha_k-1}}{z_{ij}^{\alpha_k} + \beta^{\alpha_k}} | z_{ij} > x_i\right] \\ &= \frac{4\beta^{(\alpha_1+\alpha_2)}}{1 - F_X(x_i; \alpha_1, \alpha_2, \beta)} \sum_{k=1}^2 \int_{x_i}^{\infty} \frac{\alpha_k^2 y^{\alpha_k-1} \beta^{\alpha_k-1}}{(y^{\alpha_k} + \beta^{\alpha_k})^3 (y^{\alpha_{3-k}} + \beta^{\alpha_{3-k}})} dy. \end{aligned} \quad (46)$$

In the M-step of the EM algorithm, at the $(l+1)$ iteration, the value of $\beta^{(l+1)}$ is first obtained by solving the following equation:

$$\frac{n \sum_{k=1}^2 \alpha_k^{(l)}}{\beta^{(l+1)}} - \sum_{k=1}^2 \sum_{i=1}^m \Psi(\delta_i) \left[\frac{\alpha_k^{(l)} (\beta^{(l+1)})^{\alpha_k^{(l)}-1}}{x_i^{\alpha_k^{(l)}} + (\beta^{(l+1)})^{\alpha_k^{(l)}-1}} + \sum_{j=1}^{R_i} E_3(x_i, \alpha_1^{(l)}, \alpha_2^{(l)}, \beta^{(l)}) \right] = 0.$$

The estimate of β might then be obtained by

$$\hat{\beta} = \left[n \sum_{k=1}^2 \alpha_k^{(l)} \right] \left\{ \sum_{k=1}^2 \sum_{i=1}^m \Psi(\delta_i) \left[\frac{\alpha_k^{(l)} (\beta^{(l+1)})^{\alpha_k^{(l)}-1}}{x_i^{\alpha_k^{(l)}} + (\beta^{(l+1)})^{\alpha_k^{(l)}-1}} + \sum_{j=1}^{R_i} E_3(x_i, \alpha_1^{(l)}, \alpha_2^{(l)}, \beta^{(l)}) \right] \right\}^{-1}. \quad (47)$$

After obtaining $\beta^{(l+1)}$, solve the following to obtain $\alpha_1^{(l+1)}$ and $\alpha_2^{(l+1)}$:

$$\begin{aligned} &\frac{n_k}{\alpha_k^{(l+1)}} + n \ln \beta^{(l+1)} + \sum_{i=1}^m I(\delta_i = k) \log x_i - \sum_{i=1}^m \Psi(\delta_i) \left[\frac{x_i^{\alpha_k^{(l+1)}} \log x_i + \beta^{(l+1)\alpha_k^{(l+1)}} \log \beta^{(l+1)}}{x_i^{\alpha_k^{(l+1)}} + \beta^{(l+1)\alpha_k^{(l+1)}}} \right] \\ &+ \sum_{i=1}^m I(\delta_i = k) \sum_{j=1}^{R_i} E_1(x_i, \alpha_1^{(l)}, \alpha_2^{(l)}, \beta^{(l+1)}) - \sum_{i=1}^m \Psi(\delta_i) \sum_{j=1}^{R_i} E_2(x_i, \alpha_1^{(l)}, \alpha_2^{(l)}, \beta^{(l+1)}) = 0, k = 1, 2. \end{aligned}$$

The estimate of α_k might then be obtained by

$$\hat{\alpha}_k = \frac{n_k}{\sum_{i=1}^m \Psi(\delta_i) \left[\frac{x_i^{\alpha_k^{(l+1)}} \log x_i + \beta^{(l+1)\alpha_k^{(l+1)}} \log \beta^{(l+1)}}{x_i^{\alpha_k^{(l+1)}} + (\beta^{(l+1)})^{\alpha_k^{(l)}}} \right] - \sum_{i=1}^m I(\delta_i = k) \log x_i} - n \ln \beta^{(l+1)} + \frac{n_k}{\sum_{i=1}^m \Psi(\delta_i) \sum_{j=1}^{R_i} E_2(x_i, \alpha_1^{(l)}, \alpha_2^{(l)}, \beta^{(l+1)}) - \sum_{i=1}^m I(\delta_i = k) \sum_{j=1}^{R_i} E_1(x_i, \alpha_1^{(l)}, \alpha_2^{(l)}, \beta^{(l+1)})}. \quad (48)$$

The next iteration uses $(\alpha_1^{(l+1)}, \alpha_2^{(l+1)}, \beta^{(l+1)})$ as the new value of $(\alpha_1, \alpha_2, \beta)$. Now, an iterative process can be employed to obtain the necessary maximum likelihood estimates of α_1, α_2 , and β . This iterative procedure continues until

$$\left| \alpha_1^{(l+1)} - \alpha_1^{(l)} \right| + \left| \alpha_2^{(l+1)} - \alpha_2^{(l)} \right| + \left| \beta^{(l+1)} - \beta^{(l)} \right| < \varepsilon,$$

for a predetermined small value of ε and some l . This suggested approach converges to the local maximum likelihood as the log-likelihood increases with each iteration. We use the maximum likelihood estimates of the parameters based on the entire sample as starting values in the expectation maximization technique. Henceforth, we denote the maximum likelihood estimates of α_1, α_2 , and β as their respective MLEs: $\hat{\alpha}_{1EM}, \hat{\alpha}_{2EM}$, and $\hat{\beta}_{EM}$. Furthermore, the MLEs of $S(t)$ and $H(t)$ can be straightforwardly derived using the in-variance property of the MLEs. These are given as

$$\hat{S}_{EM}(t) = \frac{4\beta_{EM}^{\hat{\alpha}_{1EM} + \hat{\alpha}_{2EM}}}{\prod_{k=1}^2 (t^{\hat{\alpha}_{kEM}} + \hat{\beta}_{kEM})}, \text{ and } \hat{H}(t) = \sum_{k=1}^2 \frac{\hat{\alpha}_{kEM} x^{\hat{\alpha}_{kEM}-1}}{(t^{\hat{\alpha}_{kEM}} + \hat{\beta}_{kEM})}, t > \beta > 0. \quad (49)$$

□

4.2. Point Estimation via Stochastic Expectation Maximization (SEM) Algorithm

Most EM algorithms follow a simple closed-form approach, iterating through two phases. However, a notable limitation of the EM method is its susceptibility to becoming trapped in a saddle point, particularly when handling high-dimensional or complex data such as censored lifespan models (see [17]). Notably, the EM equations mentioned earlier lack a closed form, necessitating numerical computation. Hence, to obtain maximum likelihood estimators, we employed the SEM method in this context. The SEM algorithm, proposed in [18], replaces the expectation step (E-step) of the EM method with a stochastic step (S-step). This modification involves substituting each missing datum with a randomly generated value from the conditional distribution of the missing data, based on the observed data and current parameter values. By incorporating random values obtained from the conditional distribution of the missing data, the SEM algorithm augments the observed sample. It is often preferred over the EM approach due to its simplicity, avoidance of complex computations, and lack of reliance on calculating conditional expectations. Moreover, studies have demonstrated that the SEM approach is robust to initial values and performs effectively with small sample sizes; see [19]. Let $W = (X, Z)$ represent the complete dataset, where $X = (x_1, x_2, \dots, x_m)$ denotes the observed data, and $Z_i = (z_{i1}, z_{i2}, \dots, z_{iR_i})$ for $i = 1, 2, \dots, m$, and $j = 1, 2, \dots, R_i$ signifies the censored data when x_i fails. Following the SEM algorithm's principle, we initially generate the missing samples $z_{ij}, i = 1, 2, \dots, m$ and $j = 1, 2, \dots, R_i$, from the truncated conditional distribution below:

$$G_Z(z_{ij}; \alpha_1, \alpha, \beta | z_{ij} > x_i) = \frac{F_Z(z_{ij}; \Theta) - F_{X_i}(x_i; \alpha_1, \alpha, \beta)}{1 - F_{X_i}(x_i; \alpha_1, \alpha, \beta)}, z_{ij} > x_i. \quad (50)$$

To obtain a random sample from Equation (50), we start by generating a random value from the uniform distribution $U(0, 1)$, denoted as u . Subsequently, the realization of z_{ij} is derived as

$$z_{ij} = F^{-1}(u + (1 - u)F_{X_i}(x_i; \Theta)), \quad (51)$$

where $F^{-1}(\cdot)$ represents the inverse function of $F(\cdot)$, as defined in Equation (7). Subsequently, in replacing each value of z_{ij} with the value generated by Equation (53), the ML estimators of α_1 , α_2 , and β at the $(l + 1)$ th stage are obtained from Equations (39) and (40) using

$$\beta^{(l+1)} = \left[n \sum_{k=1}^2 \alpha_k^{(l)} \right] \left\{ \sum_{k=1}^2 \sum_{i=1}^m \Psi(\delta_i) \left[\frac{\alpha_k^{(l)} (\beta^{(l)})^{\alpha_k^{(l)} - 1}}{x_i^{\alpha_k^{(l)}} + (\beta^{(l)})^{\alpha_k^{(l)} - 1}} + \sum_{j=1}^{R_i} \frac{\alpha_k^{(l)} (\beta^{(l)})^{\alpha_k^{(l)} - 1}}{z_{ij}^{\alpha_k^{(l)}} + (\beta^{(l)})^{\alpha_k^{(l)} - 1}} \right] \right\}^{-1}, \quad (52)$$

and

$$\begin{aligned} \alpha_k^{(l+1)} = & \frac{n_k}{\left\{ \sum_{i=1}^m \Psi(\delta_i) \left[\frac{x_i^{\alpha_k^{(l)}} \log x_i + (\beta^{(l+1)})^{\alpha_k^{(l)}} \log \beta^{(l+1)}}{x_i^{\alpha_k^{(l)}} + (\beta^{(l+1)})^{\alpha_k^{(l)}}} \right] - \sum_{i=1}^m I(\delta_i = k) \log x_i \right\} - n \ln \beta^{(l+1)}} \\ & + \frac{n_k}{\sum_{i=1}^m \Psi(\delta_i) \sum_{j=1}^{R_i} \frac{z_{ij}^{\alpha_k^{(l)}} \log z_{ij} + (\beta^{(l+1)})^{\alpha_k^{(l)}} \log \beta^{(l+1)}}{z_{ij}^{\alpha_k^{(l)}} + (\beta^{(l+1)})^{\alpha_k^{(l)}}} - \sum_{i=1}^m I(\delta_i = k) \sum_{j=1}^{R_i} \log z_{ij}}. \end{aligned} \quad (53)$$

Like in the EM algorithm, the iterations can be terminated when $|\alpha_1^{(h+1)} - \alpha_1^{(h)}| + |\alpha_2^{(h+1)} - \alpha_2^{(h)}| + |\beta^{(h+1)} - \beta^{(h)}| < \varepsilon$.

4.3. Fisher Information Matrix

This section presents the observed Fisher information matrix derived from [20]'s concept of missing values. It is worth noting that this observed Fisher information matrix can be utilized for constructing asymptotic confidence intervals. The missing information principle, referenced in various scholarly articles, can be summarized as follows:

$$\text{Observed information} = \text{Complete information} - \text{Missing information}. \quad (54)$$

Let us employ the notation shown as follows: $\Theta = (\alpha_1, \alpha_2, \beta)$; X : the observed data; W : the complete data; and $I_X(\Theta)$, $I_W(\Theta)$, and $I_{W|X}(\Theta)$ represent the observed, complete, and missing information, respectively. By the definition in [20], observed information is the difference between complete and missing information; this may be stated in the following manner:

$$I_X(\Theta) = I_W(\Theta) - I_{W|X}(\Theta). \quad (55)$$

The complete information matrix $I_W(\Theta)$ is provided as follows:

$$\begin{aligned} I_W(\Theta) = & -E \left[\frac{\partial^2 \ln L(W, \Theta)}{\partial \Theta^2} \right]_{\Theta=(\alpha_1, \alpha_2, \beta)} \\ = & \begin{bmatrix} a_{11}(\alpha_1, \alpha_2, \beta) & a_{12}(\alpha_1, \alpha_2, \beta) & a_{13}(\alpha_1, \alpha_2, \beta) \\ a_{21}(\alpha_1, \alpha_2, \beta) & a_{22}(\alpha_1, \alpha_2, \beta) & a_{23}(\alpha_1, \alpha_2, \beta) \\ a_{31}(\alpha_1, \alpha_2, \beta) & a_{32}(\alpha_1, \alpha_2, \beta) & a_{33}(\alpha_1, \alpha_2, \beta) \end{bmatrix}_{|\Theta=\Theta}. \end{aligned} \quad (56)$$

The symbol $a_{ij}(\alpha_1, \alpha_2, \beta)$ represents the matrices' elements for $I_W(\Theta)$. They are listed in the following order:

$$a_{kk} = -E \left[\frac{\partial^2 \ln L(\Theta)}{\partial \alpha_k^2} \right] = 2 \sum_{i=1}^n E \left[\frac{\beta^{\alpha_k} (\log x_i - \log \beta)^2 x_i^{\alpha_k}}{(x_i^{\alpha_k} + \beta^{\alpha_k})^2} \right] - \sum_{i=1}^n E \left[\frac{U_1(\underline{x}, \alpha_k, \alpha_{3-k}, \beta)^2}{V(\underline{x}, \alpha_k, \alpha_{3-k}, \beta)^2} \right] + \sum_{i=1}^n E \left[\frac{U_2(\underline{x}, \alpha_k, \alpha_{3-k}, \beta)}{V(\underline{x}, \alpha_k, \alpha_{3-k}, \beta)} \right],$$

where

$$U_1(\underline{x}, \alpha_k, \alpha_{3-k}, \beta) = \left[\begin{array}{c} \beta^{\alpha_{3-k}} (1 + \alpha_k \ln x_i) x_i^{\alpha_k} + \alpha_{3-k} \beta^{\alpha_k} x_i^{\alpha_{3-k}} \ln \beta \\ + (1 + (\alpha_k + \alpha_{3-k}) \ln x_i) x_i^{(\alpha_k + \alpha_{3-k})} \end{array} \right],$$

$$U_2(\underline{x}, \alpha_k, \alpha_{3-k}, \beta) = \beta^{\alpha_{3-k}} \ln x_i (2 + \alpha_k \ln x_i) x_i^{\alpha_k} + \alpha_2 \beta^{\alpha_k} x_i^{\alpha_{3-k}} \ln^2 \beta + \ln x_i [2 + (\alpha_k + \alpha_{3-k}) \ln x_i] x_i^{\alpha_k + \alpha_{3-k}},$$

and

$$V(\underline{x}, \alpha_k, \alpha_{3-k}, \beta) = \left[\alpha_k \beta^{\alpha_{3-k}} x_i^{\alpha_k} + \alpha_{3-k} \beta^{\alpha_k} x_i^{\alpha_{3-k}} + (\alpha_k + \alpha_{3-k}) x_i^{(\alpha_k + \alpha_{3-k})} \right],$$

$$a_{k(3-k)} = a_{(3-k)k} = -E \left[\frac{\partial^2 \ln L(\Theta)}{\partial \alpha_k \partial \alpha_{3-k}} \right] = -E \left[\frac{\partial^2 \ln L(\Theta)}{\partial \alpha_{3-k} \partial \alpha_k} \right] = \sum_{i=1}^n E \left[\frac{\prod_{k=1}^2 U_1(\underline{x}, \alpha_k, \alpha_{3-k}, \beta) (1 + V(\underline{x}, \alpha_k, \alpha_{3-k}, \beta))}{V(\underline{x}, \alpha_k, \alpha_{3-k}, \beta)^2} \right],$$

$$\begin{aligned} a_{k3} = a_{3k} &= -E \left[\frac{\partial^2 \ln L(\eta)}{\partial \alpha_k \partial \beta} \right] = -E \left[\frac{\partial^2 \ln L(\eta)}{\partial \beta \partial \alpha_k} \right] \\ &= \frac{-n}{\beta} + 2 \sum_{i=1}^n E \left[\frac{\beta^{\alpha_k - 1} [\beta^{\alpha_k} + (1 + \alpha_k \log \beta - \alpha_k \log x_i) x_i^{\alpha_k}]}{(x_i^{\alpha_k} + \beta^{\alpha_k})^2} \right] \\ &+ \sum_{i=1}^n \left[\frac{\alpha_{3-k} x_i^{\alpha_{3-k}} \left[x_i^{\alpha_k} \beta^{(\alpha_k + \alpha_{3-k})} (\alpha_{3-k} + \alpha_k (\alpha_k - \alpha_{3-k}) \log \beta + \alpha_k (\alpha_{3-k} - \alpha_k) \log x_i) \right]}{(V(\underline{x}, \alpha_k, \alpha_{3-k}, \beta))^2} \right] \\ &+ \sum_{i=1}^n \left[\frac{\alpha_{3-k} x_i^{\alpha_{3-k}} \left[\alpha_{3-k} \beta^{\alpha_{3-k}} x_i^{2\alpha_k} + \alpha_{3-k} \beta^{2\alpha_k} x_i^{\alpha_{3-k}} \right]}{V(\underline{x}, \alpha_k, \alpha_{3-k}, \beta)^2} \right] \\ &+ \sum_{i=1}^n \frac{\alpha_{3-k} x_i^{\alpha_{3-k}} \left[(\beta^{\alpha_k} (\alpha_{3-k} + \alpha_k (\alpha_k + \alpha_{3-k}) \log \beta - \alpha_k (\alpha_k + \alpha_{3-k}) \log x_i)) x_i^{(\alpha_k + \alpha_{3-k})} \right]}{V(\underline{x}, \alpha_k, \alpha_{3-k}, \beta)^2}, \end{aligned}$$

and

$$\begin{aligned} a_{33} &= -E \left[\frac{\partial^2 \ln L(\eta)}{\partial \beta^2} \right] = \frac{n}{\beta^2} \sum_{k=1}^2 \alpha_k + 2 \sum_{k=1}^2 \sum_{i=1}^n E \left[\frac{\alpha_k \beta^{\alpha_k - 2} [(\alpha_k - 1) x_i^{\alpha_k} - \beta^{\alpha_k}]}{(x_i^{\alpha_k} + \beta^{\alpha_k})^2} \right] \\ &- \sum_{i=1}^n E \left[\frac{\alpha_k^2 \alpha_{3-k}^2 (\beta^{\alpha_k} x_i^{\alpha_{3-k}} + \beta^{\alpha_{3-k}} x_i^{\alpha_k})^2}{\beta^2 [V(\underline{x}, \alpha_k, \alpha_{3-k}, \beta)]^2} \right] \\ &+ \sum_{i=1}^n E \left[\frac{\alpha_k \alpha_{3-k} [(\alpha_{3-k} - 1) \beta^{\alpha_{3-k}} x_i^{\alpha_k} + (\alpha_k - 1) \beta^{\alpha_k} x_i^{\alpha_{3-k}}]}{\beta^2 [V(\underline{x}, \alpha_k, \alpha_{3-k}, \beta)]} \right], \end{aligned}$$

where for every function $g(y)$, the expected value is provided by

$$E[g(y)] = \int_0^{\infty} g(y) f(y, \alpha_1, \alpha_2, \beta) dy = 4\beta^{(\alpha_1 + \alpha_2)} \sum_{k=1}^2 \alpha_k \int_{\beta}^{\infty} \frac{g(y) y^{\alpha_k - 1}}{(y^{\alpha_k} + \beta^{\alpha_k})^2 (y^{\alpha_{3-k}} + \beta^{\alpha_{3-k}})} dy. \quad (57)$$

Moreover, when considering a single observation that was censored at the time of the i th failure, the Fisher information matrix is obtained as follows:

$$\begin{aligned} I_{W|X}^i(\Theta) &= -E_{Z_i|X_i} \left[\frac{\partial^2 \ln f_{Z|X}(z_i | z_i > x_i, \alpha_1, \alpha_2, \beta)}{\partial \Theta^2} \right]_{\Theta=(\theta, \beta, \lambda)} \\ &= \begin{bmatrix} b_{11}(x_i, \alpha_1, \alpha_2, \beta) & b_{12}(x_i, \alpha_1, \alpha_2, \beta) & b_{13}(x_i, \alpha_1, \alpha_2, \beta) \\ b_{21}(x_i, \alpha_1, \alpha_2, \beta) & b_{22}(x_i, \alpha_1, \alpha_2, \beta) & b_{23}(x_i, \alpha_1, \alpha_2, \beta) \\ b_{31}(x_i, \alpha_1, \alpha_2, \beta) & b_{32}(x_i, \alpha_1, \alpha_2, \beta) & b_{33}(x_i, \alpha_1, \alpha_2, \beta) \end{bmatrix}_{|\Theta=\hat{\Theta}}. \end{aligned} \quad (58)$$

wherein $f_{Z|X}(z_i | z_i > x_i, \alpha_1, \alpha_2, \beta)$ is given by Equation (43). It is therefore simple to retrieve the expected missing information as

$$I_{W|X}(\Theta) = \sum_{i=1}^m R_i I_{W|X}^i(\Theta), \quad (59)$$

where $I_{W|X}^i(\Theta)$ and $I_{W|X}^m(\Theta)$ are the information matrix of a single observation for the truncated NP distribution with left truncation at x .

For brevity and simplicity, assume that $F = f_{Z|X}(z_i | z_i > x_i, \alpha_1, \alpha_2, \beta)$. From Equations (7), (8), and (43), the logarithm of the PDF of the truncated NP distribution with left truncation at x_i can be listed as the PDF of the truncated NP distribution with left truncation at x_i , which can be listed as

$$\ln F = \ln \left[(\alpha_1 + \alpha_2) z_i^{(\alpha_1 + \alpha_2)} + \alpha_2 \beta^{\alpha_1 z_i \alpha_2} + \alpha_1 \beta^{\alpha_2 z_i \alpha_1} \right] - B + \sum_{k=1}^2 \ln(x_i^{\alpha_k} + \beta^{\alpha_k}), \quad (60)$$

where

$$B = \ln(z_i) + 2 \sum_{k=1}^2 \ln(x_i^{\alpha_k} + \beta^{\alpha_k}).$$

The negative expected value of the second partial derivatives with respect to α_1 , α_2 , and β are obtained by

$$\begin{aligned} b_{kk}(\cdot) &= -E \left[\frac{\partial^2 \ln F}{\partial \alpha_k^2} \right] = -\frac{x_i^{\alpha_k} \beta^{\alpha_k} (\ln x_i - \ln \beta)^2}{(x_i^{\alpha_k} + \beta^{\alpha_k})^2} + 2 \sum_{k=1}^2 E \left[\frac{z_{ij}^{\alpha_k} \beta^{\alpha_k} (\ln z_{ij} - \ln \beta)^2}{(z_{ij}^{\alpha_k} + \beta^{\alpha_k})^2} \right] \\ &+ E \left[\frac{z_i^{2\alpha_1} (z_{ij}^{\alpha_2} + \beta^{\alpha_2})^2 - z_i^{(\alpha_1 + \alpha_2)} \alpha_2 \beta^{\alpha_1} (z_{ij}^{\alpha_2} (\alpha_1 + \alpha_2) + \alpha_1 \beta^{\alpha_2}) \ln^2 z_{ij}}{[(\alpha_1 + \alpha_2) z_i^{(\alpha_1 + \alpha_2)} + \alpha_2 \beta^{\alpha_1 z_i \alpha_2} + \alpha_1 \beta^{\alpha_2 z_i \alpha_1}]^2} \right] \\ &+ E \left[\frac{2 z_i^{(\alpha_1 + \alpha_2)} \alpha_2 \beta^{\alpha_1} (z_{ij}^{\alpha_2} + \beta^{\alpha_2}) \ln \beta - z_i^{(\alpha_1 + \alpha_2)} \alpha_2 \beta^{\alpha_1} (z_{ij}^{\alpha_2} (\alpha_1 + \alpha_2) + \alpha_1 \beta^{\alpha_2}) \ln^2 z_{ij}}{[(\alpha_1 + \alpha_2) z_i^{(\alpha_1 + \alpha_2)} + \alpha_2 \beta^{\alpha_1 z_i \alpha_2} + \alpha_1 \beta^{\alpha_2 z_i \alpha_1}]^2} \right] \\ &+ E \left[\frac{2 z_i^{(\alpha_1 + \alpha_2)} \alpha_2 \beta^{\alpha_1} \ln z_{ij} (-z_{ij}^{\alpha_2} - \beta^{\alpha_2} + (z_{ij}^{\alpha_2} (\alpha_1 + \alpha_2) + \alpha_1 \beta^{\alpha_2}) \ln \beta)}{[(\alpha_1 + \alpha_2) z_i^{(\alpha_1 + \alpha_2)} + \alpha_2 \beta^{\alpha_1 z_i \alpha_2} + \alpha_1 \beta^{\alpha_2 z_i \alpha_1}]^2} \right], \end{aligned}$$

$$b_{k3}(\cdot) = b_{3k}(\cdot) = -E \left[\frac{\partial^2 \ln F}{\partial \alpha_k \partial \beta} \right] = -\frac{\beta^{\alpha_k-1} [x_i^{\alpha_k} + \beta^{\alpha_k} + \alpha_k x_i^{\alpha_k} (\ln z_{ij} - \ln \beta)]}{(x_i^{\alpha_k} + \beta^{\alpha_k})^2} \\ + E \left[\frac{\beta^{\alpha_k-1} [z_{ij}^{\alpha_k} + \beta^{\alpha_k} + \alpha_k z_{ij}^{\alpha_k} (\ln z_{ij} - \ln \beta)]}{(z_{ij}^{\alpha_k} + \beta^{\alpha_k})^2} \right], k = 1, 2, \\ b_{33}(\cdot) = -E \left[\frac{\partial^2 \ln F}{\partial \beta^2} \right] = -\frac{\alpha_1 \beta^{\alpha_1-2} [(\alpha_1 - 1) x_i^{\alpha_1} - \beta^{\alpha_1}]}{(x_i^{\alpha_1} + \beta^{\alpha_1})^2} + E \left[\frac{\alpha_1 \beta^{\alpha_1-2} [(\alpha_1 - 1) z_{ij}^{\alpha_1} - \beta^{\alpha_1}]}{(z_{ij}^{\alpha_1} + \beta^{\alpha_1})^2} \right].$$

The inverse of the observed Fisher information matrix $I_X(\Theta)$ at the maximum likelihood estimates $\hat{\Theta}_{EM} = (\hat{\Theta}_{1EM}, \hat{\Theta}_{2EM}, \hat{\Theta}_{EM})$ is the asymptotic variance–covariance matrix of the $\hat{\Theta}_{EM}$.

$$I_X^{-1}(\hat{\Theta}_{EM}) = [I_W(\hat{\Theta}_{EM}) - I_{W|X}(\hat{\Theta}_{EM})]^{-1}. \quad (61)$$

Accordingly, for $\Theta = (\alpha_1, \alpha_2, \beta)$, the $100(1 - \gamma)\%$ asymptotic confidence interval is then given by

$$\left(\hat{\alpha}_{1EM} \pm Z_{\gamma/2} \sqrt{\text{Var}(\hat{\alpha}_{1EM})} \right), \left(\hat{\alpha}_{2EM} \pm Z_{\gamma/2} \sqrt{\text{Var}(\hat{\alpha}_{2EM})} \right), \text{ and } \left(\hat{\beta}_{EM} \pm Z_{\gamma/2} \sqrt{\text{Var}(\hat{\beta}_{EM})} \right), \quad (62)$$

and the $100(1 - \gamma)\%$ log-transformed MLE confidence intervals of $\Theta = (\alpha_1, \alpha_2, \beta)$ through the EM algorithm are as follows:

$$\left[\hat{\Theta}_{EM} \exp \left(-Z_{(\gamma/2)} \sqrt{\frac{\text{Var}(\hat{\Theta}_{EM})}{\hat{\Theta}_{EM}^2}} \right), \hat{\Theta}_{EM} \exp \left(Z_{(\gamma/2)} \sqrt{\frac{\text{Var}(\hat{\Theta}_{EM})}{\hat{\Theta}_{EM}^2}} \right) \right], \hat{\Theta} = (\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}). \quad (63)$$

In using the delta method in the same way as presented in Section 3, the estimators of $S(t)$ and $H(t)$ can be obtained.

5. Bayesian Estimation

In this section, Bayesian estimates for the NP distributions are generated based on progressively censored competing risks samples, focusing on the unknown parameters α_1 , α_2 , and β , alongside the reliability function $S(t)$ and the hazard function $H(t)$. Estimations are derived considering three distinct loss functions: the general entropy loss function (GELF), LINEX loss function (LLF), and squared error loss function (SELF). While the LINEX and GE loss functions exhibit asymmetry, the squared error loss function is symmetric, affording equal weight to both over- and under-estimations. Given its prevalence in the Bayesian literature, the SELF is often chosen for its symmetry; however, in cases where over-estimation or under-estimation holds varying degrees of significance, asymmetric loss functions such as the LINEX come into play. In Bayesian estimation, prior distributions for parameters are essential, drawing from previous experiences and available parameter information. Since the NP distribution lacks a natural conjugate prior, and joint conjugate priors are not feasible for unknown parameters, independent gamma priors with hyperparameters (a_i, b_i) , $i = 1, 2, 3$, are recommended for all positive parameters in the model.

$$\pi_1(\alpha_1) \propto \alpha_1^{a_1-1} e^{-b_1 \alpha_1}, \pi_2(\alpha_2) \propto \alpha_2^{a_2-1} e^{-b_2 \alpha_2} \text{ and } \pi_3(\beta) \propto \beta^{a_3-1} e^{-b_3 \beta}; \alpha_1, \alpha_2, \beta > 0. \quad (64)$$

In this context, it is presumed that each hyper-parameter, denoted as (a_i, b_i) , $i = 1, 2, 3$, is known and non-negative. This assumption leads to the proposition of a specific prior distribution, representing a scenario where non-informative priors for the parameters are available. Consequently, the joint prior distribution of $(\alpha_1, \alpha_2, \beta)$ can be expressed as follows:

$$\pi(\alpha_1, \alpha_2, \beta) \propto \alpha_1^{a_1-1} \alpha_2^{a_2-1} \beta^{a_3-1} e^{-(b_1 \alpha_1 + b_2 \alpha_2 + b_3 \beta)}. \quad (65)$$

In combining the prior information from Equation (65) with the likelihood function presented in Equation (14), the joint posterior distribution can be represented as follows:

$$\begin{aligned} \pi^*(\alpha_1, \alpha_2, \beta | \underline{x}) &\propto \alpha_1^{a_1-1} \alpha_2^{a_2-1} \beta^{a_3-1} e^{-(b_1\alpha_1+b_2\alpha_2+b_3\beta)} \prod_{i=1}^m \prod_{k=1}^2 \left\{ \frac{\alpha_k \beta^{\alpha_k} \beta^{\alpha_{3-k}} x_i^{\alpha_k-1}}{(x_i^{\alpha_k} + \beta^{\alpha_k})^2 (x_i^{\alpha_{3-k}} + \beta^{\alpha_{3-k}})} \right\}^{I(\delta_i=k)} \\ &\times \prod_{i=1}^m \prod_{k=1}^2 \left\{ \frac{\beta^{\alpha_k}}{(x_i^{\alpha_k} + \beta^{\alpha_k})} \right\}^{R_i}. \end{aligned} \quad (66)$$

The Bayes estimator of any function of α_1, α_2 , and β , say $\Phi(\alpha_1, \alpha_2, \beta)$ under the SE, LINEX, and GE loss functions, can be expressed as

$$\hat{\Phi}_{BS} = E(\Phi(\alpha_1, \alpha_2, \beta) | \underline{x}) = \frac{\int_{x_1}^{\infty} \int_0^{\infty} \int_0^{\infty} \Phi(\alpha_1, \alpha_2, \beta) \pi^*(\alpha_1, \alpha_2, \beta | \underline{x}) d\alpha_1 d\alpha_2 d\beta}{\int_{x_1}^{\infty} \int_0^{\infty} \int_0^{\infty} \pi^*(\alpha_1, \alpha_2, \beta | \underline{x}) d\alpha_1 d\alpha_2 d\beta}, \quad (67)$$

$$\hat{\Phi}_{LINEX} = -\frac{1}{c} \log \frac{\int_{x_1}^{\infty} \int_0^{\infty} \int_0^{\infty} e^{-c\Phi(\alpha_1, \alpha_2, \beta)} \pi^*(\alpha_1, \alpha_2, \beta | \underline{x}) d\alpha_1 d\alpha_2 d\beta}{\int_{x_1}^{\infty} \int_0^{\infty} \int_0^{\infty} \pi^*(\alpha_1, \alpha_2, \beta | \underline{x}) d\alpha_1 d\alpha_2 d\beta}, \quad (68)$$

and

$$E_{\alpha_1, \alpha_2, \beta | \underline{x}} (g(\alpha_1, \alpha_2, \beta)^{-q}) = \frac{\int_{x_1}^{\infty} \int_0^{\infty} \int_0^{\infty} \Phi(\alpha_1, \alpha_2, \beta)^{-q} \pi^*(\alpha_1, \alpha_2, \beta | \underline{x}) d\alpha_1 d\alpha_2 d\beta}{\int_{x_1}^{\infty} \int_0^{\infty} \int_0^{\infty} \pi^*(\alpha_1, \alpha_2, \beta | \underline{x}) d\alpha_1 d\alpha_2 d\beta}. \quad (69)$$

Numerically solving the Bayes estimators becomes necessary as they are not explicitly obtained under the SE or LINEX loss functions. Hence, we propose deriving the Bayes estimators of α_1, α_2 , and β using the MCMC method. Initially, the fully conditional posterior distribution $\pi^*(\alpha_k | \underline{x}, \beta)$ of α_k is

$$\pi^*(\alpha_k | \underline{x}, \beta) \propto \alpha_k^{a_k+m_k-1} \beta^{n\alpha_k} e^{-b_k\alpha_k} \prod_{i=1}^m x_i^{(\alpha_k-1)I(\delta_i=k)} \prod_{i=1}^m (x_i^{\alpha_k} + \beta^{\alpha_k})^{-(\Delta_k(\delta_i, R_i))}, k = 1, 2, \quad (70)$$

and the fully conditional posterior distribution $\pi^*(\beta | \underline{x}, \alpha_1, \alpha_2)$ of β is

$$\begin{aligned} \pi^*(\beta | \underline{x}, \alpha_1, \alpha_2) &\propto \beta^{a_3-1+n(\alpha_k+\alpha_{3-k})} e^{-b_3\beta} \prod_{i=1}^m \prod_{k=1}^2 \left\{ \frac{x_i^{\alpha_k-1}}{(x_i^{\alpha_k} + \beta^{\alpha_k})^2 (x_i^{\alpha_{3-k}} + \beta^{\alpha_{3-k}})} \right\}^{I(\delta_i=k)} \\ &\times \prod_{i=1}^m \prod_{k=1}^2 \left\{ \frac{1}{(x_i^{\alpha_k} + \beta^{\alpha_k})} \right\}^{R_i}. \end{aligned} \quad (71)$$

The absence of closed-form solutions for the three posterior distributions necessitates the utilization of the Metropolis–Hastings algorithm to derive our Bayes estimators from posterior samples. The Metropolis–Hastings (M-H) algorithm is a valuable tool for generating random samples from the posterior distribution by leveraging a proposal density, as detailed in [21,22]. The Algorithm 1 proceeds through the following stages:

Algorithm 1 MCMC algorithm

- Step 1:** Choose an initial guess of $(\alpha_1, \alpha_2, \beta)$ denoted by $(\alpha_1^{(0)}, \alpha_2^{(0)}, \beta^{(0)})$, and set $i = 1$.
- Step 2:** Using the following Metropolis–Hustings, generate $\beta^{(i)}$ from π^* , with the normal proposal distribution, $N(\beta^{(i-1)}, Var(\beta))$, where $(\beta^{(i-1)})$ is the current value of β , and $Var(\beta)$ is a variance of β . Perform the following:
- Generate a proposal β^* from $N(\beta^{(i-1)}, Var(\beta))$;
 - Evaluate the acceptance probability $\rho = \min\left\{1, \frac{\pi^*(\beta^*|\alpha_1^{(i-1)}, \alpha_2^{(i-1)}, \underline{x})}{\pi^*(\beta^{(i-1)}|\alpha_1^{(i-1)}, \alpha_2^{(i-1)}, \underline{x})}\right\}$;
 - Generate a u from a Uniform(0, 1) distribution;
 - If $u \leq \rho$, accept the proposal and set $\beta^{(i)} = \beta^*$; else, set $\beta^{(i)} = \beta^{(i-1)}$.
- Step 3:** In the same way as in the previous step, generate $\alpha_k^{(i)}$ from $\pi^*(\alpha_k^{(i-1)}|\underline{x}, \beta^{(i)})$ with the normal proposal distribution $N(\alpha_k^{(i-1)}, Var(\alpha_k))$, where $(\alpha_k^{(i-1)})$ is the current value of α_k , and $Var(\alpha_k)$ is a variance of α_k .
- Step 4:** For given t , compute the reliability and hazard functions:

$$S(t, \alpha_1^{(i)}, \alpha_2^{(i)}, \beta^{(i)}) = 4 \prod_{j=1}^2 \left[\frac{\beta^{(i)\alpha_k^{(i)}}}{t^{\alpha_k^{(i)}} + \beta^{(i)\alpha_k^{(i)}}} \right],$$

and

$$H(t, \Theta) = \sum_{k=1}^2 \left[\frac{\beta^{(i)\alpha_k^{(i)}}}{t^{\alpha_k^{(i)}} + \beta^{(i)\alpha_k^{(i)}}} \right].$$

- Step 5:** Set $i = i + 1$.
- Step 6:** Repeat steps (2 – 4) N times and obtain the desired number of samples. After discarding the first M burn-in samples, the remaining $N - M$ samples are used to obtain the Bayesian estimates.
- Step 7:** It is now possible to calculate the Bayes estimate of $\Phi = \Phi(\alpha_1, \alpha_2, \beta)$ under the SE, LINEX, and GE loss functions as

$$\hat{\Phi}_S = \frac{1}{N - M} \sum_{i=M+1}^N \Phi(\alpha_1^{(i)}, \alpha_2^{(i)}, \beta^{(i)}),$$

$$\hat{\Phi}_{LINEX} = \frac{-1}{c} \log \left[\frac{1}{N - M} \sum_{i=M+1}^N e^{-c\Phi(\alpha_1^{(i)}, \alpha_2^{(i)}, \beta^{(i)})} \right],$$

and

$$\hat{\beta}_{GE} = \left[\frac{1}{N - M} \sum_{i=M+1}^N \Phi(\alpha_1^{(i)}, \alpha_2^{(i)}, \beta^{(i)})^{-q} \right]^{-\frac{1}{q}},$$

where $\Phi = \Phi(\alpha_1, \alpha_2, \beta)$ denotes parameters $\alpha_1, \alpha_2, \beta$, $S(t)$, and $H(t)$.

- Step 8:** Order $\Phi^{(M+1)}, \Phi^{(M+2)}, \dots, \Phi^{(N)}$ as $\Phi_{(1)} < \Phi_{(2)} < \dots < \Phi_{(N-M)}$. Consequently,

$$\left(\Phi_{[(N-M)\alpha/2]}, \Phi_{[(N-M)(1-\alpha/2)]} \right)$$

yields the $100(1 - \gamma)\%$ Bayesian credible interval of Φ . In this case, $[q]$ stands for the integer part of q .

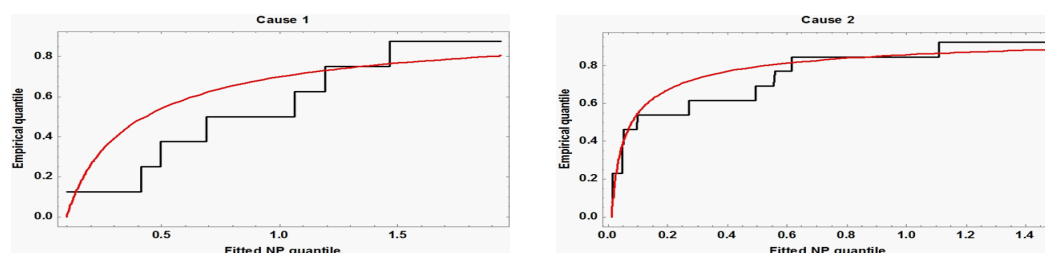
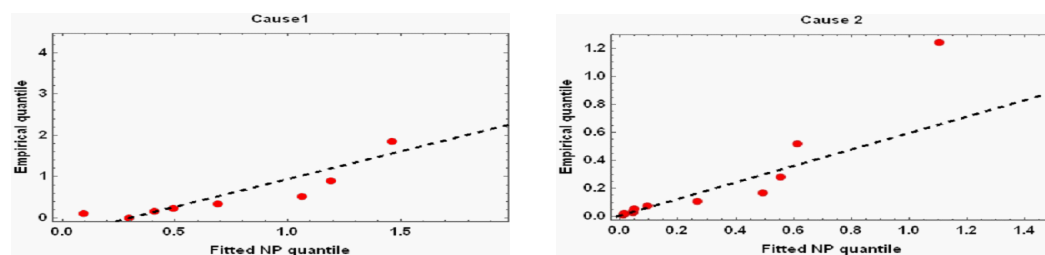
6. Actual Data Illustration

The main objective of this part is to illustrate how the recommended techniques may be used in real-world situations. The first example includes the failure times of electrical appliances. The source of this dataset is listed in [23]. This real data collection, which [24] first looked at, contains 18 electrical equipment failure times. Only failure mode 11 is seen to occur more than twice out of all the failure modes found in these data. In this article, failure mode 11 was regarded as reason number one (cause-1), and the other failure modes as cause number two (cause-2)). In this instance, we treat X_1 and X_2 as failing because of two competing risks under cause-1 and cause-2, respectively. It is observed that there are 8 failures attributed to cause-1 and 13 failures attributed to cause-2. Data analysis begins with dividing the original set of data by 100. Since this transformation is only a scale change, it will not have an impact on this study's results. Table 1 displays the electrical appliance survival times results.

Table 1. Electrical appliance survival times.

Cause1	0.98	4.13	4.95	6.92	10.65	11.93	14.67	19.37
Cause2	0.12	4.95	0.16	5.57	0.16	6.16	0.46	11.07
	0.46	14.67	0.52	0.98	2.70			

We first look at whether or not these datasets can be analyzed using the NP distribution before moving on. It is commonly known that the Kolmogorov–Smirnov (K–S) test may be used with both extremely small and large samples. The Kolmogorov–Smirnov distances and the accompanying p -values (with bracket) for cause-1 and -2 are, respectively, 0.3637 (0.2405) and 0.1792 (0.7981). Because both causes' associated p -values are clearly higher than the significance level of $\gamma = 0.05$, we are unable to reject the null hypothesis that the data for electronic applications originate from the NP distribution. Figure 2 displays the plot of the empirical vs. the fitted CDFs. Furthermore, the most widely used graphical approach for model validation is quantile–quantile (Q–Q), which determines if the fitted model agrees with the provided data. Let $\hat{F}(x)$ be the estimated value of $F(x)$ given $x_1, x_2, \dots, x_n$. Point $\hat{F}(i/n)$ vs. $x_{i:n}$, where $i = 1, 2, \dots, n$, was plotted as a Q–Q scatter diagram. The estimated and observed quantiles are displayed in a Q–Q plot. The points on the Q–Q plot should demonstrate a 45 straight line if the model matches the data well. Figure 3 displays the appliance failure data together with the quantile–quantile (Q–Q) plot. As can be seen in Figure 3, there is a rough straight-line pattern that suggests a good match for the NP model. Furthermore, one crucial graphical technique for determining whether or not the data may be applied to a particular distribution is the total time test (TTT) plot. The TTT plot's empirical form is provided by $\tau(k/n) = [\sum_{i=1}^k x_{i:n} - (n-k)x_{k:n}] / (\sum_{i=1}^n x_i)$, $k = 1, 2, \dots, n$, where $x_{i:n}$ denotes the sample's order statistics. In a graphic representation of the TTT plot, the HRF is growing (or decreasing) if the TTT plot is concave (or convex), and it is constant if the TTT plot is graphically portrayed as a straight diagonal. When the TTT plot is convex and then concave, the HRF is U-shaped; otherwise, it is unimodal. Figure 4 shows the TTT plot for the two datasets. The failure data's total time on the test (TTT) plot in Figure 4 illustrates the increasing and decreasing hazard rate for cause-1 and cause-2, respectively, which is consistent with the NP distribution.

**Figure 2.** Fitted and empirical CDFs associated with NP distribution for two causes.**Figure 3.** Q–Q plots associated with NP distribution for two causes.

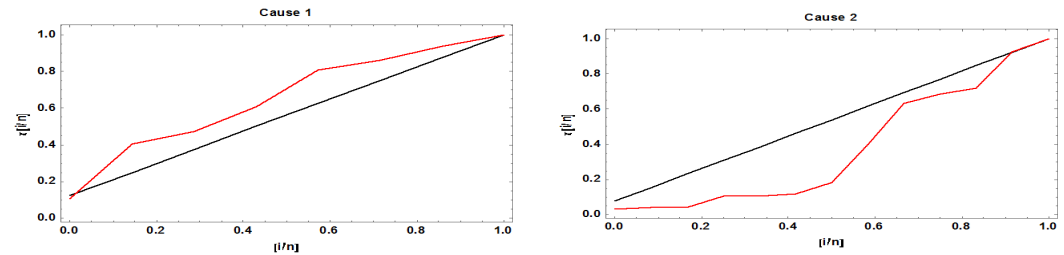


Figure 4. T-T plots associated with NP distribution for two causes.

Three groups of progressive Type-II samples are formed based on the total failure data from Table 1 and are displayed in Table 2. These generated samples are used to estimate the MLEs and Bayes estimates of $\alpha_1, \alpha_2, \beta, S(t)$, and $H(t)$ together with their standard errors (in parenthesis).

Table 2. Progressive Type-II data for electrical appliance data, with $n = 21$ and $m = 15$.

Sc. I	$(6, 0^{14})$	$(0.12, 2)$	$(0.16, 2)$	$(0.16, 2)$	$(0.46, 2)$	$(0.46, 2)$
		$(0.52, 2)$	$(0.98, 2)$	$(0.98, 1)$	$(2.70, 2)$	$(4.95, 2)$
		$(55.7, 2)$	$(61.6, 2)$	$(10.65, 1)$	$(11.93, 1)$	$(14.67, 2)$
Sc. II	$(1^6, 0^9)$	$(0.12, 2)$	$(0.16, 2)$	$(0.46, 2)$	$(0.46, 2)$	$(0.52, 2)$
		$(0.98, 2)$	$(0.98, 1)$	$(2.70, 2)$	$(4.13, 1)$	$(4.95, 2)$
		$(4.95, 1)$	$(6.92, 1)$	$(11.93, 1)$	$(14.67, 2)$	$(14.67, 1)$
Sc. III	$(0^{14}, 6)$	$(0.12, 2)$	$(0.16, 2)$	$(0.16, 2)$	$(0.46, 2)$	$(0.46, 2)$
		$(0.52, 2)$	$(0.98, 2)$	$(0.98, 1)$	$(2.70, 2)$	$(4.13, 1)$
		$(4.95, 2)$	$(4.95, 1)$	$(5.57, 2)$	$(6.16, 2)$	$(10.65, 1)$

In real-world data analysis, choosing the starting parameters for an NR algorithm is a challenging task. For Sc. I, the initial guess of β is taken to be the x_1 , and a graphical approach presented in [25] is utilized to calculate the MLE shape parameters α_1 and α_2 , where, from Equation (19), the curves of $\frac{1}{\alpha_k}$ and $\Phi(\alpha_k)$ are plotted in Figure 5. According to Figure 5, for $k = 1, 2$, the crossing points of the two functions ($\frac{1}{\alpha_k}, \Phi(\alpha_k)$) are around 0.1543 and 0.4605, respectively. To acquire the MLEs of α_1 and α_2 , we thus recommend starting the iteration with $\alpha_1 = 0.15$ and $\alpha_2 = 0.46$ as initial values. Moreover, a contour plot of the log-likelihood function of α_1 and α_2 with respect to $\beta = x_1$ is shown in Figure 6.

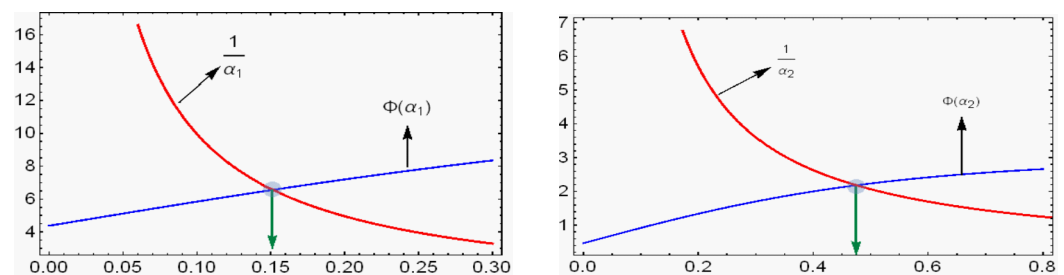


Figure 5. The MLE shape parameters with NP distribution.

Based on these initial values, we first calculated the MLEs of $\alpha_1, \alpha_2, \beta, S(t)$, and $H(t)$ using the NR approach. Table 3 presents the values of the MLEs and estimated standard errors of the MLEs of $\alpha_1, \alpha_2, \beta, S(t)$, and $H(t)$, where $t = 0.4$ is the calculation point for the reliability function $S(t)$ and the hazard rate $H(t)$ function. We also analyzed the dataset (Sc. I) using the EM and SEM methods developed in Section 4. The NR approach is used to set the starting values of α_1, α_2 and β for the EM and SEM algorithms as related MLEs. Table 3 also includes a list of all point estimation results for the MLEs using the EM and SEM methods, along with estimated standard errors. It can be seen from the standard errors that the estimations produced by the NR approach are often higher than those produced

by the EM or SEM algorithms. Next, the 95% asymptotic confidence intervals of α_1 , α_2 , β , $S(t)$, and $H(t)$ were computed using the asymptotic normality of the MLEs and missing information principal techniques.

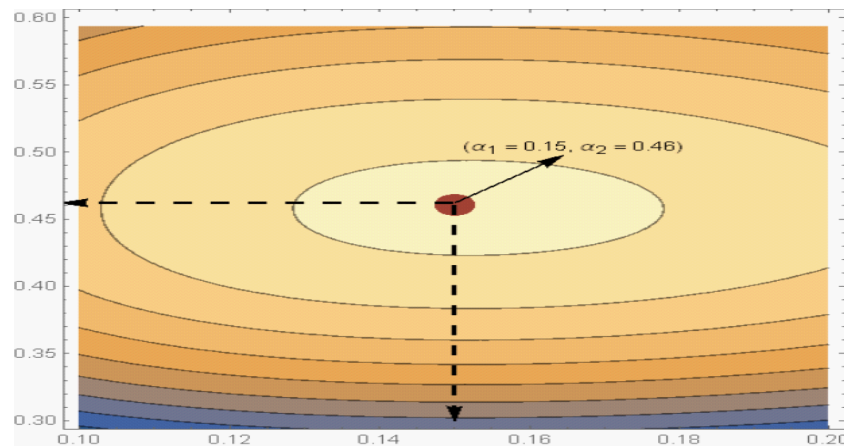


Figure 6. Contour plot of the log-likelihood function of α_1 and α_2 .

The next step would be to compute the Bayes estimates of α_1 , α_2 , β , $S(t)$, and $H(t)$ against the SE, LINEX, and GE loss functions using the MCMC samples. Selecting the hyper-parameters in a real dataset while using Bayesian estimating is a very difficult operation since the true value is not known in advance. Hence, we combine the non-informative prior assumptions ($a_i = b_i = 0, i = 1, 2, 3$) with the MCMC approach to build Bayes estimators. The initial assumptions for executing the MCMC algorithm are thought to be the MLEs estimations of α_1 , α_2 , and β . In employing 30,000 posterior sample points and ignoring the burn-in of the first 5000 times, the Bayes estimates of α_1 , α_2 , β , $S(t)$, and $H(t)$ are evaluated using the M-H method. A proposed density function that follows a normal distribution may be built in order to use the MCMC approaches, using the variance–covariance matrix and mean equal to the MLEs of the unknown parameters.

Table 3. ML and Bayesian point estimates (first row) and their standard errors (second row) of α_1 , α_2 , β , $S(t)$, and $H(t)$ for the real data.

Scheme	Parameter	MLE				Bayes			
		NR	EM	SEM	SEL	Loss Function			
						LINEX		GEL	
						$c = -2$	$c = 2$	$q = -2$	$c = 2$
Sc. I: $(6, 0^{14})$	α_1	0.1523	0.1111	0.1091	0.1501	0.1510	0.1498	0.1513	0.1466
		(0.0218)	(0.0061)	(0.0065)	(0.0048)	(0.0521)	(0.0072)	(0.0014)	(3.1353)
	α_2	0.4638	0.4917	0.4835	0.4567	0.4650	0.4534	0.4603	0.4458
		(0.0356)	(0.0319)	(0.0291)	(0.0146)	(0.7900)	(0.0117)	(0.0137)	(0.3379)
	β	0.1200	0.1200	0.1200	0.1141	0.1142	0.1140	0.1143	0.1139
		(0.0146)	(0.0143)	(0.0133)	(0.0009)	(0.0077)	(0.0014)	(0.0002)	(1.1689)
	$S(t = 0.5)$	0.6070	0.6104	0.6162	0.6006	0.6038	0.5994	0.6017	0.5974
		(0.0224)	(0.0237)	(0.0217)	(0.0092)	(0.9349)	(0.0055)	(0.0110)	(0.0874)
	$H(t = 0.5)$	0.7807	0.7773	0.7616	0.7742	0.7987	0.7650	0.7802	0.7561
		(0.0611)	(0.0474)	(0.0432)	(0.0247)	(0.7132)	(0.0105)	(0.0392)	(0.1149)

Table 3. Cont.

Scheme	Parameter	MLE				Bayes			
		NR	EM	SEM	SEL	Loss Function			
						LINEX		GEL	
						$c = -2$	$c = 2$	$q = -2$	$c = 2$
Sc. II: $(1^6, 0^9)$	α_1	0.2318	0.1722	0.1834	0.1560	0.1569	0.1557	0.1572	0.1527
		(0.0234)	(0.0093)	(0.0118)	(0.0048)	(0.0537)	(0.0071)	(0.0015)	(2.7634)
	α_2	0.3056	0.3561	0.3782	0.4108	0.4181	0.4080	0.4142	0.4005
		(0.0263)	(0.0204)	(0.0241)	(0.0136)	(0.5889)	(0.0119)	(.0115)	(0.4275)
	β	0.1200	0.1200	0.1200	0.1156	0.1156	0.1155	0.1156	0.1154
		(0.0157)	(0.0157)	(0.0137)	(0.0009)	(0.0077)	(0.0014)	(0.0002)	(1.1245)
	$S(t = 0.5)$	0.6566	0.6594	0.6406	0.6276	0.6304	0.6265	0.6285	0.6249
		(0.0213)	(0.0246)	(0.0210)	(0.0085)	(0.9958)	(0.0049)	(0.0107)	(0.0722)
	$H(t = 0.5)$	0.6411	0.6378	0.6851	0.7066	0.7268	0.6992	0.7119	0.6910
		(0.0492)	(0.0306)	(0.0369)	(0.0223)	(0.9597)	(0.0107)	(0.0323)	(0.1335)
Sc. III: $(0^{14}, 6)$	α_1	0.1290	0.2161	0.2236	0.1481	0.1489	0.1478	0.1492	0.1447
		(0.0160)	(0.0113)	(0.0158)	(0.0047)	(0.0497)	(0.0069)	(0.0014)	(3.1217)
	α_2	0.2950	0.2182	0.2261	0.3992	0.4054	0.3967	0.4022	0.3900
		(0.0227)	(0.0114)	(0.0165)	(0.0130)	(0.5229)	(0.0116)	(0.0106)	(0.4435)
	β	0.1200	0.1200	0.1200	0.1157	0.1157	0.1157	0.1157	0.1155
		(0.0170)	(0.0181)	(0.0155)	(0.0009)	(0.0076)	(0.0014)	(0.0002)	(1.1088)
	$S(t = 0.5)$	0.7198	0.7162	0.7071	0.6389	0.6413	0.6378	0.6396	0.6364
		(0.0206)	(0.0250)	(0.0204)	(0.0083)	(1.0171)	(0.0047)	(0.0105)	(0.0661)
	$H(t = 0.5)$	0.4971	0.5010	0.5212	0.6784	0.6953	0.6721	0.6832	0.6644
		(0.0367)	(0.0204)	(0.0296)	(0.0211)	(0.96918)	(0.0107)	(0.0293)	(0.1419)

It is not a secret that, for the LINEX loss function, overestimation has a larger penalty than underestimation when $c > 0$, and the opposite is true when $c < 0$. Additionally, when c approaches 0, the LINEX loss function becomes symmetric and behaves somewhat like the SE loss function. The Bayes estimates were produced under the LINEX loss function discussed here for two different values of $c = -5$ and $+2$. Furthermore, in the general entropy (GE) loss function, the values of parameter q were chosen to be -2 and $+2$. The point Bayes estimates for α_1 , α_2 , β , $S(t)$, and $H(t)$ were been computed and are presented in Table 3, along with the estimated standard errors and the corresponding 95% credible ranges in Table 4.

Table 4. The 95% interval estimates of α_1 , α_2 , β , $S(t)$, and, $H(t)$ for the real data.

Scheme	Method	α_1	α_2	β	$S(t = 0.5)$	$H(t = 0.5)$
Sc. I: $(6, 0^{14})$	ACI	(−0.0127, 0.3172)	(0.1936, 0.7340)	(0.0094, 0.2306)	(0.4370, 0.7768)	(0.3165, 1.2448)
	LACI	(0.0515, 0.4498)	(0.2590, 0.8305)	(0.0477, 0.3016)	(0.4588, 0.8030)	(0.4308, 1.4147)
	EM	(0.0730, 0.1689)	(0.0679, 0.8047)	(0.0486, 0.2963)	(0.4548, 0.8194)	(0.4891, 1.2353)
	SEM	(0.0690, 0.1727)	(0.3062, 0.7636)	(0.0519, 0.2775)	(0.4717, 0.8047)	(0.4949, 1.1718)
	HPD	(0.1149, 0.1892)	(0.3496, 0.5749)	(0.1077, 0.1208)	(0.5305, 0.6692)	(0.5986, 0.9751)
Sc. II: $(1^6, 0^9)$	ACI	(0.0535, 0.4101)	(0.1064, 0.5049)	(0.0010, 0.2390)	(0.4951, 0.8180)	(0.2679, 1.0142)
	LACI	(0.1074, 0.5002)	(0.1592, 0.5866)	(0.0445, 0.3235)	(0.5135, 0.8396)	(0.3582, 1.1474)
	EM	(0.1144, 0.2591)	(0.1114, 0.5502)	(0.0445, 0.3235)	(0.4969, 0.8751)	(0.4430, 0.9185)
	SEM	(0.1124, 0.2993)	(0.2332, 0.6136)	(0.0505, 0.2852)	(0.4995, 0.8216)	(0.4550, 1.0317)
	HPD	(0.1206, 0.1949)	(0.3145, 0.5210)	(0.1090, 0.1221)	(0.5594, 0.6896)	(0.5536, 0.8917)

Table 4. Cont.

Scheme	Method	α_1	α_2	β	$S(t = 0.5)$	$H(t = 0.5)$
Sc. III: $(0^{14}, 6)$	ACI	(0.0068, 0.2511)	(0.12215, 0.4679)	(−0.0089, 0.2489)	(0.5637, 0.8758)	(0.2186, 0.7755)
	LACI	(0.0500, 0.3325)	(0.1642, 0.5301)	(0.0410, 0.3514)	(0.5795, 0.894)	(0.2839, 0.8704)
	EM	(0.1452, 0.3216)	(0.1452, 0.3247)	(0.0382, 0.3770)	(0.5495, 0.9334)	(0.3673, 0.6833)
	SEM	(0.1306, 0.3829)	(0.1299, 0.3934)	(0.0449, 0.3204)	(0.5678, 0.8805)	(0.3389, 0.8016)
	HPD	(0.1144, 0.1860)	(0.3082, 0.5035)	(0.1092, 0.1223)	(0.5743, 0.6981)	(0.535, 0.8502)

To assess the convergence of the Markov chain Monte Carlo (MCMC) method for the studied dataset, we present density plots as well as trace plots of the MCMC outputs for the parameters α_1 , α_2 , β , $S(t = 0.4)$, and $H(t = 0.4)$ in Figure 7. These graphs demonstrate the good convergence of all the parameters considered for different chains. Furthermore, it appears that the samples come from the same posterior density because all of the plots display a very strong overlap of the density plots for different chains. Moreover, as Figure 6 illustrates, the estimated estimations of α_1 , α_2 , β , $S(t = 0.4)$, and $H(t = 0.4)$ for each sample were typically symmetrical. Thus, the trace and density graphs appear to indicate that the MCMC approach has a high degree of convergence.

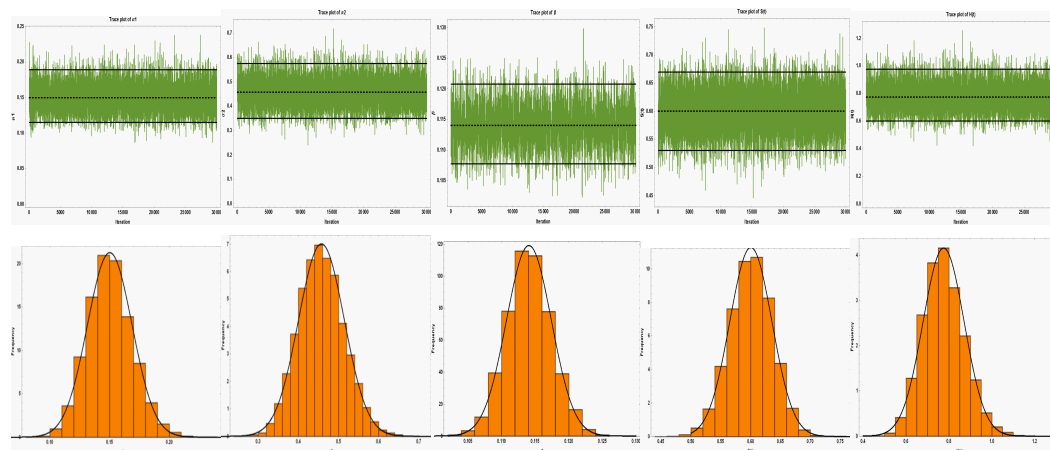


Figure 7. MCMC trace plot (first row) and histogram (second row) of α_1 , α_2 , β , $S(t)$, and $H(t)$ for electrical appliance dataset. Dashed lines (---) represent the posterior means and solid lines (—) represent the lower and upper bounds of the 95% probability interval.

The results are shown in Tables 3 and 4, where the recommended Bayes estimates outperform the most frequent estimates in terms of the lowest standard errors. Moreover, the HPD credible intervals estimates outperform the ACIs, LACIs, EM, and SEM estimates with respect to the shortest ILs. This example also shows how similar the results of all the traditional estimations are to one another. However, it should be mentioned that the MLEs that employed the SEM approach had the lowest standard errors. As a result, machine learning estimates obtained from SEM algorithms frequently performed better than estimates obtained from the NR and EM techniques. Furthermore, it is important to keep in mind that the techniques utilized to compute the point and interval estimators based on Sc. I were also applied to Sc. II and Sc. III; Tables 4 and 5 present the results.

Table 5. The average confidence lengths (AL) and the corresponding coverage percentages (CP) of α_1 , α_2 , β , $S(t)$, and $H(t)$ based on ML and Bayes estimates for censoring Scheme I.

(n, m)		MLE								MCMC			
		Normal		Log-Normal		EM		SEM		NIP		IP	
		AL	CP	AL	CP	AL	CP	AL	CP	AL	CP	AL	CP
(30, 20)	α_1	1.6237	0.813	2.1950	0.875	0.9999	0.859	0.8271	0.879	0.6282	0.821	0.3709	0.953
	α_2	1.8054	0.829	2.1247	0.894	1.1388	0.894	0.9577	0.893	0.6601	0.893	0.3912	0.962
	β	1.1514	0.939	0.9871	0.907	0.6561	0.993	0.5441	0.965	0.1149	0.955	0.0888	0.962
	$S(t)$	0.6043	0.931	0.4201	0.918	0.3986	0.974	0.3984	0.993	0.2481	0.887	0.1547	0.955
	$H(t)$	1.7221	0.814	2.4435	0.945	1.1505	0.957	1.0194	0.989	0.6702	0.821	0.3983	0.947
	α_1	1.3480	0.958	1.7874	0.958	0.9596	0.942	0.7644	0.930	0.6000	0.921	0.3610	0.951
	α_2	1.4875	0.959	2.0442	0.967	1.0825	0.975	0.9265	0.925	0.6509	0.932	0.3891	0.973
	β	0.9636	0.967	0.9715	0.981	0.6422	0.992	0.5230	0.975	0.1086	0.940	0.0818	0.962
	$S(t)$	0.5822	0.919	0.4166	0.964	0.3851	0.981	0.3891	0.961	0.2344	0.937	0.1510	0.964
	$H(t)$	1.3756	0.934	2.1249	0.972	1.0145	0.946	0.9082	0.958	0.6618	0.939	0.3929	0.947
(50, 35)	α_1	1.2839	0.918	1.4703	0.92	0.7550	0.976	0.6222	0.975	0.5180	0.933	0.3367	0.963
	α_2	1.4564	0.933	1.6739	0.913	0.8297	0.988	0.6838	0.981	0.5379	0.932	0.3562	0.957
	β	0.7670	0.941	0.8492	0.987	0.4937	0.982	0.4152	0.992	0.0658	0.941	0.0537	0.951
	$S(t)$	0.3253	0.939	0.3310	0.933	0.3032	0.979	0.3018	0.998	0.1907	0.954	0.1299	0.959
	$H(t)$	1.5651	0.925	1.6949	0.913	0.8331	0.990	0.7230	0.986	0.5360	0.955	0.3619	0.975
(50, 45)	α_1	0.8591	0.984	0.8793	0.982	0.7458	0.98	0.5786	0.927	0.4609	0.962	0.3255	0.960
	α_2	0.9730	0.991	0.9849	0.973	0.7969	0.982	0.6612	0.960	0.4832	0.947	0.3414	0.933
	β	0.5536	0.956	0.5688	0.988	0.4890	0.963	0.4065	0.971	0.0640	0.956	0.0513	0.960
	$S(t)$	0.3050	0.950	0.3090	0.961	0.3021	0.955	0.3029	0.965	0.1754	0.963	0.1225	0.967
	$H(t)$	1.0225	0.980	0.9891	0.979	0.8227	0.968	0.6432	0.968	0.4920	0.953	0.3474	0.968
(80, 60)	α_1	0.8289	0.982	0.8434	0.981	0.5849	0.967	0.4650	0.945	0.3932	0.923	0.2984	0.955
	α_2	0.9331	0.969	0.9519	0.973	0.6289	0.952	0.5183	0.936	0.4051	0.920	0.3142	0.946
	β	0.5308	0.990	0.5584	0.980	0.3909	0.967	0.3256	0.976	0.0398	0.979	0.0349	0.954
	$S(t)$	0.2513	0.964	0.2535	0.967	0.2408	0.982	0.2397	0.987	0.1462	0.942	0.1120	0.972
	$H(t)$	0.9486	0.965	0.9531	0.972	0.6359	0.980	0.5298	0.973	0.4116	0.963	0.3193	0.969
(80, 75)	α_1	0.6538	0.981	0.6768	0.945	0.5681	0.962	0.4429	0.973	0.3748	0.929	0.2913	0.962
	α_2	0.7093	0.957	0.7327	0.980	0.6158	0.957	0.4813	0.964	0.3895	0.937	0.3038	0.952
	β	0.4266	0.966	0.4394	0.982	0.3827	0.981	0.3171	0.958	0.0362	0.963	0.0337	0.967
	$S(t)$	0.2441	0.947	0.2461	0.971	0.2378	0.964	0.2396	0.972	0.1356	0.965	0.1066	0.960
	$H(t)$	0.7276	0.993	0.7456	0.994	0.6214	0.989	0.4642	0.987	0.3995	0.971	0.3128	0.946

7. Simulation Study

This section compares the performances of suggested estimate methods under progressive Type-II censoring using Monte Carlo simulations. A Monte Carlo simulation analysis was carried out using the Mathematica version 11 statistical program. A comparison is made between point estimators for competing risk lifetimes parameters based on the following perspectives:

(i) Bias = $\frac{1}{N_s} \sum_{i=1}^{N_s} |\Theta_i - \hat{\Theta}_i|$, where Θ_i and $\hat{\Theta}_i$ stand for the unknown parameters and the associated estimations, and N_s is the number of simulation repeats. The higher agreement of the experimental data with the estimated model is shown by the smaller value of the average bias.

(ii) Mean squared error (MSE) = $\frac{1}{N_s} \sum_{i=1}^{N_s} (\Theta_i - \hat{\Theta}_i)^2$. An improved estimate performance is shown by a lower MSE value. Each and every result is derived from 1000 replications.

Additionally, average confidence lengths (ACLs) (a better interval estimate performance is correlated with a smaller width) and coverage percentages (CPs) (the probability that the real parameter values lie within the range of the interval estimations) are used to evaluate interval estimators of asymptotic and HPD intervals. If the CP is 95%, then a confidence interval estimator works well. The significance threshold that we employed was $\gamma = 0.05$. We created progressive Type-II censored competing risk data, where the parameters' true values of the NP distribution are arbitrarily assumed to be $(\alpha_1, \alpha_2, \beta) = (0.7, 0.8, 0.5)$. The hazard rate functions in Figure 1 were used to determine the values of α_1 , α_2 , and β . In order to conduct this simulation research, we chose three distinct censoring schemes, which are detailed below, for $(n, m) = (30, 20), (30, 25), (50, 35), (50, 45)$, and $(80, 75)$:

Sc. I: $R_1 = R_2 = \dots = R_{m-1} = 0$ and $R_m = n - m$;

Sc. II: $R_1 = n - m$ and $R_2 = R_3 = \dots = R_{m-1} = 0$;

Sc. III: $R_1 = R_2 = \dots = R_{m-1} = 1$ and $R_{n-m+1} = \dots = R_m = 0$.

Here, Sc. I is equivalent to the conventional Type-II censoring technique, and Sc. II is Type-I censoring. After the samples are created, we use the NR, EM, and SEM algorithms to obtain the MLEs of the parameters α_1 , α_2 , β , $S(t)$, and $H(t)$. The EM and SEM procedures began, in each case, with the true values of the parameters, and the iteration ended when the absolute value of the difference between the two successive iterations for each of the three parameters was less than 10^{-5} . Figures 8–12 display the ABs and MSEs of the MLEs for α_1 , α_2 , β , $S(t)$, and $H(t)$, respectively. Additionally, we use the missing information principle, the asymptotic distribution of the MLEs, and the log-transformed MLEs to generate the 95% confidence intervals for α_1 , α_2 , β , $S(t)$, and $H(t)$ (at time $t = 1.0$).

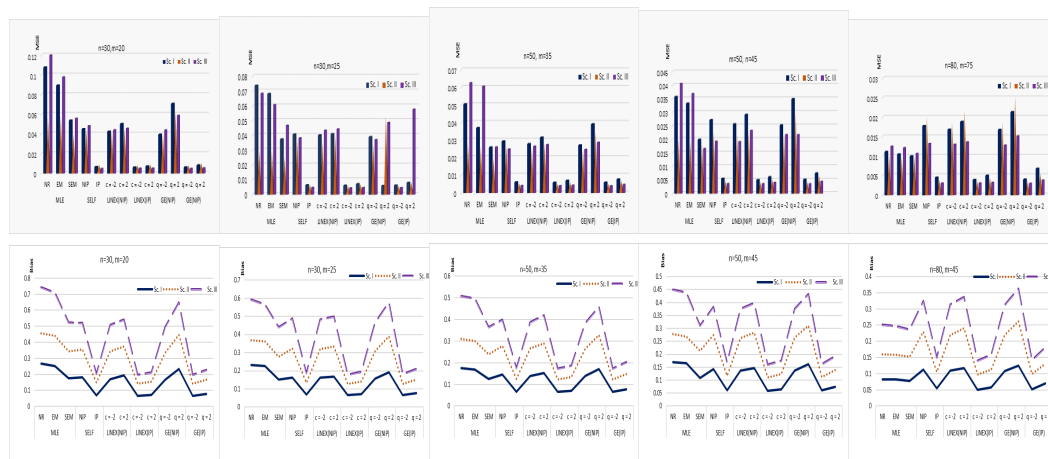


Figure 8. Relationship between the three censoring schemes, bias, and MSE values for different estimates of α_1 at different sample sizes.

In Bayesian calculations, the non-informative prior (NIP) and informative prior (IP) for unknown parameters $(\alpha_1, \alpha_2, \beta, S(t), H(t))$ are taken into account. The hyper-parameters allocated to proper Bayes are $a_1 = 4$, $a_2 = 5$, $a_3 = 15$, and $b_1 = b_2 = b_3 = 10$, for which the corresponding true means and prior means are the same. The hyper-parameters are allocated values of $a_i = b_i = 0$ (NIP) for an improper prior case. Prior I and prior II are used to denote informative and non-informative Bayes estimators, respectively. Additionally, squared error loss (SEL), LINEX loss, and general entropy loss (GEL) are the three distinct loss functions that are taken into consideration when evaluating all Bayes estimators of unknown parameters $(\alpha_1, \alpha_2, \beta, S(t), H(t))$. It was assumed that the values for the LINEX and GE loss parameters were ± 2 for both c and q . The Bayesian estimates are obtained using the M-H technique described in Section 4. In order to construct α_1 , α_2 , and β for this procedure, we used normal proposal densities from Equations (70) and (71). The initial values of α_1 , α_2 , and β in this procedure are the MLEs, or actual parameters values. Following Section 5's instruction, we created $N = 10,000$ MCMC samples and discarded

the first $M = 1000$ values as the burn-in period. Thus, the SE, LINEX, and GE loss functions were used to construct the Bayes estimates of α_1 , α_2 , β , $S(t)$, and $H(t)$, based on 9000 M-H sample data points. Using the M-H samples, we obtained the 95% symmetric credible intervals of α_1 , α_2 , β , $S(t)$, and $H(t)$. Figures 8–12 show the relationships between the progressively censoring schemes (Cs.), bias, and MSE values for different estimates of α_1 , α_2 , β , $S(t)$, and $H(t)$ at different sample sizes.



Figure 9. Relationship between the three censoring schemes, bias, and MSE values for different estimates of α_2 at different sample sizes.



Figure 10. Relationship between the three censoring schemes, bias, and MSE values for different estimates of β at different sample sizes.

Overall, the numerical results displayed in each figure demonstrate that the recommended estimates generally outperform the expectations as n increases. In every instance, it is noted that biases and mean square errors decrease with increasing sample size. For all unknown parameters, the Bayesian estimators outperform the MLE estimators, primarily because the Bayesian approach considers both the data and prior information about the unknown parameters, whereas the MLEs just consider the data. Informative prior Bayesian estimators perform better than non-informative prior estimators for the same reason. When m increases with a fixed n or n increases with a fixed m , the estimator performances of all the unknown parameters improve. Both the MSE and the bias of the Bayesian estimators and MLEs decrease. Therefore, one strategy to improve the outcome estimate is to increase the effective sample size. The best Bayes estimates for the reliability function $S(t)$ are under the GE loss function with $q = +2$, whereas the Bayesian estimates under the LINEX loss function with $c = -2$ are better than those under the SE or GE loss functions for α_1 , α_2 , β , and $H(t)$. The values of the MSE and bias of all Bayes estimates for α_1 , α_2 , β , $S(t)$, and $H(t)$,

are presented also in Figures 8–12. Finally, the ALs and associated CPs of the estimated confidence intervals for α_1 , α_2 , β , $S(t)$, and $H(t)$, are also presented in Tables 5–7.



Figure 11. Relationship between the three censoring schemes, bias, and MSE values for different estimates of $S(t)$ at different sample sizes.



Figure 12. Relationship between the three censoring schemes, bias, and MSE values for different estimates of $H(t)$ at different sample sizes.

We found that the SEM strategy produces better confidence intervals than either the NR or EM strategies in terms of average length. Still, when compared to the five recommended methods (NR, log normal, EM, SEM, and HPD), credible intervals from the HPD perform best. This suggests that their ALs are smaller than at the nominal level, and their CPs are closer to those at the nominal level. Furthermore, informative prior Bayesian credible intervals outperform non-informative prior intervals. Additionally, Bayesian credible intervals with an informative prior perform better than intervals with a non-informative prior. This is true for all parameters α_1 , α_2 , β , $S(t)$, and $H(t)$.

Finally, it can be seen that for point and interval estimations with various combinations of sample size (m, n) and censoring schemes, the Bayesian technique with an informative prior provides excellent results. Additionally, varied estimates of the parameters exhibit varied performances under varied loss functions. As a result, if it is possible to access previous knowledge about the unknown parameters, the Bayesian approach is preferable. If not, we often select the Bayesian approach with a non-informative prior when the sample size is modest.

Table 6. The average confidence lengths (ALs) and the corresponding coverage percentages (CPs) of $\alpha_1, \alpha_2, \beta, S(t)$, and $H(t)$ based on ML and Bayes estimates for censoring Scheme II.

(n, m)	SC. II.	MLE								MCMC			
		Normal		Log-Normal		EM		SEM		NIP		IP	
		AL	CP	AL	CP	AL	CP	AL	CP	AL	CP	AL	CP
(30, 20)	α_1	1.5193	0.895	1.6258	0.882	0.9935	0.928	0.7426	0.893	0.6698	0.939	0.3817	0.957
	α_2	1.7420	0.921	1.6777	0.915	1.0981	0.948	0.8834	0.909	0.6671	0.946	0.3898	0.961
	β	0.9044	0.958	1.1778	0.935	0.6592	0.950	0.5508	0.955	0.1093	0.963	0.0793	0.953
	$S(t)$	0.4201	0.975	0.4715	0.960	0.3997	0.970	0.3899	0.973	0.2534	0.971	0.1528	0.958
	$H(t)$	1.8289	0.961	1.8724	0.965	1.0964	0.974	0.9507	0.964	0.7084	0.980	0.4041	0.980
(30, 25)	α_1	1.2936	0.941	1.4567	0.967	0.9519	0.938	0.7178	0.941	0.6087	0.933	0.3673	0.948
	α_2	1.4650	0.965	1.5402	0.955	1.0203	0.925	0.8544	0.953	0.6261	0.972	0.3835	0.936
	β	0.8076	0.989	0.9287	0.971	0.6531	0.996	0.5227	0.962	0.1076	0.981	0.0784	0.968
	$S(t)$	0.4157	0.986	0.4258	0.989	0.3897	0.986	0.3733	0.971	0.2331	0.967	0.1508	0.955
	$H(t)$	1.5254	0.957	1.6910	0.976	1.0679	0.950	0.8992	0.964	0.6516	0.942	0.3949	0.970
(50, 35)	α_1	1.2055	0.948	1.3211	0.957	0.7105	0.965	0.6148	0.893	0.5241	0.963	0.3396	0.965
	α_2	1.3169	0.963	1.4668	0.969	0.8546	0.970	0.6863	0.925	0.5227	0.957	0.3530	0.932
	β	0.7246	0.959	0.8016	0.979	0.5434	0.979	0.4267	0.962	0.0656	0.951	0.0511	0.950
	$S(t)$	0.3171	0.974	0.3375	0.963	0.3165	0.983	0.3110	0.960	0.1921	0.976	0.1285	0.912
	$H(t)$	1.4375	0.968	1.5127	0.982	0.9394	0.974	0.8413	0.907	0.5463	0.980	0.3603	0.956
(50, 45)	α_1	0.9863	0.939	1.1346	0.986	0.6559	0.954	0.5835	0.964	0.4596	0.978	0.3251	0.976
	α_2	1.0898	0.944	1.2108	0.982	0.8202	0.959	0.6485	0.962	0.4750	0.969	0.3395	0.957
	β	0.5891	0.919	0.6406	0.987	0.4996	0.960	0.4037	0.974	0.0545	0.983	0.0442	0.948
	$S(t)$	0.3082	0.941	0.3248	0.965	0.3079	0.952	0.3039	0.978	0.1716	0.982	0.1254	0.969
	$H(t)$	1.0969	0.963	1.2396	0.984	0.8263	0.903	0.6334	0.980	0.4870	0.985	0.3460	0.973
(80, 60)	α_1	0.9161	0.942	1.0169	0.940	0.5611	0.949	0.4416	0.931	0.3910	0.947	0.2987	0.920
	α_2	0.9987	0.950	1.1559	0.936	0.6181	0.957	0.4962	0.935	0.4153	0.952	0.3161	0.927
	β	0.5661	0.923	0.6106	0.967	0.3879	0.964	0.3240	0.941	0.0407	0.962	0.0358	0.940
	$S(t)$	0.2579	0.946	0.2619	0.977	0.2391	0.959	0.2311	0.940	0.1478	0.968	0.1116	0.953
	$H(t)$	1.0316	0.962	1.1103	0.981	0.6150	0.961	0.4751	0.953	0.4163	0.972	0.3184	0.946
(80, 75)	α_1	0.6010	0.961	0.6203	0.948	0.5501	0.959	0.4264	0.973	0.3425	0.958	0.2758	0.949
	α_2	0.6664	0.947	0.6861	0.962	0.6082	0.971	0.4835	0.964	0.3604	0.967	0.2889	0.951
	β	0.4067	0.956	0.4177	0.967	0.3757	0.961	0.3124	0.957	0.0381	0.955	0.0343	0.953
	$S(t)$	0.2405	0.969	0.2424	0.982	0.2294	0.974	0.2206	0.970	0.1303	0.959	0.1046	0.960
	$H(t)$	0.6694	0.965	0.6841	0.980	0.6066	0.98	0.4728	0.952	0.3575	0.969	0.2904	0.957

Table 7. The average confidence lengths (ALs) and the corresponding coverage percentages (CPs) of $\alpha_1, \alpha_2, \beta, S(t)$, and $H(t)$ based on ML and Bayes estimates for censoring Scheme III.

(n, m)	SC. II.	MLE								MCMC			
		Normal		Log-Normal		EM		SEM		NIP		IP	
		AL	CP	AL	CP	AL	CP	AL	CP	AL	CP	AL	CP
(30, 15)	α_1	1.5068	0.899	1.4165	0.858	0.9673	0.885	0.7862	0.944	0.6975	0.962	0.3918	0.961
	α_2	1.6727	0.901	1.7901	0.874	1.1628	0.909	0.8341	0.958	0.8879	0.954	0.4046	0.958
	β	1.0934	0.912	0.9106	0.906	0.6737	0.924	0.5371	0.966	0.0941	0.927	0.0761	0.974
	$S(t)$	0.4288	0.933	0.4575	0.961	0.4020	0.964	0.4025	0.984	0.2780	0.932	0.1532	0.966
	$H(t)$	1.7791	0.947	1.8364	0.931	1.1477	0.950	0.7730	0.975	0.9280	0.974	0.4366	0.970

Table 7. Cont.

(n, m)	SC. II.	MLE								MCMC			
		Normal		Log-Normal		EM		SEM		NIP		IP	
		AL	CP	AL	CP	AL	CP	AL	CP	AL	CP	AL	CP
(30, 25)	α_1	1.3599	0.913	1.3795	0.925	0.9205	0.921	0.7257	0.967	0.6561	0.887	0.3765	0.964
	α_2	1.5155	0.925	1.5960	0.930	1.0907	0.955	0.7793	0.935	0.7043	0.927	0.3968	0.949
	β	0.8601	0.973	0.8983	0.941	0.6328	0.956	0.5146	0.969	0.0871	0.947	0.0660	0.964
	$S(t)$	0.4246	0.924	0.4310	0.966	0.3961	0.971	0.3949	0.978	0.2374	0.927	0.1516	0.973
	$H(t)$	1.5901	0.927	1.6530	0.965	1.0548	0.975	0.7520	0.967	0.7243	0.913	0.4109	0.982
(50, 35)	α_1	1.1810	0.927	1.2832	0.934	0.7542	0.935	0.5438	0.961	0.5691	0.943	0.3529	0.958
	α_2	1.2960	0.932	1.3718	0.939	0.8244	0.960	0.6138	0.972	0.5942	0.937	0.3789	0.952
	β	0.7910	0.965	0.7164	0.951	0.4985	0.957	0.4045	0.954	0.0518	0.985	0.0505	0.961
	$S(t)$	0.3211	0.919	0.3403	0.945	0.3054	0.968	0.3171	0.956	0.2032	0.972	0.1310	0.964
	$H(t)$	1.3949	0.941	1.4298	0.967	0.8321	0.974	0.5927	0.964	0.6323	0.968	0.3892	0.956
(50, 45)	α_1	0.9155	0.929	0.9655	0.947	0.7336	0.951	0.5136	0.960	0.4960	0.939	0.3374	0.971
	α_2	1.0021	0.932	0.8673	0.954	0.8185	0.956	0.6092	0.941	0.5172	0.923	0.3542	0.957
	β	0.5830	0.970	0.6341	0.962	0.4952	0.970	0.3910	0.965	0.0465	0.947	0.0460	0.962
	$S(t)$	0.3160	0.961	0.3228	0.969	0.3026	0.968	0.3116	0.971	0.1774	0.912	0.1278	0.967
	$H(t)$	1.0408	0.968	1.1734	0.959	0.8248	0.973	0.5837	0.959	0.5487	0.913	0.3681	0.960
(80, 60)	α_1	0.7033	0.949	0.7603	0.944	0.5849	0.949	0.4429	0.962	0.4814	0.940	0.3286	0.967
	α_2	0.7783	0.938	0.8430	0.957	0.6654	0.907	0.5021	0.956	0.4995	0.912	0.3537	0.981
	β	0.4680	0.968	0.5810	0.962	0.3909	0.968	0.3082	0.972	0.0362	0.973	0.0336	0.973
	$S(t)$	0.3098	0.977	0.3194	0.977	0.2413	0.983	0.2397	0.966	0.1621	0.867	0.1172	0.965
	$H(t)$	0.9821	0.951	1.0190	0.964	0.6566	0.987	0.4692	0.961	0.5280	0.867	0.3666	0.947
(80, 75)	α_1	0.6538	0.955	0.6768	0.978	0.5726	0.968	0.3996	0.969	0.3748	0.953	0.2913	0.962
	α_2	0.7093	0.943	0.7327	0.980	0.6289	0.965	0.4813	0.973	0.3895	0.987	0.3045	0.957
	β	0.4266	0.976	0.4394	0.981	0.3878	0.971	0.3050	0.968	0.0301	0.960	0.0296	0.967
	$S(t)$	0.2441	0.981	0.2461	0.979	0.2408	0.967	0.2378	0.971	0.1356	0.959	0.1067	0.961
	$H(t)$	0.7276	0.963	0.7456	0.994	0.6359	0.989	0.4640	0.987	0.3995	0.981	0.3126	0.948

8. Conclusions

In this study, we took into consideration progressive Type-II censoring data for the statistical inference of unknown lifespan parameters when competing failure mechanisms are present but independent. The lifespan distribution for each cause of failure is assumed to correspond to a NP distribution. The maximum likelihood and Bayesian estimate techniques are taken into consideration in order to accomplish our goal. This study was conducted on the point and approximate confidence interval estimations for the hazard rate functions, reliability, and unknown parameters. We note that in using both the SEM and EM techniques, the complexity related to the numerical computation of the MLEs may have been significantly reduced. To produce the Bayesian estimates under the squared error, LINEX, and GE loss functions, and the corresponding credible intervals, the Metropolis–Hastings method was used in the Bayesian paradigm. We analyzed a real-world dataset to illustrate the techniques presented in this article. We then carried out an extensive simulation analysis to contrast the effectiveness of different estimators. When there is no subjective information, it was found that the MLEs perform rather well. With subjective information available, the Bayesian estimators perform better than the MLEs, as predicted.

While units failing because of two competing risks are subject to progressive Type-II censoring, it should be noted that the results also apply to cases where there are multiple causes of failure and to other types of failure data cases, including complete data, Type-II censoring, and progressive first-failure censoring. The optimal design and sample plan of

progressive censoring under the competing risk model appear to be interesting topics for additional investigation and will be looked into in the future.

Author Contributions: Methodology, E.A.A. and M.S.E.; software, E.A.A. and M.S.E.; validation, T.S.A.; formal analysis, E.A.A. and M.S.E.; investigation, E.A.A. and T.S.A.; resources, T.S.A.; data curation, E.A.A.; writing—review and editing, E.A.A. and M.S.E.; visualization, T.S.A. and M.S.E. All authors have read and agreed to the published version of the manuscript.

Funding: Qassim University under project number (QU-APC-2024-9/1).

Data Availability Statement: All datasets are listed within the article.

Acknowledgments: The researchers would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support (QU-APC-2024-9/1).

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

References

1. Bourguignon, M.; Saulo, H.; Fernandez, R.N. A new Pareto-type distribution with applications in reliability and income data. *Phys. A Stat. Mech. Appl.* **2016**, *457*, 166–175. [\[CrossRef\]](#)
2. Sarabia, J.M.; Jorda, V.; Prieto, F. On a new Pareto-type distribution with applications in the study of income inequality and risk analysis. *Phys. A Stat. Mech. Appl.* **2019**, *527*, 121277. [\[CrossRef\]](#)
3. Abd Raof, A.S.; Haron, M.A.; Safari, M.A.M.; Siri, Z. Modeling the incomes of the upper-class group in Malaysia using new Pareto-type distribution. *Sains Malays.* **2022**, *51*, 3437–3448. [\[CrossRef\]](#)
4. Nik, A.S.; Asgharzadeh, A.; Nadarajah, S. Comparisons of methods of estimation for a new Pareto-type distribution. *Statistica* **2019**, *79*, 291–319.
5. Nik, A.S.; Asgharzadeh, A.; Raqab, M.Z. Estimation and prediction for a new Pareto-type distribution under progressive type-II censoring. *Math. Comput. Simul.* **2021**, *190*, 508–530.
6. Nik, A.S.; Asgharzadeh, A.; Baklizi, A. Inference based on new Pareto-type records with applications to precipitation and COVID-19 data. *Stat. Optim. Inf. Comput.* **2023**, *11*, 243–257. [\[CrossRef\]](#)
7. Li, F.; Wei, S.; Zhao, M. Bayesian estimation of a new Pareto-type distribution based on mixed Gibbs sampling algorithm. *Mathematics* **2023**, *12*, 18. [\[CrossRef\]](#)
8. Safari, M.A.M.; Masseran, N. Robust estimation techniques for the tail index of the new Pareto-type distribution. *Empir. Econ.* **2024**, *66*, 1161–1189. [\[CrossRef\]](#)
9. Balakrishnan, N.; Aggarwala, R. *Progressive Censoring: Theory, Methods, and Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2000.
10. Mao, S.; Shi, Y.; Wang, L. Exact inference for two exponential populations with competing risks data. *J. Syst. Eng. Electron.* **2014**, *25*, 711–720.
11. Azizi, F.; Haghighi, F.; Tabibi, G.N. Statistical inference for competing risks model under progressive interval censored Weibull data. *Commun. Stat. Simul. Comput.* **2020**, *49*, 1931–1944. [\[CrossRef\]](#)
12. Dutta, S.; Kayal, S. Bayesian and non-Bayesian inference of Weibull lifetime model based on partially observed competing risks data under unified hybrid censoring scheme. *Qual. Reliab. Eng. Int.* **2022**, *38*, 3867–3891. [\[CrossRef\]](#)
13. Dutta, S.; Ng, H.K.T.; Kayal, S. Inference for a general family of inverted exponentiated distributions under unified hybrid censoring with partially observed competing risks data. *J. Comput. Appl. Math.* **2023**, *422*, 114934. [\[CrossRef\]](#)
14. Greene, W.H. *Econometric Analysis*, 4th ed.; Prentice-Hall: New York, NY, USA, 2000.
15. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B Stat. Methodol.* **1977**, *39*, 1–22. [\[CrossRef\]](#)
16. 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. [\[CrossRef\]](#)
17. Zhang, M.; Zhisheng, Y.; Min, X. A stochastic EM algorithm for progressively censored data analysis. *Qual. Reliab. Eng. Int.* **2013**, *30*, 711–722. [\[CrossRef\]](#)
18. Celeux, G.; Diebolt, J. The SEM algorithm: A probabilistic teacher algorithm derived from the EM algorithm for the mixture problem. *Comput. Stat. Quar.* **1985**, *2*, 73–82.
19. Cariou, C.; Chehdi, K. Unsupervised texture segmentation/classification using 2-d autoregressive modeling and the stochastic expectation-maximization algorithm. *Pattern Recognit. Lett.* **2008**, *29*, 905–917. [\[CrossRef\]](#)
20. Louis, T.A. Finding the observed information matrix using the EM algorithm. *J. R. Stat. Soc. B Stat. Methodol.* **1982**, *44*, 226–233. [\[CrossRef\]](#)
21. Metropolis, N.; Rosenbluth, A.; Rosenbluth, M.; Teller, A.; Teller, E. Equations of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092. [\[CrossRef\]](#)
22. Hastings, W.K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **1970**, *57*, 97–109. [\[CrossRef\]](#)
23. Lawless, J.F. *Statistical Models and Methods for Lifetime Data*, 2nd ed.; John Wiley and Sons: New York, NY, USA, 2003.

24. Park, C.; Kulasekera, K.B. Parametric inference of incomplete data with competing risks among several groups. *IEEE Trans. Reliab.* **2004**, *53*, 11–21. [[CrossRef](#)]
25. Balakrishnan, N.; Kateri, M. On the maximum likelihood estimation of parameters of Weibull distribution based on complete and censored data. *Stat. Probab. Lett.* **2008**, *78*, 2971–2975. [[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.