

## Article

# On Estimation of Shannon's Entropy of Maxwell Distribution Based on Progressively First-Failure Censored Data

Kapil Kumar <sup>1,†</sup> , Indrajeet Kumar <sup>2,†</sup> and Hon Keung Tony Ng <sup>3,\*,†</sup> <sup>1</sup> Department of Statistics, Central University of Haryana, Mahendergarh 123031, India; kapilstats@cuh.ac.in<sup>2</sup> Department of Statistics, Central University of South Bihar, Gaya 824236, India; indrajeetkumar@cusb.ac.in<sup>3</sup> Department of Mathematical Sciences, Bentley University, Waltham, MA 02452, USA

\* Correspondence: tng@bentley.edu; Tel.: +1-781-216-7102

† These authors contributed equally to this work.

**Abstract:** Shannon's entropy is a fundamental concept in information theory that quantifies the uncertainty or information in a random variable or data set. This article addresses the estimation of Shannon's entropy for the Maxwell lifetime model based on progressively first-failure-censored data from both classical and Bayesian points of view. In the classical perspective, the entropy is estimated using maximum likelihood estimation and bootstrap methods. For Bayesian estimation, two approximation techniques, including the Tierney-Kadane (T-K) approximation and the Markov Chain Monte Carlo (MCMC) method, are used to compute the Bayes estimate of Shannon's entropy under the linear exponential (LINEX) loss function. We also obtained the highest posterior density (HPD) credible interval of Shannon's entropy using the MCMC technique. A Monte Carlo simulation study is performed to investigate the performance of the estimation procedures and methodologies studied in this manuscript. A numerical example is used to illustrate the methodologies. This paper aims to provide practical values in applied statistics, especially in the areas of reliability and lifetime data analysis.

**Keywords:** Maxwell distribution; Shannon's entropy; bootstrap confidence intervals; bayesian estimation; T-K approximation



**Citation:** Kumar, K.; Kumar, I.; Ng, H.K.T. On Estimation of Shannon's Entropy of Maxwell Distribution Based on Progressively First-Failure Censored Data. *Stats* **2024**, *7*, 138–159. <https://doi.org/10.3390/stats7010009>

Academic Editor: Wei Zhu

Received: 29 December 2023

Revised: 5 February 2024

Accepted: 6 February 2024

Published: 8 February 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

Information theory offers a feasible way for quantifying uncertainty and reciprocal information of random variables in which entropy plays a significant role. Shannon [1] established the concept of statistical entropy, namely Shannon's entropy, which measures the average level of missing information in a random source, as a fundamental idea in information theory. The relationship between Shannon's entropy and the information available in classical and quantum systems is well established. In addition, statistical mechanics interpretations can be made through Shannon's entropy. Shannon's entropy is typically used in statistical mechanics to represent the entropy for classical systems with configurations taken from canonical ensembles. Moreover, Shannon's entropy is useful in many disciplines, such as computer science, molecular biology, hydrology, and meteorology, to solve different scientific problems involving uncertainties in random quantities. For example, molecular biologists use the principle of Shannon's entropy to study trends in gene sequences. Recently, Saraiva [2] exemplified Shannon's entropy in biological diversity and student migration studies and provided an intuitive introduction to Shannon's entropy. For more details, one may refer to the excellent monograph by Cover [3] on the theory and implications of entropy with applications in various disciplines. Shannon's entropy is one of the most widely used entropies in statistics and information theory.

Suppose a random variable  $X$  has probability density function (pdf)  $f(\cdot)$ , the Shannon's entropy of the random variable  $X$ , is given by

$$H(f) = E[-\ln f(X)] = - \int_{-\infty}^{\infty} f(x) \ln(f(x)) dx. \quad (1)$$

Please note that the entropy in Equation (1) is a continuous entropy, and the measurement is relative to the coordinate system ([1] Section 20). In other words, the continuous entropy is not invariant under a variable change in general. However, the entropy based on the variable change can be expressed as the original entropy less the expected logarithm of the Jacobian of the variable change. Recently, researchers studied parametric statistical inference for measuring the entropy under different lifetime models based on complete or censored data. For example, entropy for several shifted exponential populations was studied by Kayal and Kumar [4]. Another author Cho et al. [5] studied the entropy estimation for Rayleigh distribution under doubly generalized Type-II hybrid censored data. Another paper Du et al. [6] developed statistical inference for the information entropy of the log-logistic distribution based on progressively Type-I interval-censored data, and Liu and Gui [7] investigated the entropy estimation for Lomax distribution based on generalized progressively hybrid censoring. Another paper Yu et al. [8] developed statistical inference on Shannon's entropy for inverse Weibull distribution based on progressively first-failure censored data.

The Maxwell–Boltzmann distribution, popularly known as the Maxwell distribution (MWD), was initially developed by James Clerk Maxwell and Ludwig Boltzmann in the late 1800s as a distribution of velocities in a gas at a specific temperature. In this paper, we consider the MWD due to its simplicity and nice physical interpretation. For further information about the reliability characteristics of the MWD, see Bekker and Roux [9]. The MWD is frequently employed in physics and chemistry for several reasons, including the fact that the MWD can be used to describe several important gas properties, such as pressure and diffusion. The MWD has become a well-known lifespan model in recent years. Many researchers studied this statistical distribution in depth for modeling lifespan data. As Bekker and Roux [9] pointed out, the MWD is useful in life-testing and reliability studies because of its desirable properties, especially for situations where the assumption of a constant failure rate is not realistic. Recently, in the reliability report for Gallium Nitride (GaN) power devices by Pozo et al. [10], they found that the Maxwell–Boltzmann distribution fits the hot carrier distribution in the high energy regime tails well.

The classical estimation of the model parameters of MWD has been studied by [11,12] for complete and censored samples, respectively. Following this, Krishna and Malik [13] studied the MWD under progressive censoring, Krishna et al. [14] discussed the MWD under randomly censored data, Tomer and Panwar [15] developed an estimation procedure for the MWD based on Type-I progressive hybrid censored data, Panwar and Tomer [16] discussed robust Bayesian analysis for the MWD, and Kumari et al. [17] discussed classical and Bayesian estimation for the MWD based on adaptive progressive Type-II censored data.

The pdf and cumulative distribution function (cdf), respectively, for a random variable  $X$  that follows the Maxwell distribution (MWD) with parameter  $\lambda$  (denoted as  $MWD(\lambda)$ ) are given by

$$f(x; \lambda) = \frac{4}{\sqrt{\pi}} \frac{1}{\lambda^{3/2}} x^2 e^{-\frac{x^2}{\lambda}}; \quad 0 < x < \infty, \lambda > 0, \quad (2)$$

$$F(x; \lambda) = \Gamma\left(\frac{x^2}{\lambda}, \frac{3}{2}\right); \quad 0 < x < \infty, \lambda > 0. \quad (3)$$

where  $\Gamma(x, b) = \frac{1}{\Gamma(b)} \int_0^x e^{-t} t^{b-1} dt$  is the incomplete gamma ratio. The MWD has an increasing failure rate function [13]. The MWD is a special case of many generalized distributions, such as the generalized gamma distribution [18], generalized Rayleigh distribution ([19] p.452), and generalized Weibull distribution [20].

Assume that the lifetime of the items of interest follow the MWD with pdf in Equation (2), from Equation (1), the Shannon's entropy is given by

$$H(f) = \frac{1}{2} \ln \lambda + \gamma + \frac{1}{2} \ln \pi - \frac{1}{2} \stackrel{def.}{=} H(\lambda) \quad \text{say,} \quad (4)$$

where  $\gamma = 0.5772156649 \dots$  is the Euler–Mascheconi constant.

Censoring often occurs in reliability engineering or life testing procedures when the actual lifetimes of the items of interest are not observed; for instance, when subjects or experimental units are taken out of the experiments intentionally or accidentally. Pre-planned censoring saves time and costs for experiments, and it has been applied in various fields, including but not limited to engineering, survival analysis, clinical trials, medical research, etc. Several censoring schemes, such as Type-I and Type-II censoring schemes (also known as time and item censoring schemes, respectively), are often utilized in life-testing experiments. When products or items have a long lifespan, a life-testing experiment may require a long time to complete, even with Type-I or Type-II censoring. For these situations, Balasooriya [21] established the first-failure censoring plan that provides a time and cost-saving strategy for life-testing experiments. In this censoring scheme, an experimenter examines  $k \times n$  units by grouping them into  $n$  groups, each having  $k$  items, and then conducts all the tests jointly until the first failure is seen in each group. Although the first-failure censoring plan saves time and costs, it does not allow for the periodic removal of surviving units or subjects during the life-testing experiment. Therefore, the progressive censoring scheme proposed by Cohen [22] that allows for removing units throughout the life-testing experiment can be considered. For this reason, first-failure and progressive censoring were combined to create a more flexible life-testing strategy known as the progressive first-failure censoring (PFFC) scheme [23]. Due to its compatible features with other censoring plans, the PFFC scheme has gained much attention in the literature. For instance, the Lindley distribution based on PFFC data was studied by Dube et al. [24], the exponentiated exponential distribution based on PFFC data was studied by Mohammed et al. [25], the estimation of stress-strength reliability for generalized inverted exponential distribution based on PFFC data was studied by Krishna et al. [26]. One author Kayal et al. [27] studied the PFFC scheme and developed statistical inference on the Chen distribution. Following on this, Saini et al. [28] studied the estimation of stress-strength reliability for generalized Maxwell distribution based on PFFC data, and Kumar et al. [29] discussed the reliability estimation in inverse Pareto distribution based on PFFC data, etc.

The PFFC scheme can be described as follows: suppose there are  $n$  independent groups, each with  $k$  items, these are placed on a life test, and we have a prespecified number of observed failures  $m$  ( $\leq n$ ) and a prefixed progressive censoring plan  $\mathcal{G} = (G_1, G_2, \dots, G_m)$ , where  $G_i$  ( $i = 1, 2, \dots, m$ ) are the prefixed number of groups without item failure to be removed at  $i$ -th failure such that  $n = m + G_1 + G_2 + \dots + G_m$ . In other words, the test will be terminated when the  $m$ -th failure is observed. When the first failure occurs at time  $X_{1:m:n:k}$ ,  $G_1$  groups without item failure and the group where the first failure is observed are removed from the experiment. When the second failure occurs at time  $X_{2:m:n:k}$ ,  $G_2$  groups without item failure and the group containing the second failure are removed from the experiment, and so on. Finally, when the  $m$ -th failure occurs at time  $X_{m:m:n:k}$ , the remaining  $G_m$  groups without item failure and the group containing  $m$ -th failure are removed from the experiment. Consequently, the observed failure times,  $X_{1:m:n:k}^{(\mathcal{G})} < X_{2:m:n:k}^{(\mathcal{G})} < \dots < X_{m:m:n:k}^{(\mathcal{G})}$  are called progressively first-failure censored order statistics with the progressive censoring plan  $\mathcal{G} = (G_1, G_2, \dots, G_m)$ . Figure 1 represents the schematic diagram of the PFFC scheme. Note that the PFFC scheme bears the following special cases: (i) it reduces to a complete

sample case when  $k = 1, n = m$  and  $G_i = 0; i = 1, 2, \dots, m$ ; (ii) it reduces to a conventional Type-II censoring plan, if  $k = 1$  and  $G_i = 0; i = 1, 2, \dots, m - 1$ , and  $G_m = n - m$ ; (iii) it becomes a progressive Type-II censoring plan when  $k = 1$ , and (iv) it reduces to a first-failure censoring plan, when  $m = n$  and  $G_i = 0; i = 1, 2, \dots, m$ . Moreover, the applicability or practicality of PFFC may be a particular setup of a parallel-series system in which  $n$  homogeneous system like a batch of electric bulbs are connected parallel, each batch has  $k$  bulbs or electronics components connected in series. For the testing procedures or collecting lifetimes, the same mechanism presented in Figure 1 can be used.

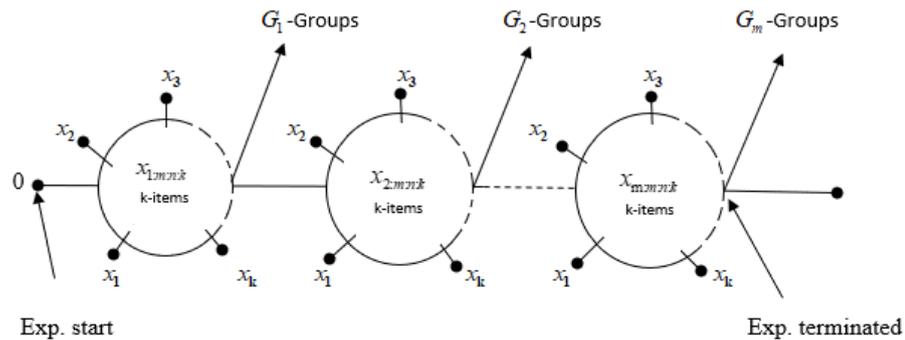


Figure 1. Schematic diagram of PFFC scheme.

Let  $\underline{x} = (x_{1:m:n:k}^{(G)}, x_{2:m:n:k}^{(G)}, \dots, x_{m:m:n:k}^{(G)})$  be a PFFC sample drawn from a continuous population with cdf  $F(x)$  and pdf  $f(x)$ . For notation convenience, we suppress the notations  $G$  and  $(m : n : k)$  in the observed data and denote the observed data as  $\underline{x} = (x_1, x_2, \dots, x_m)$ . The likelihood function can be expressed as [23]

$$L(\underline{x}) = Ak^m \prod_{i=1}^m f(x_i) [1 - F(x_i)]^{k(G_i+1)-1}, \quad 0 < x_1 < x_2 < \dots < x_m < \infty, \quad (5)$$

where  $A = n(n - G_1 - 1)(n - G_1 - G_2 - 2) \dots (n - G_1 - G_2 - \dots - G_{m-1} - m + 1)$ .

To establish any information based on the available data, we can consider a statistical probability model involving some parameter(s), and first, we need to estimate the model parameter(s). The two popular estimation methods in the literature are classical and Bayesian. The Bayesian estimation can be used when the prior information is available. In this study, we consider both estimation methods for feasibility purposes. The main objective of this study is to estimate the associate parameter and Shannon’s entropy for the MWD based on progressively first-failure censored data. Through the study in this manuscript, we aim to provide practical values in applied statistics, especially in the areas of reliability and lifetime data analysis.

The rest of this paper is organized as follows. Section 2 develops the frequentist estimation techniques, including the maximum likelihood estimation and asymptotic and bootstrap confidence intervals. Section 3 is devoted to Bayesian estimation techniques using the Tierney-Kadane (T-K) approximation and Markov Chain Monte Carlo (MCMC) methods. In Section 4, a Monte Carlo simulation study is used to evaluate the performance of the estimation procedures developed in this manuscript. A numerical example is provided in Section 5 to illustrate the methodologies developed in this manuscript. Finally, in Section 6, some concluding remarks are presented.

## 2. Frequentist Estimation Approach

In this section, we develop the expectation-maximization (EM) algorithm to obtain the maximum likelihood estimates (MLEs) of the related parameters  $\lambda$  and Shannon’s entropy  $H(\lambda)$ . The EM algorithm is one of the popularly used iterative methods for the censored data; for more details, one may refer to Dempster et al. [30], McLachlan and Krishnan [31],

and Casella and Berger [32]. Additionally, we consider the asymptotic and bootstrap confidence intervals for  $\lambda$  and  $H(\lambda)$ .

2.1. Maximum Likelihood Estimation

Based on the observed sample under the PFFC scheme with censoring plan  $G$ ,  $\underline{x} = (x_1, x_2, \dots, x_m)$ , and assuming that the lifetimes of the items follow the MWD( $\lambda$ ), the likelihood function in Equation (5) can be obtained as

$$L(\underline{x}; \lambda) = Ak^m \left( \frac{4}{\sqrt{\pi}} \right)^m \lambda^{-\frac{3m}{2}} \exp \left\{ -\frac{1}{\lambda} \sum_{j=1}^m x_j^2 \right\} \prod_{j=1}^m x_j^2 \left[ 1 - \Gamma \left( \frac{x_j^2}{\lambda}, \frac{3}{2} \right) \right]^{k(G_j+1)-1}. \tag{6}$$

The MLE of the parameter  $\lambda$  can be obtained by maximizing the likelihood function in Equation (6). Instead of directly maximizing the likelihood function, we take advantage of the closed-form MLE of  $\lambda$  based on a complete sample by considering the EM algorithm to obtain the MLE of  $\lambda$ .

Suppose the observed and censored data are  $\underline{X} = \underline{x} = (x_1, x_2, \dots, x_m)$ , and  $\underline{Z} = (Z_{11}, \dots, Z_{1[k(G_1+1)-1]}, \dots, Z_{m1}, \dots, Z_{m[k(G_m+1)-1]})$ , respectively, the combined forms of complete data set is given by  $\underline{Y} = (\underline{X}, \underline{Z})$ . If a complete sample  $\underline{Y} = (y_1, y_2, \dots, y_{nk})$  from MWD is available, the MLE of  $\lambda$ , denoted as  $\hat{\lambda}_C$  can be expressed in a closed form as

$$\hat{\lambda}_C = \frac{2}{3nk} \sum_{i=1}^{nk} y_i^2.$$

Based on the complete sample  $\underline{Y} = (\underline{X}, \underline{Z})$ , the log-likelihood function can be expressed as

$$\ln L_C(\underline{Y}; \lambda) = constant - \frac{3nk}{2} \ln \lambda - \frac{1}{\lambda} \sum_{i=1}^m x_i^2 - \frac{1}{\lambda} \sum_{i=1}^m \sum_{j=1}^{k(G_i+1)-1} Z_{ij}^2. \tag{7}$$

The EM algorithm treats the censored data as missing observations and replaces those missing values with their corresponding expected values. In the expectation step (E-step) of the EM algorithm, we replace the functions of the unobserved (censored) data  $\underline{Z}$  with the corresponding expected values given the observed data  $\underline{x}$ . Hence, the pseudo-complete log-likelihood function can be obtained as

$$\ln L_C(\underline{Y}; \lambda) = constant - \frac{3nk}{2} \ln \lambda - \frac{1}{\lambda} \sum_{i=1}^m x_i^2 - \frac{1}{\lambda} \sum_{i=1}^m \sum_{j=1}^{k(G_i+1)-1} E \left[ Z_{ij}^2 | Z_{ij} > x_i \right]. \tag{8}$$

For given  $X_i = x_i$ , the conditional distribution of  $Z_{ij}$ , follows a truncated MWD with left truncation at  $x_i$ ,

$$f(z_{ij} | x_i, \lambda) = \frac{f(z_{ij}, \lambda)}{1 - F(x_i, \lambda)}; \quad z_{ij} > x_i, \quad i = 1, 2, \dots, m; \quad j = 1, 2, \dots, [k(G_i + 1) - 1]. \tag{9}$$

Then, the conditional expectations involved in Equation (9) can be computed as

$$A(c, \lambda) = E \left[ Z_{ij}^2 | Z_{ij} > c \right] = \frac{3\lambda}{2[1 - F(x_i, \lambda)]} \left[ 1 - \Gamma \left( \frac{c^2}{\lambda}, \frac{3}{2} \right) \right]. \tag{10}$$

In the maximization step (M-step) of the EM algorithm, suppose the estimate of  $\lambda$  in the  $r$ -th stage of the EM algorithm is  $\hat{\lambda}^{(r)}$ , then the updated estimate in the  $(r + 1)$ -th stage can be obtained as

$$\hat{\lambda}^{(r+1)} = \frac{2}{3nk} \left[ \sum_{i=1}^m x_i^2 + \sum_{i=1}^m \sum_{j=1}^{k(G_i+1)-1} A(x_i, \lambda^{(r)}) \right]. \tag{11}$$

With an initial estimate  $\lambda^{(0)}$  of parameter  $\lambda$ , the E-step and M-step iterate until convergence occurs. Here, we consider convergence as  $|\lambda^{(r+1)} - \lambda^{(r)}| < \epsilon$ , where  $\epsilon$  is a small value. Once the MLE  $\hat{\lambda}$  of parameter  $\lambda$  is computed, using the invariance property of MLE, the MLE of Shannon’s entropy  $H(\lambda)$  is obtained as

$$\hat{H}(\hat{\lambda}) = \frac{1}{2} \ln \hat{\lambda} + \gamma + \frac{1}{2} \ln \pi - \frac{1}{2}.$$

### 2.2. Asymptotic Confidence Interval

In this subsection, we construct confidence intervals for the parameter  $\lambda$  and Shannon’s entropy  $H(\lambda)$  based on the asymptotic theory of the MLE. The standard errors of the MLE can be obtained by direct computation or numerical differentiation by evaluating the negative of the second derivative of the observed log-likelihood function Tanner [33] [Section 4.4]. However, the observed information can be difficult to compute because of the incomplete data. Therefore, the missing information principle computes the observed Fisher information by using the difference between the information in the pseudo-complete sample, and the information in the missing observations can be used to obtain the asymptotic variance of the MLE  $\lambda$ . The missing information principle presented by Louis [34] (see also, [33]) can be described as follows:

$$\text{Observe Information} = \text{Complete Information} - \text{Missing Information} \tag{12}$$

Let  $I_X(\lambda)$  be the Fisher information based on the observed data,  $I_Y(\lambda)$  be the Fisher information based on a complete sample, and  $I_{Y|X}(\lambda)$  be the missing information, then Equation (12) can be expressed as

$$I_X(\lambda) = I_Y(\lambda) - I_{Y|X}(\lambda). \tag{13}$$

For MWD, the Fisher information based on complete data is given by

$$I_Y(\lambda) = -E \left[ \frac{\partial^2 L_c(Y; \lambda)}{\partial \lambda^2} \right] = \frac{3nk}{2\lambda^2} - \frac{1}{\lambda^2} \sum_{i=1}^m x_i^2. \tag{14}$$

At the time of the  $i$ -th failure at time  $x_i$ , the Fisher information in a censored observation can be obtained as

$$I_{Y|X}^{(i)}(\lambda) = -E_{Z_{ij}|x_i} \left[ \frac{\partial^2 \ln f(Z_{ij}|x_i, \lambda)}{\partial \lambda^2} \right] = \frac{3}{2} \frac{1}{\lambda^2} + \psi'(\lambda)$$

where  $\psi(\lambda) = -\frac{3}{2\lambda \left[ 1 - \Gamma\left(\frac{x_i^2}{\lambda}, \frac{3}{2}\right) \right]} \left\{ \frac{3}{2} \Gamma\left(\frac{x_i^2}{\lambda}, \frac{5}{2}\right) - \Gamma\left(\frac{x_i^2}{\lambda}, \frac{3}{2}\right) \right\},$

$$\psi'(\lambda) = \frac{\partial \psi(\lambda)}{\partial \lambda} = -\frac{3}{2\lambda^2} \left\{ \frac{[1 - \Gamma(x_i^2/\lambda, 3/2)] Q_1 + Q_2}{1 - \Gamma(x_i/\lambda, 3/2)} \right\},$$

$$Q_1 = \frac{5}{2} \left[ \Gamma(x_i^2/\lambda, 7/2) - 2\Gamma(x_i^2/\lambda, 5/2) + \Gamma(x_i^2/\lambda, 3/2) \right],$$

$$Q_2 = \frac{3}{2} \left[ \Gamma(x_i^2/\lambda, 5/2) - \Gamma(x_i^2/\lambda, 3/2) \right].$$

Consequently, the missing information can be obtained as

$$I_{Y|X}(\lambda) = \sum_{i=1}^m [k(G_i + 1) - 1] I_{Y|X}^{(i)}(\lambda). \tag{15}$$

From Equation (12), the Fisher information of the observed data  $I_X(\lambda)$  can be obtained by using Equations (14) and (15). Then, under suitable regularity conditions (e.g., the true param-

eter value of  $\lambda$  must be interior to the parameter space, the derivatives of the log-likelihood function with respect to  $\lambda$  exist up to third order and bounded), the asymptotic variance of the MLE of  $\lambda$  can be obtained as  $\widehat{Var}(\hat{\lambda}) = I_X^{-1}(\lambda)|_{\lambda=\hat{\lambda}} = \left[ I_Y(\lambda) - I_{Y|X}(\lambda) \right]^{-1}|_{\lambda=\hat{\lambda}}$ . Therefore, an asymptotic  $100(1 - \alpha)\%$  confidence interval of  $\lambda$  is  $\hat{\lambda} \pm z_{\alpha/2} \sqrt{\widehat{Var}(\hat{\lambda})}$ , where  $z_{\alpha/2}$  is the upper  $(\alpha/2)$ -th quantile of the standard normal distribution  $N(0, 1)$ .

To construct an asymptotic confidence interval of Shannon's entropy  $H(\lambda)$ , we first use the delta method to approximate the variance of the MLE of  $H(\lambda)$  [32]. Specifically, suppose  $\hat{\lambda}$  is the MLE of  $\lambda$ , the asymptotic variance of  $\hat{H}(\lambda)$  using the delta method (see, for example, Krishnamoorthy and Lin [35]) is given by

$$Var(H) = [b'_C I_X^{-1}(\lambda) b_C],$$

where  $b_C = \frac{\partial H(\lambda)}{\partial \lambda} = \frac{1}{2\lambda}$ . Hence, the approximate variance of  $\hat{H}(\lambda)$

$$\widehat{Var}(\hat{H}) \simeq [b'_C I_X^{-1}(\lambda) b_C]_{\lambda=\hat{\lambda}}.$$

Based on the asymptotic theory of MLE,  $\frac{\hat{H}-H}{\sqrt{\widehat{Var}(\hat{H})}} \sim N(0, 1)$  asymptotically. Therefore, an asymptotic  $100(1 - \alpha)\%$  confidence interval of  $H$  is given by  $\hat{H} \pm z_{\alpha/2} \sqrt{\widehat{Var}(\hat{H})}$ .

### 2.3. Bootstrap Confidence Intervals

The bootstrap method is a well-known resampling technique used in statistics to estimate the sampling distribution of a statistic by repeatedly resampling with replacement from the observed data. The main objective of this method is to provide an empirical approximation of the sampling distribution of a statistic, which can be useful for making inferences and constructing confidence intervals. In the literature, the concept of the bootstrap technique was first introduced by Efron [36], and since then, several applications have been discussed. For more details, one may refer to Efron and Tibshirani [37]. Here, we construct confidence intervals for the parameter  $\lambda$  and Shannon's entropy using two parametric bootstrap methods—the percentile bootstrap (boot-p) method [38] and the bootstrap- $t$  (boot-t) method [39]. Following Efron and Tibshirani [37], the procedures to obtain the two parametric bootstrap confidence intervals are described as follows.

#### 2.3.1. Percentile Bootstrap (Boot-P) Confidence Intervals

**Step A1:** Based on the observed PFFC sample  $\underline{X} = (x_1, x_2, \dots, x_m)$  with effective sample size  $m$  from  $n$  groups each having  $k$  items and progressive censoring plan  $\underline{G}$ , compute the MLEs  $\hat{\lambda}$  and  $\hat{H}$  of parameter  $\lambda$  and entropy  $H(\lambda)$ , respectively.

**Step A2:** Generate an independent bootstrap PFFC sample  $\underline{X}^{(b)} = (x_1^{(b)}, x_2^{(b)}, \dots, x_m^{(b)})$  from  $MWD(\hat{\lambda})$  with effective sample size  $m$  from  $n$  groups each having  $k$  items and progressive censoring plan  $\underline{G}$ .

**Step A3:** Based on the bootstrap sample  $\underline{X}^{(b)}$ , compute the MLEs  $\hat{\lambda}^{(b)}$  and  $\hat{H}^{(b)}$  of parameter  $\lambda$  and entropy  $H(\lambda)$ , respectively.

**Step A4:** Repeat **Step A2** and **Step A3**  $B$  times to obtain a set of bootstrap estimates as  $(\hat{\lambda}^{(b)}, \hat{H}^{(b)}); b = 1, 2, \dots, B$ .

**Step A5:** Order the bootstrap estimates  $(\hat{\lambda}^{(1)}, \hat{\lambda}^{(2)}, \dots, \hat{\lambda}^{(B)})$  and  $(\hat{H}^{(1)}, \hat{H}^{(2)}, \dots, \hat{H}^{(B)})$  in ascending order as  $(\hat{\lambda}^{[1]} < \hat{\lambda}^{[2]} < \dots < \hat{\lambda}^{[B]})$  and  $(\hat{H}^{[1]} < \hat{H}^{[2]} < \dots < \hat{H}^{[B]})$ , respectively. The  $100(1 - \alpha)\%$  percentile bootstrap confidence intervals for  $\lambda$  and  $H$ , respectively, are given by

$$\left( \hat{\lambda}^{[(\alpha/2)B]}, \hat{\lambda}^{[(1-\alpha/2)B]} \right) \text{ and } \left( \hat{H}^{[(\alpha/2)B]}, \hat{H}^{[(1-\alpha/2)B]} \right),$$

where  $[a]$  is the integral part of  $a$ .

### 2.3.2. Bootstrap-*t* (Boot-*t*) Confidence Intervals

**Step B1:** Based on the observed PFFC sample  $\underline{X} = (x_1, x_2, \dots, x_m)$  with effective sample size  $m$  from  $n$  groups each having  $k$  items and progressive censoring plan  $\underline{G}$ , compute the MLEs  $\hat{\lambda}$  and  $\hat{H}$  of parameter  $\lambda$  and entropy  $H(\lambda)$ , respectively.

**Step B2:** Generate an independent bootstrap PFFC sample  $\underline{X}^{(b)} = (x_1^{(b)}, x_2^{(b)}, \dots, x_m^{(b)})$  from  $MWD(\hat{\lambda})$  with effective sample size  $m$  from  $n$  groups each having  $k$  items and progressive censoring plan  $\underline{G}$ .

**Step B3:** Compute the bootstrap-*t* statistics for  $\lambda$  as

$$\tau_{\lambda}^{(b)} = \frac{\hat{\lambda}^{(b)} - \hat{\lambda}}{\sqrt{\widehat{Var}(\hat{\lambda}^{(b)})}},$$

and the bootstrap-*t* statistic for  $H$  as

$$\tau_H^{(b)} = \frac{\hat{H}^{(b)} - \hat{H}}{\sqrt{\widehat{Var}(\hat{H}^{(b)})}}.$$

**Step B4:** Repeat **Step B2** and **Step B3**  $B$  times to obtain a set of bootstrap statistics as  $(\tau_{\lambda}^{(b)}, \tau_H^{(b)}); b = 1, 2, \dots, B$ .

**Step B5:** Order the bootstrap statistics  $(\tau_{\lambda}^{(1)}, \tau_{\lambda}^{(2)}, \dots, \tau_{\lambda}^{(B)})$  and  $(\tau_H^{(1)}, \tau_H^{(2)}, \dots, \tau_H^{(B)})$  in ascending order as  $(\tau_{\lambda}^{[1]} < \tau_{\lambda}^{[2]} < \dots < \tau_{\lambda}^{[B]})$  and  $(\tau_H^{[1]} < \tau_H^{[2]} < \dots < \tau_H^{[B]})$ , respectively.

**Step B6:** The  $100(1 - \alpha)\%$  bootstrap-*t* confidence intervals for  $\lambda$  and  $H$ , respectively, are given by

$$\left( \hat{\lambda} - \tau_{\lambda}^{[(1-\alpha/2)B]} \sqrt{\widehat{Var}(\hat{\lambda})}, \hat{\lambda} + \tau_{\lambda}^{[(1-\alpha/2)B]} \sqrt{\widehat{Var}(\hat{\lambda})} \right)$$

and

$$\left( \hat{H} - \tau_H^{[(1-\alpha/2)B]} \sqrt{\widehat{Var}(\hat{H})}, \hat{H} + \tau_H^{[(1-\alpha/2)B]} \sqrt{\widehat{Var}(\hat{H})} \right).$$

### 3. Bayesian Estimation Approach

In this section, we derive Bayes estimators under the linear exponential (LINEX) loss function and construct the highest posterior density (HPD) credible intervals for the parameters  $\lambda$  and Shannon’s entropy  $H(\lambda)$ . For details of Bayesian statistical inference and data analysis methods, one may refer to the books by Box and Tiao [40] and Gelman et al. [41]. The Bayesian approach to reliability analysis involves previous knowledge of lifespan parameters, technical knowledge of failure mechanisms, as well as experimental data to be incorporated into the inferential procedure. As pointed out by Tian et al. [42], employing the Bayesian approach in reliability analysis has the advantages of making statistical inferences using information from prior experience with the failure mechanism or physics-of-failure and avoiding making inferences based on plausibly inaccurate large-sample theory in frequentist approach. For more details about Bayesian inference methods and specifying prior distribution in reliability applications, one may refer to Tian et al. [42]. As a result, Bayesian techniques are frequently applied to small sample data, which is highly advantageous in the case of pricey life testing tests. For Bayesian estimation, the inverted gamma distribution is commonly used as the natural conjugate prior density for the parameter  $\lambda$  of MWD (see, for example, Bekker and Roux [9], Chaudhary et al. [43]). Following Bekker

and Roux [9] and Chaudhary et al. [43], we consider the prior distribution of the unknown parameter  $\lambda$  is the inverted gamma distribution with pdf

$$g(\lambda) \propto \frac{1}{\lambda^{a+1}} \exp(-b/\lambda); \quad \lambda > 0, a > 0, b > 0, \quad (16)$$

where  $a$  and  $b$  are hyper-parameters. Thus, by incorporating the prior information in Equation (16) to the likelihood function in Equation (6), the posterior distribution of  $\lambda$  can be expressed as

$$\begin{aligned} \pi(\lambda|\underline{x}) &= \frac{L(\underline{x}|\lambda)g(\lambda)}{\int_0^{\infty} L(\underline{x},\lambda)g(\lambda)d\lambda}, \\ &= \frac{K}{\lambda^{\frac{3m}{2}+a+1}} \exp\left[-\frac{1}{\lambda}\left(\sum_{i=1}^m x_i^2 + b\right)\right] \prod_{i=1}^m \left[1 - \Gamma\left(\frac{x_i^2}{\lambda}, \frac{3}{2}\right)\right]^{k(G_i+1)-1}, \end{aligned} \quad (17)$$

where  $K^{-1}$  is the normalizing constant given by

$$K^{-1} = \int_0^{\infty} \frac{1}{\lambda^{\frac{3m}{2}+a+1}} \exp\left[-\frac{1}{\lambda}\left(\sum_{i=1}^m x_i^2 + b\right)\right] \prod_{i=1}^m \left[1 - \Gamma\left(\frac{x_i^2}{\lambda}, \frac{3}{2}\right)\right]^{k(G_i+1)-1} d\lambda.$$

Here, we consider the LINEX loss function proposed by [44], which is one of the most commonly used asymmetric loss functions. The LINEX loss is defined as

$$L(\Delta) = \exp(c\Delta) - c\Delta - 1; \quad c \neq 0, \quad \Delta = \tilde{\lambda} - \lambda, \quad (18)$$

where the loss function's scaling parameter is  $c$  and  $\tilde{\lambda}$  is an estimate of  $\lambda$ . The LINEX loss function provides greater weight to overestimation or underestimation, depending on whether the value of  $c$  is positive or negative, and for small values of  $c \simeq 0$ , it is virtually identical to the squared error loss function. This loss function is appropriate when overestimation is more expensive than underestimating. Thus, under the LINEX loss function, the Bayes estimator of any function of the parameter  $\lambda$ , say  $\phi(\lambda)$ , is given by

$$E[\phi(\lambda)] = -\frac{1}{c} \ln \left[ \frac{\int_0^{\infty} e^{-c\phi(\lambda)} \pi(\lambda|\underline{x}) d\lambda}{\int_0^{\infty} \pi(\lambda|\underline{x}) d\lambda} \right]. \quad (19)$$

From Equation (19), the Bayes estimators take the ratio of two integrals for which has no closed-form solution. To obtain the ratio of the two integrals in Equation (19), we suggest using two approximation techniques—the T-K approximation and the MCMC methods. The T-K approximation method is one of the oldest deterministic approximation techniques, whereas the MCMC method is one of the newest popularized techniques based on the posterior sample algorithm. The MCMC could be expensive to compute, especially for large sample sizes  $n$ . Moreover, many MCMC algorithms require a rough estimate of key posterior quantities, such as the posterior variance. Compared to MCMC methods, the T-K approximation cannot be reduced by running the algorithm longer. However, deterministic approximation is typically very fast to compute and sufficiently reliable in several applied contexts. These issues motivate us to develop both techniques in this study. The details of these two approximation techniques are presented in the following subsections.

### 3.1. Tierney-Kadane (T-K) Approximation Technique

According to T-K approximation’s technique proposed by Tierney and Kadane [45], the approximation of the posterior mean of the function of the parameter, say  $\phi(\lambda)$  is given by

$$E[\phi(\lambda)|x] = \frac{\int_0^\infty e^{n\delta_\phi^*(\lambda)} d\lambda}{\int_0^\infty e^{n\delta(\lambda)} d\lambda} \simeq \left( \frac{|\Sigma_\phi^*|}{|\Sigma|} \right)^{\frac{1}{2}} e^{n[\delta_\phi^*(\hat{\lambda}_\phi^*) - \delta(\hat{\lambda}_\phi)]}, \tag{20}$$

where  $\delta(\lambda) = \frac{1}{n}[l(\lambda) + \rho(\lambda)]$ ,  $\delta^*(\lambda) = \delta(\lambda) + \frac{1}{n} \ln \phi(\lambda)$ ,  $l(\lambda)$  is the log-likelihood function,  $\rho(\lambda) = \ln g(\lambda)$ , and  $|\Sigma_\phi^*|$  and  $|\Sigma|$  are the determinants of inverse of the negative Hessian of  $\delta^*(\lambda)$  and  $\delta(\lambda)$  at  $\hat{\lambda}_{\delta^*}$  and  $\hat{\lambda}_\delta$ , respectively. Here,  $\hat{\lambda}_\delta$  and  $\hat{\lambda}_{\delta^*}$  maximize  $\delta(\lambda)$  and  $\delta^*(\lambda)$ , respectively. We observe that

$$\delta(\lambda) = \frac{1}{n} \left[ - \left( \frac{3m}{2} + a + 1 \right) \ln \lambda - \frac{1}{\lambda} \left( \sum_{i=1}^m x_i^2 + b \right) + 2 \sum_{i=1}^m \ln x_i + \sum_{i=1}^m [k(G_i + 1) - 1] \ln \left( 1 - \Gamma \left( \frac{x_i^2}{\lambda}, \frac{3}{2} \right) \right) \right].$$

To determine the value of  $\hat{\lambda}_\delta$ , we solve the following non-linear equation:

$$\frac{\partial \delta(\lambda)}{\partial \lambda} = \frac{1}{n} \left[ - \left( \frac{3m}{2} + a + 1 \right) \frac{1}{\lambda} + \frac{1}{\lambda^2} \left( \sum_{i=1}^m x_i^2 + b \right) + \sum_{i=1}^m [k(G_i + 1) - 1] \psi(\lambda) \right] = 0,$$

where

$$\psi(\lambda) = \frac{\partial \ln \left[ 1 - \Gamma \left( \frac{x_i^2}{\lambda}, \frac{3}{2} \right) \right]}{\partial \lambda} = - \frac{3}{2\lambda \left[ 1 - \Gamma \left( \frac{x_i^2}{\lambda}, \frac{3}{2} \right) \right]} \left\{ \frac{3}{2} \Gamma \left( \frac{x_i^2}{\lambda}, \frac{5}{2} \right) - \Gamma \left( \frac{x_i^2}{\lambda}, \frac{3}{2} \right) \right\}.$$

Now,  $|\Sigma|$  can be obtained from

$$\begin{aligned} \Sigma^{-1} &= \frac{1}{n} \left( - \frac{\partial^2 \delta(\lambda)}{\partial \lambda^2} \right) \\ &= - \frac{1}{n} \left[ \left( \frac{3m}{2} + a + 1 \right) \frac{1}{\lambda^2} - \frac{2}{\lambda^3} \left( \sum_{i=1}^m x_i^2 + b \right) + \sum_{i=1}^m [k(G_i + 1) - 1] \psi'(\lambda) \right], \end{aligned}$$

where

$$\begin{aligned} \psi'(\lambda) &= \frac{\partial \psi(\lambda)}{\partial \lambda} = - \frac{3}{2\lambda^2} \left\{ \frac{[\{1 - \Gamma(x_i^2/\lambda, 3/2)\} Q_1 + Q_2]}{\{1 - \Gamma(x_i/\lambda, 3/2)\}} \right\}, \\ Q_1 &= \frac{5}{2} \left[ \Gamma(x_i^2/\lambda, 7/2) - 2\Gamma(x_i^2/\lambda, 5/2) + \Gamma(x_i^2/\lambda, 3/2) \right], \\ Q_2 &= \frac{3}{2} \left[ \Gamma(x_i^2/\lambda, 5/2) - \Gamma(x_i^2/\lambda, 3/2) \right]. \end{aligned}$$

To calculate the Bayes estimator of  $\lambda$  under the LINEX loss function, we take  $\phi(\lambda) = e^{-c\lambda}$ , consequently the function  $\delta^*(\lambda)$  becomes

$$\delta^*(\lambda) = \delta(\lambda) - \frac{c\lambda}{n}$$

and then  $\hat{\lambda}_{\delta^*}^*$ , is computed as solutions of the following non-linear equation

$$\frac{\partial \delta^*(\lambda)}{\partial \lambda} = \frac{\partial \delta(\lambda)}{\partial \lambda} - \frac{c}{n} = 0, \quad \text{and obtain } |\Sigma^*| \text{ from } \Sigma_\lambda^{*-1} = -\frac{1}{n} \left( \frac{\partial^2 \delta^*(\lambda)}{\partial \lambda^2} \right).$$

Thus, the approximate Bayes estimator of  $\lambda$  under the LINEX loss function is given by

$$\hat{\lambda}_{TK} = -\frac{1}{c} \ln \left[ \left( \frac{|\Sigma_\lambda^*|}{|\Sigma|} \right)^{\frac{1}{2}} \exp \{ n [\delta_\lambda^*(\hat{\lambda}_\delta^*) - \delta(\hat{\lambda}_\delta)] \} \right].$$

Similarly, the Bayes estimator of Shannon’s entropy  $H(\lambda)$  under the LINEX loss function is given by

$$\hat{H}_{TK} = -\frac{1}{c} \ln \left[ \left( \frac{|\Sigma_H^*|}{|\Sigma|} \right)^{\frac{1}{2}} \exp \{ n [\delta_H^*(\hat{H}_\delta^*) - \delta(\hat{H}_\delta)] \} \right].$$

### 3.2. Markov Chain Monte Carlo (MCMC) Techniques

In this subsection, we use the MCMC techniques to obtain the Bayes estimates of the parameter  $\lambda$  and Shannon’s entropy  $H(\lambda)$  under the LINEX loss function. The Metropolis-Hastings (M-H) algorithm was initially established by Metropolis et al. [46] and subsequently extended by Hastings [47] and popularized as one of the most commonly used MCMC techniques. The candidate points are created from a normal distribution to a sample from the posterior distribution of  $\lambda$  using the observed data  $\mathcal{X}$  in (17). The following steps are used to obtain MCMC sequences:

**Step C1.** For parameter  $\lambda$ , set  $\lambda^{(0)}$  as the initial guess value

**Step C2.** From proposal density  $\eta(\lambda^{(j)} | \lambda^{(j-1)})$ , generate a candidate point  $\lambda_c^{(j)}$ .

**Step C3.** Generate  $u$  from uniform distribution in  $(0, 1)$ .

**Step C4.** Compute  $\alpha(\lambda_c^{(j)} | \lambda^{(j-1)}) = \min \left\{ \frac{\pi(\lambda_c^{(j)} | \mathcal{w}) \eta(\lambda^{(j-1)} | \lambda_c^{(j)})}{\pi(\lambda^{(j-1)} | \mathcal{w}) \eta(\lambda_c^{(j)} | \lambda^{(j-1)})}, 1 \right\}$ .

**Step C5.** If  $u \leq \alpha$ , set  $\lambda^{(j)} = \lambda_c^{(j)}$  with acceptance probability  $\alpha$ ; otherwise, set  $\lambda^{(j)} = \lambda^{(j-1)}$ .

**Step C6.** Compute  $H^{(j)} = H(\lambda^{(j)})$  from Equation (1).

**Step C7.** Repeat Steps C2–C6  $M$  times to obtain the sequence of the parameter  $\lambda$  as  $(\lambda_1, \lambda_2, \dots, \lambda_M)$  and Shannon’s entropy  $H$  as  $(H_1, H_2, \dots, H_M)$ .

To acquire an independent sample from the stationary distribution of the Markov chain, we consider a burn-in period of size  $M_0$  by discarding the first  $M_0$  values in the MCMC sequences. Thus, the Bayes estimators of  $\lambda$  and  $H(\lambda)$  under the LINEX loss function, respectively, are given by

$$\hat{\lambda}_{MH} = -\frac{1}{c} \ln \left[ \frac{1}{M - M_0} \sum_{j=M_0+1}^M e^{-c\lambda_j} \right],$$

$$\text{and } \hat{H}_{MH} = -\frac{1}{c} \ln \left[ \frac{1}{M - M_0} \sum_{j=M_0+1}^M e^{-cH(\lambda_j)} \right].$$

Based on the MCMC samples, we can obtain the HPD credible interval for the parameter  $\lambda$  and Shannon’s entropy  $H(\lambda)$ . Suppose  $(\lambda_{[1]} < \lambda_{[2]} < \dots < \lambda_{[M^*]})$  and  $(H_{[1]} < H_{[2]} < \dots < H_{[M^*]})$  denotes the ordered values of  $(\lambda_1, \lambda_2, \dots, \lambda_{M^*})$  and  $(H_1, H_2, \dots, H_{M^*})$ , respectively, after the burn-in period, where  $M^* = M - M_0$ . Then, following Chen and Shao [48], the  $100(1 - \alpha)\%$  HPD credible interval for  $\lambda$  can be obtained as  $(\lambda_{(j)}, \lambda_{(j+[(1-\alpha)M^*])})$ , where  $j$  is chosen such that

$$\lambda_{j+[(1-\alpha)M^*]} - \lambda_{(j)} = \min_{1 \leq i \leq \alpha M^*} (\lambda_{(i+[(1-\alpha)M^*])} - \lambda_{(i)}); \quad j = 1, 2, \dots, M^*.$$

Similarly, the  $100(1 - \alpha)\%$  HPD credible interval for  $H(\lambda)$  is given by  $(H_{(j)}, H_{(j+[(1-\alpha)M^*])})$ , where  $j$  is chosen such that

$$H_{j+[(1-\alpha)M^*]} - H_{(j)} = \min_{1 \leq i \leq \alpha M^*} (H_{(i+[(1-\alpha)M^*])} - H_{(i)}); \quad j = 1, 2, \dots, M^*.$$

#### 4. Monte Carlo Simulation Study

In this section, a Monte Carlo simulation study is conducted to evaluate the performance of the proposed estimation procedures. The frequentist and Bayesian point estimation procedures for the parameter  $\lambda$  and Shannon’s entropy  $H(\lambda)$  are compared by means of the average estimates (AE) and mean squared errors (MSE). For interval estimation procedures, we compare the asymptotic confidence intervals (Asym), the percentile bootstrap confidence intervals (boot-p), the bootstrap- $t$  confidence intervals (boot-t), and the HPD credible intervals in terms of their simulated average lengths (AL) and the simulated coverage probabilities (CP). For the bootstrap confidence intervals, the intervals are obtained based on  $B = 1000$  bootstrap samples.

In the simulation study, we consider that the PFFC samples are generated from the MWD with parameter  $\lambda = 0.75$  and  $1.5$  (the corresponding entropy  $H(\lambda)$  are  $H(0.75) = 0.5057$  and  $H(1.5) = 0.8523$ , respectively) with various combinations of a number of groups  $n$ , effective sample size  $m$ , group size  $k$ , and censoring scheme  $\underline{G}$ . We consider group sizes  $k = 3$  and  $5$ , the number of groups  $n = 20$  and  $50$ , and effective sample sizes  $m = 0.4n$  and  $0.8n$ . Three different censoring schemes  $\underline{G}$  for each combination of  $n$  and  $m$  are considered:

- (I)  $[(k, n, m), (G_1 = n - m, G_j = 0, \quad \forall j = 2, 3, \dots, m)]:$   $(n - m)$  groups are removed from the experiment at the first failure only;
- (II)  $[(k, n, m), (G_j = 0, \quad \forall j = 1, 2, \dots, m - 1, G_m = n - m)]:$   $(n - m)$  groups are removed at  $m$ th failure;
- (III)  $[(k, n = m), G_j = 0, \quad \forall j = 1, 2, \dots, m]:$  first-failure censored sample.

The censoring schemes ([CS]) used in the Monte Carlo simulation study are summarized in Table 1. Note that simplified notations are used to denote the censoring schemes, for example,  $(0_7)$  denotes  $(0, 0, 0, 0, 0, 0, 0)$  and  $(4_3)$  stands for  $(4, 4, 4)$ .

**Table 1.** Censoring plans used in the simulation study.

n	m	[CS]	Schemes	n	m	[CS]	Schemes
20	8	[1]	$(12_1, 0_7)$	50	20	[7]	$(30_1, 0_{19})$
		[2]	$(4_3, 0_5)$			[8]	$(5_6, 0_{14})$
		[3]	$(0_7, 12_1)$			[9]	$(0_{19}, 30_1)$
20	16	[4]	$(4_1, 0_{15})$	50	40	[10]	$(10_1, 0_{39})$
		[5]	$(2_2, 0_{14})$			[11]	$(5_2, 0_{38})$
		[6]	$(0_{15}, 4_1)$			[12]	$(0_{39}, 10_1)$

For the Bayesian estimation approach, the Bayes estimates of Shannon’s entropy are computed with informative inverted gamma prior under the LINEX loss function. The hyper-parameters  $(a, b)$  are selected for Bayesian computations of the parameter  $\lambda$  and Shannon’s entropy in such a manner that the prior mean is precisely identical to the true values of the parameter, i.e.,  $\lambda = a/b$ . Specifically, we consider  $(a, b) = (3, 4)$  and  $(3, 2)$  for  $\lambda = 0.75$  and  $1.5$ , respectively. When computing Bayes estimators under the LINEX loss function, we consider the loss function parameter  $c = -0.5$  and  $0.5$ . We use  $M = 10,000$  with a burn-in period  $M_0 = 2000$  for the M-H algorithm. The simulation results are based on 1000 repetitions in this study. All the computations are performed using the statistical software R (<https://www.r-project.org/>) [49].

The simulated results for point estimation are presented in Tables 2 and 3, and the simulation results for interval estimation are presented in Tables 4 and 5. For point estimation, from Tables 2 and 3, we observe that the MLEs and Bayes estimates of Shannon’s entropy are performing well with small MSEs. The simulated MSEs decrease as  $n$  or  $m$

increases. The Bayes estimates perform better than the MLEs in terms of MSEs when the prior information matches the true value of  $\lambda$ . Among the two approaches to obtain the Bayes estimates, the Bayes estimates obtained by the M-H algorithm techniques outperform the Bayes estimates obtained by the T-K approximation in terms of the MSEs.

For interval estimation, from Tables 4 and 5, it can be seen that the simulated average lengths of the 95% asymptotic, percentile bootstrap, bootstrap- $t$  confidence intervals, and the Bayesian HPD credible intervals are decreasing as the number of failures ( $m$ ) increases. All the interval estimation procedures provide reasonable simulated coverage probability (CP) that are close to the nominal level 95%. The HPD credible intervals have smaller simulated average lengths than those frequentist confidence intervals.

**Table 2.** Simulated average MLEs and Bayes estimates and the MSEs of Shannon’s entropy  $H(\lambda)$  when  $\lambda = 1.5$  and  $H(1.5) = 0.8523$ .

$(k, n, m)$	[CS]	MLE		TK Approximation				MCMC			
				$c = -0.50$		$c = 0.50$		$c = -0.50$		$c = 0.50$	
		AE	MSE	AE	MSE	AE	MSE	AE	MSE	AE	MSE
(3,20,8)	[1]	0.8372	0.0189	0.8066	0.0174	0.7991	0.0181	0.8329	0.0024	0.8298	0.0025
	[2]	0.8371	0.0185	0.8074	0.0171	0.8000	0.0178	0.8333	0.0024	0.8302	0.0025
	[3]	0.8371	0.0172	0.8101	0.0160	0.8031	0.0166	0.8342	0.0023	0.8314	0.0024
(3,20,16)	[4]	0.8475	0.0091	0.8305	0.0087	0.8263	0.0089	0.8413	0.0016	0.8394	0.0017
	[5]	0.8475	0.0091	0.8305	0.0087	0.8263	0.0089	0.8413	0.0016	0.8394	0.0017
	[6]	0.8475	0.0088	0.8312	0.0084	0.8272	0.0086	0.8416	0.0016	0.8398	0.0016
(3,50,20)	[7]	0.8485	0.0074	0.8348	0.0071	0.8314	0.0073	0.8433	0.0014	0.8418	0.0014
	[8]	0.8485	0.0072	0.8353	0.0070	0.8320	0.0071	0.8435	0.0014	0.8420	0.0014
	[9]	0.8483	0.0066	0.8366	0.0064	0.8336	0.0065	0.8442	0.0013	0.8428	0.0013
(3,50,40)	[10]	0.8496	0.0039	0.8425	0.0039	0.8407	0.0039	0.8472	0.0009	0.8463	0.0009
	[11]	0.8496	0.0039	0.8425	0.0039	0.8407	0.0039	0.8472	0.0009	0.8463	0.0009
	[12]	0.8495	0.0038	0.8428	0.0037	0.8411	0.0038	0.8474	0.0008	0.8465	0.0008
(5,20,8)	[1]	0.8371	0.0180	0.8505	0.0124	0.8011	0.0174	0.8483	0.0023	0.8306	0.0025
	[2]	0.8370	0.0177	0.8504	0.0122	0.8020	0.0171	0.8484	0.0023	0.8310	0.0024
	[3]	0.8369	0.0166	0.8503	0.0117	0.8046	0.0161	0.8485	0.0022	0.8319	0.0023
(5,20,16)	[4]	0.8474	0.0087	0.8536	0.0072	0.8276	0.0085	0.8511	0.0016	0.8400	0.0016
	[5]	0.8474	0.0086	0.8536	0.0072	0.8276	0.0085	0.8511	0.0016	0.8400	0.0016
	[6]	0.8474	0.0084	0.8536	0.0070	0.8283	0.0082	0.8511	0.0015	0.8403	0.0016
(5,50,20)	[7]	0.8485	0.0070	0.8535	0.0061	0.8325	0.0069	0.8515	0.0013	0.8422	0.0014
	[8]	0.8484	0.0069	0.8535	0.0060	0.8329	0.0068	0.8515	0.0013	0.8424	0.0014
	[9]	0.8482	0.0064	0.8534	0.0056	0.8342	0.0063	0.8515	0.0012	0.8431	0.0013
(5,50,40)	[10]	0.8495	0.0038	0.8521	0.0035	0.8413	0.0037	0.8517	0.0008	0.8466	0.0008
	[11]	0.8495	0.0038	0.8521	0.0035	0.8413	0.0037	0.8517	0.0008	0.8466	0.0008
	[12]	0.8495	0.0036	0.8520	0.0034	0.8416	0.0036	0.8517	0.0008	0.8468	0.0008

**Table 3.** Simulated average MLEs and Bayes estimates and the MSEs of Shannon’s entropy  $H(\lambda)$ , when  $\lambda = 0.75$  and  $H = 0.5057$ .

$(k, n, m)$	[CS]	MLE		TK Approximation				MCMC			
				$c = -0.50$		$c = 0.50$		$c = -0.50$		$c = 0.50$	
		AE	MSE	AE	MSE	AE	MSE	AE	MSE	AE	MSE
(3,20,8)	[1]	0.4906	0.0189	0.5823	0.0151	0.5746	0.0139	0.5371	0.0037	0.5338	0.0035
	[2]	0.4905	0.0185	0.5811	0.0148	0.5735	0.0137	0.5365	0.0036	0.5333	0.0035
	[3]	0.4905	0.0172	0.5768	0.0138	0.5696	0.0129	0.5347	0.0034	0.5317	0.0032
(3,20,16)	[4]	0.5009	0.0091	0.5507	0.0083	0.5465	0.0079	0.5252	0.0021	0.5232	0.0021
	[5]	0.5009	0.0091	0.5506	0.0083	0.5464	0.0079	0.5251	0.0021	0.5232	0.0020
	[6]	0.5009	0.0088	0.5493	0.0080	0.5453	0.0077	0.5246	0.0020	0.5227	0.0020
(3,50,20)	[7]	0.5020	0.0074	0.5424	0.0068	0.5390	0.0066	0.5219	0.0017	0.5203	0.0017
	[8]	0.5019	0.0072	0.5417	0.0067	0.5384	0.0064	0.5215	0.0017	0.5200	0.0017
	[9]	0.5018	0.0066	0.5391	0.0062	0.5360	0.0060	0.5204	0.0016	0.5190	0.0015
(3,50,40)	[10]	0.5030	0.0039	0.5242	0.0037	0.5225	0.0037	0.5142	0.0009	0.5134	0.0009
	[11]	0.5030	0.0039	0.5242	0.0037	0.5224	0.0037	0.5142	0.0009	0.5134	0.0009
	[12]	0.5030	0.0038	0.5235	0.0036	0.5218	0.0035	0.5139	0.0009	0.5131	0.0009

Table 3. Cont.

$(k, n, m)$	[CS]	MLE		TK Approximation				MCMC			
		AE	MSE	$c = -0.50$		$c = 0.50$		$c = -0.50$		$c = 0.50$	
				AE	MSE	AE	MSE	AE	MSE	AE	MSE
(5,20,8)	[1]	0.4905	0.0180	0.5795	0.0144	0.5721	0.0134	0.5359	0.0036	0.5327	0.0034
	[2]	0.4904	0.0177	0.5785	0.0142	0.5712	0.0132	0.5354	0.0035	0.5323	0.0033
	[3]	0.4903	0.0166	0.5749	0.0134	0.5679	0.0125	0.5339	0.0033	0.5309	0.0031
(5,20,16)	[4]	0.5008	0.0087	0.5488	0.0079	0.5448	0.0076	0.5243	0.0020	0.5225	0.0020
	[5]	0.5008	0.0086	0.5488	0.0079	0.5447	0.0076	0.5243	0.0020	0.5225	0.0020
	[6]	0.5008	0.0084	0.5476	0.0077	0.5437	0.0074	0.5238	0.0020	0.5220	0.0019
(5,50,20)	[7]	0.5019	0.0070	0.5409	0.0065	0.5376	0.0063	0.5212	0.0017	0.5197	0.0016
	[8]	0.5018	0.0069	0.5403	0.0064	0.5371	0.0062	0.5209	0.0016	0.5194	0.0016
	[9]	0.5016	0.0064	0.5381	0.0060	0.5351	0.0058	0.5199	0.0015	0.5185	0.0015
(5,50,40)	[10]	0.5029	0.0038	0.5233	0.0035	0.5216	0.0035	0.5138	0.0009	0.5130	0.0009
	[11]	0.5029	0.0038	0.5233	0.0035	0.5216	0.0035	0.5138	0.0009	0.5130	0.0009
	[12]	0.5029	0.0036	0.5227	0.0034	0.5211	0.0034	0.5135	0.0009	0.5128	0.0009

Table 4. Simulated average lengths and coverage probabilities of 95% asymptotic, percentile bootstrap, and bootstrap- $t$  confidence interval and Bayesian HPD credible intervals of Shannon’s entropy  $H(\lambda)$ , when  $\lambda = 1.5$  and  $H = 0.8523$ .

$(k, n, m)$	[CS]	Asym		Bootstrap				HPD	
		AL	CP	Boot-p		Boot-t		AL	CP
				AL	CP	AL	CP		
(3,20,8)	[1]	0.5257	0.948	0.5324	0.927	0.5335	0.952	0.3055	0.999
	[2]	0.5198	0.947	0.5271	0.927	0.5279	0.952	0.3027	0.999
	[3]	0.5007	0.951	0.5105	0.925	0.5103	0.951	0.2941	0.999
(3,20,16)	[4]	0.3746	0.948	0.3721	0.945	0.3721	0.950	0.2387	0.998
	[5]	0.3743	0.948	0.3719	0.945	0.3718	0.950	0.2385	0.998
	[6]	0.3680	0.945	0.3657	0.946	0.3658	0.951	0.2350	0.998
(3,50,20)	[7]	0.3343	0.945	0.3458	0.955	0.3459	0.954	0.2170	0.997
	[8]	0.3304	0.944	0.3419	0.956	0.3416	0.954	0.2146	0.998
	[9]	0.3165	0.944	0.3273	0.954	0.3273	0.953	0.2065	0.997
(3,50,40)	[10]	0.2372	0.941	0.2472	0.950	0.2472	0.950	0.1602	0.987
	[11]	0.2371	0.941	0.2471	0.950	0.2471	0.950	0.1601	0.987
	[12]	0.2327	0.942	0.2427	0.950	0.2427	0.950	0.1573	0.986
(5,20,8)	[1]	0.5129	0.948	0.5210	0.925	0.5219	0.952	0.3050	0.997
	[2]	0.5081	0.947	0.5162	0.925	0.5175	0.953	0.3027	0.997
	[3]	0.4917	0.950	0.5024	0.923	0.5023	0.951	0.2950	0.997
(5,20,16)	[4]	0.3651	0.946	0.3628	0.945	0.3629	0.951	0.2345	0.997
	[5]	0.3649	0.946	0.3626	0.945	0.3627	0.951	0.2344	0.997
	[6]	0.3592	0.945	0.3576	0.946	0.3577	0.951	0.2312	0.997
(5,50,20)	[7]	0.3260	0.946	0.3369	0.953	0.3370	0.953	0.2130	0.997
	[8]	0.3228	0.946	0.3336	0.954	0.3336	0.953	0.2111	0.997
	[9]	0.3109	0.944	0.3215	0.953	0.3214	0.954	0.2041	0.996
(5,50,40)	[10]	0.2312	0.942	0.2409	0.950	0.2409	0.950	0.1565	0.988
	[11]	0.2311	0.942	0.2408	0.950	0.2408	0.950	0.1564	0.988
	[12]	0.2272	0.942	0.2368	0.949	0.2368	0.950	0.1539	0.988

**Table 5.** Simulated average lengths and coverage probabilities of 95% asymptotic, bootstrap confidence and HPD credible intervals of entropy  $H(\lambda)$ , when  $\lambda = 0.75$  and  $H = 0.5057$ .

$(k, n, m)$	Asym	Bootstrap						HPD	
		[CS]	Boot-p		Boot-t		AL	CP	
		AL	CP	AL	CP	AL	CP	AL	CP
(3,20,8)	[1]	0.5219	0.948	0.5256	0.927	0.5312	0.952	0.3158	0.987
	[2]	0.5163	0.947	0.5207	0.927	0.5258	0.952	0.3133	0.988
	[3]	0.4981	0.951	0.5055	0.925	0.5089	0.951	0.3051	0.988
(3,20,16)	[4]	0.3746	0.948	0.3721	0.945	0.3722	0.950	0.2413	0.988
	[5]	0.3743	0.948	0.3719	0.945	0.3718	0.950	0.2411	0.988
	[6]	0.3680	0.945	0.3657	0.946	0.3658	0.951	0.2377	0.987
(3,50,20)	[7]	0.3343	0.945	0.3458	0.955	0.3459	0.954	0.2189	0.989
	[8]	0.3304	0.945	0.3419	0.956	0.3416	0.954	0.2167	0.989
	[9]	0.3165	0.944	0.3273	0.954	0.3273	0.953	0.2089	0.989
(3,50,40)	[10]	0.2372	0.941	0.2472	0.950	0.2472	0.950	0.1607	0.983
	[11]	0.2371	0.941	0.2471	0.950	0.2471	0.950	0.1606	0.983
	[12]	0.2327	0.942	0.2427	0.950	0.2427	0.950	0.1579	0.985
(5,20,8)	[1]	0.5098	0.948	0.5151	0.925	0.5201	0.952	0.3104	0.988
	[2]	0.5052	0.947	0.5107	0.924	0.5158	0.953	0.3083	0.988
	[3]	0.4895	0.950	0.4979	0.923	0.5011	0.951	0.3012	0.988
(5,20,16)	[4]	0.3651	0.946	0.3628	0.945	0.3629	0.951	0.2362	0.989
	[5]	0.3649	0.946	0.3625	0.945	0.3626	0.951	0.2361	0.989
	[6]	0.3592	0.945	0.3576	0.946	0.3577	0.951	0.2331	0.99
(5,50,20)	[7]	0.3260	0.946	0.3369	0.953	0.3370	0.953	0.2143	0.989
	[8]	0.3228	0.946	0.3336	0.954	0.3336	0.953	0.2125	0.989
	[9]	0.3109	0.944	0.3215	0.952	0.3214	0.954	0.2056	0.989
(5,50,40)	[10]	0.2312	0.942	0.2409	0.950	0.2409	0.950	0.1569	0.985
	[11]	0.2311	0.942	0.2408	0.950	0.2408	0.950	0.1568	0.985
	[12]	0.2272	0.942	0.2368	0.949	0.2369	0.950	0.1544	0.985

### 5. Practical Data Analysis

To demonstrate the effectiveness of the MWD in modeling lifetime data and illustrate the methodologies developed in the paper, a practical data analysis of a real data set is conducted. We consider the tensile strength (in GPa) of 100 carbon fibers. This data set was originally reported by Nichols and Padgett [50] and further studied by Mohammed et al. [25] and Xie and Gui [51]. The data set is presented in Table 6.

**Table 6.** Tensile strength of 100 carbon fibers originally reported by Nichols and Padgett [50].

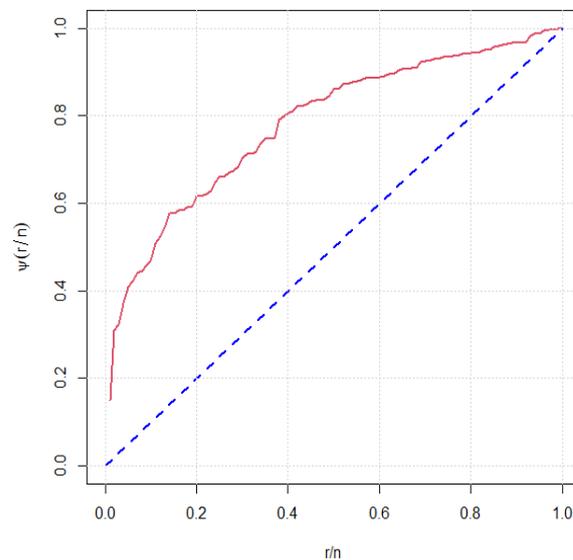
3.70	3.11	4.42	3.28	3.75	2.96	3.39	3.31	3.15	2.81	1.41	2.76	3.19	1.59	2.17
3.51	1.84	1.61	1.57	1.89	2.74	3.27	2.41	3.09	2.43	2.53	2.81	3.31	2.35	2.77
2.68	4.91	1.57	2.00	1.17	2.17	0.39	2.79	1.08	2.88	2.73	2.87	3.19	1.87	2.95
2.67	4.20	2.85	2.55	2.17	2.97	3.68	0.81	1.22	5.08	1.69	3.68	4.70	2.03	2.82
2.50	1.47	3.22	3.15	2.97	2.93	3.33	2.56	2.59	2.83	1.36	1.84	5.56	1.12	2.48
1.25	2.48	2.03	1.61	2.05	3.60	3.11	1.69	4.90	3.39	3.22	2.55	3.56	2.38	1.92
0.98	1.59	1.73	1.71	1.18	4.38	0.85	1.80	2.12	3.65					

First, we use the scaled total time on test (TTT) transform to understand the behavior of the failure rate function of the data set. The scaled TTT transform is given by

$$\psi(r/n) = \left[ \sum_{j=1}^r t_{(j)} + (n-r)t_r \right] / \left( \sum_{j=1}^r t_{(j)} \right), \quad r = 1, 2, \dots, n,$$

where  $t_{(i)}, i = 1, 2, \dots, n$  represent the  $i$ -th order statistic of the sample. If the plot  $(r/n, \psi(r/n))$  is convex (concave), the failure rate function has a decreasing (increasing) shape. For more details about scaled TTT transform, see, for example, Mudholkar et al. [52].

The scaled TTT plot of the data set in Table 6 is displayed in Figure 2. Figure 2 shows that the considered data set follows an increasing failure rate function. This empirical behavior of the failure rate function indicates that the MWD model can be considered a suitable model for this data set.



**Figure 2.** TTT plots under consideration of real data set.

Furthermore, we check whether the MWD is well-fit for the data set in Table 6 using two goodness-of-fit tests. We consider the Kolmogorov–Smirnov (KS) and Anderson–Darling (AD) test statistics and obtain the corresponding  $p$ -values. The KS and AD statistics with their corresponding  $p$ -values (in parentheses) are 0.0884 (0.4145) and 0.7977 (0.4824), respectively. According to these  $p$ -values of the goodness-of-fit tests, the MWD fits quite well for the data set in Table 6. In addition to the goodness-of-fit test, we also assess the feasibility of fitting the data set using MWD graphically using the empirical and fitted cdfs plot and the probability-probability (P-P) plots in Figure 3. These plots are tools used in statistics to assess the goodness-of-fit of a statistical model to observed data. A good fit is indicated when the points in the P-P plot lie close to a straight line (usually a 45-degree line), suggesting that the empirical and theoretical cumulative probabilities are a good fit or similar. From Figure 3, one can observe that the observed data points show almost similar patterns to the theoretical distribution, i.e., the MWD fits the data set in Table 6 reasonably well.

To illustrate the methodologies developed in this paper, we generate first-failure censored data based on the data set in Table 6. After grouping the 100 carbon fiber into  $n = 25$  groups and  $k = 4$  individuals within each group. The grouped data and the corresponding first-failure censored samples are reported in Table 7. The items with “+” within each group indicate the first failure. Then, we obtain six different PFFC samples using different censoring schemes for  $m = 10$  and 20 based on the first-failure censored data in Table 7. The censoring schemes and the corresponding PFFC samples are presented in Table 8. To avoid ambiguity with the censoring schemes in Table 1, we named these censored schemes [CS1]–[CS6].

Based on each PFFC sample in Table 8, we compute the MLEs and Bayes estimates of the parameter  $\lambda$  and Shannon’s entropy  $H(\lambda)$ . As we do not have prior information on the parameter  $\lambda$ , we use a non-informative prior for obtaining the Bayes estimates. The Bayes estimates are computed using the T-K approximation and MCMC methods under the LINEX loss function at two values of loss parameter  $c = -0.5$  and 0.5. We construct 95% asymptotic, percentile bootstrap, bootstrap- $t$  confidence intervals and the Bayesian

HPD credible intervals for the parameter  $\lambda$  and Shannon’s entropy  $H(\lambda)$ . The point and interval estimation results are presented in Tables 9 and 10, respectively.

For the Bayesian estimation procedures, we validate the convergence of the generated MCMC sequence of the parameter  $\lambda$  samples from the posterior distribution by using the M-H algorithm for their stationary distributions using graphical diagnostic tools, such as the trace plot, boxplot, and histogram with Gaussian density plots, as shown in Figure 4 for  $c = -0.5$ . The trace plot shows a random scatter around the mean (shown by a thick red line) and a fine mixture of parameter chains. The posterior distribution is almost symmetric, as seen by the boxplots and histograms of produced samples, implying that the posterior mean can be used as a reasonable Bayes estimate of the parameter  $\lambda$ . For illustration, we also present the simulated posterior predictive densities in Figure 5.

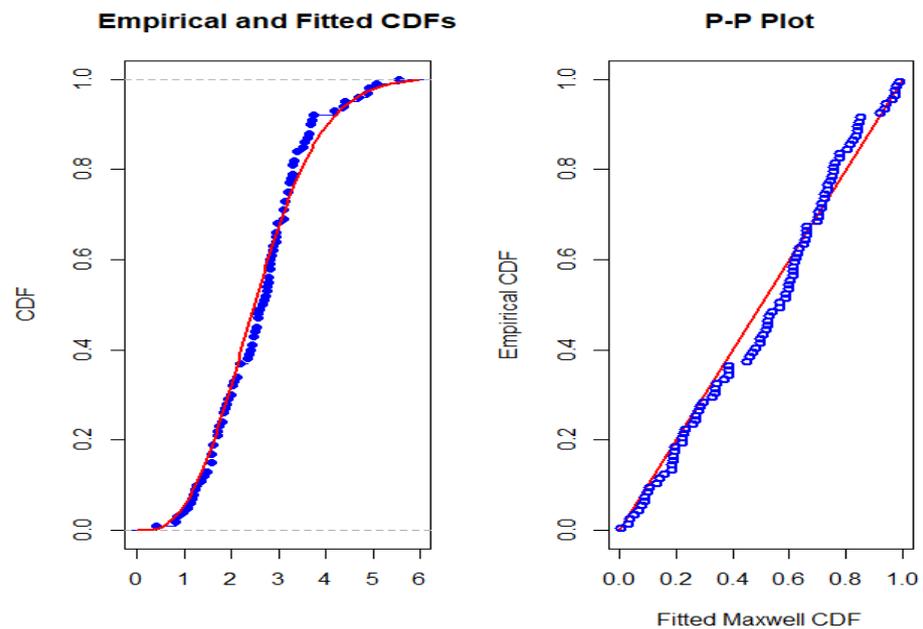


Figure 3. Empirical and fitted Maxwell distribution plots for real data.

Table 7. Grouped real data set (Observation with “+” indicates the first failure (FF) in the group).

Groups → Items↓	1	2	3	4	5	6	7	8	9	10	11	12	13
I	3.27	2.05	3.33	1.87	2.03	3.68	2.87	2.67 +	1.84	1.73	2.82	2.77	2.73
II	2.97	1.61	4.38	2.97	2.12	2.68	1.59 +	2.96	0.39 +	3.19	2.41 +	2.17	2.88
III	3.11	3.11	1.69	1.57 +	0.85 +	4.90	1.89	3.09	2.17	1.57 +	3.60	3.51	3.75
IV	2.03 +	1.25 +	1.18 +	1.59	1.84	2.38 +	2.43	4.20	2.35	2.93	3.22	1.08 +	1.69 +
FF Obs.	2.03	1.25	1.18	1.57	0.85	2.38	1.59	2.67	0.39	1.57	2.41	1.08	1.69
Groups → Items↓	14	15	16	17	18	19	20	21	22	23	24	25	
I	5.56	2.81	2.55	4.70	1.36 +	3.22	1.61	3.15	4.91	3.31	2.76	0.98 +	
II	1.41 +	3.39	2.17	2.59	2.83	1.71 +	2.85	4.42	1.17	1.92	5.08	3.39	
III	2.48	3.68	3.56	3.19	2.74	3.65	1.47 +	2.00 +	1.12 +	1.80 +	3.28	2.50	
IV	2.95	0.81 +	1.22 +	2.56 +	2.53	3.15	2.79	2.81	3.70	2.55	2.48 +	3.31	
FF Obs.	1.41	0.81	1.22	2.56	1.36	1.71	1.47	2.00	1.12	1.80	2.48	0.98	

**Table 8.** Censoring schemes and progressively first-failure censored samples corresponding to considered real data set.

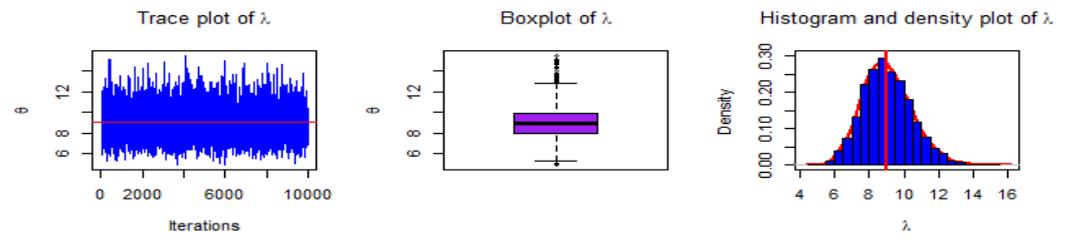
$(k, n, m)$	[CS]	Schemes	Progressively First-Failure Censored Samples
(4,25,10)	[CS1]	$(15_1, 0_9)$	0.39, 1.80, 1.84, 2.03, 2.12, 2.17, 2.48, 2.50, 2.73, 2.77
	[CS2]	$(5_3, 0_7)$	0.39, 1.18, 1.57, 2.03, 2.12, 2.17, 2.48, 2.50, 2.73, 2.77
	[CS3]	$(0_9, 15_1)$	0.39, 0.81, 0.85, 0.98, 1.08, 1.12, 1.18, 1.22, 1.25, 1.36
(4,25,20)	[CS4]	$(5_1, 0_{19})$	0.39, 1.18, 1.22, 1.25, 1.36, 1.41, 1.47, 1.57, 1.59, 1.61, 1.69, 1.80, 1.84, 2.03, 2.12, 2.17, 2.48, 2.50, 2.73, 2.77
	[CS5]	$(2_1, 3_1, 0_{18})$	0.39, 0.98, 1.22, 1.25, 1.36, 1.41, 1.47, 1.57, 1.59, 1.61, 1.69, 1.80, 1.84, 2.03, 2.12, 2.17, 2.48, 2.50, 2.73, 2.77
	[CS6]	$(0_{19}, 5_1)$	0.39, 0.81, 0.85, 0.98, 1.08, 1.12, 1.18, 1.22, 1.25, 1.36, 1.41, 1.47, 1.57, 1.59, 1.61, 1.69, 1.80, 1.84, 2.03, 2.12

**Table 9.** MLEs and Bayes estimates of  $\lambda$  and  $H(\lambda)$  for the PFFC samples in Table 8 with  $k = 4, n = 25, m = 10$  and  $20$ , and  $c = -0.5$  and  $0.5$ .

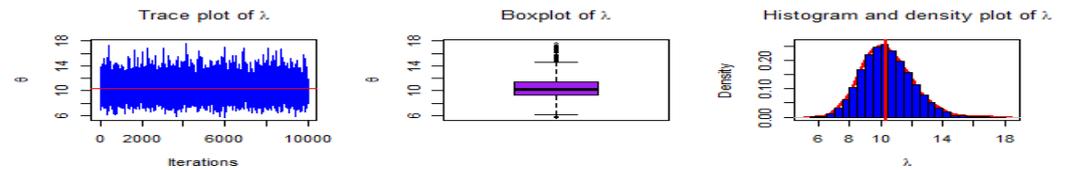
$(k, n, m)$	[CS]	T-K Bayes						M-H Bayes			
		MLE		$c = -0.5$		$c = 0.5$		$c = -0.5$		$c = 0.5$	
		$\hat{\lambda}$	$\hat{H}$	$\hat{\lambda}$	H	$\hat{\lambda}$	$\hat{H}$	$\hat{\lambda}$	$\hat{H}$	$\hat{\lambda}$	$\hat{H}$
(4,25,10)	[CS1]	9.2897	1.7640	13.4153	1.7800	8.7078	1.7714	9.6163	1.7467	8.5906	1.7436
	[CS2]	10.6695	1.8333	11.5620	1.8493	9.9031	1.8411	11.1367	1.8167	9.8262	1.8137
	[CS3]	5.6674	1.5169	6.5958	1.5330	5.5463	1.5304	5.7261	1.5016	5.3784	1.4988
(4,25,20)	[CS4]	6.6806	1.5992	7.2764	1.6083	6.5572	1.6046	6.7445	1.5903	6.4575	1.5887
	[CS5]	6.7637	1.6054	7.3707	1.6145	6.6362	1.6107	6.8301	1.5965	6.5371	1.5949
	[CS6]	5.7635	1.5254	6.2046	1.5344	5.6944	1.5308	5.8011	1.5169	5.5937	1.5153

**Table 10.** Ninety-five percent asymptotic (Asym), percentile bootstrap (boot-p), bootstrap- $t$  (boot-t) confidence intervals, and the Bayesian HPD credible intervals for the parameter  $\lambda$  and Shannon’s entropy  $H(\lambda)$  for the PFFC samples in Table 8 with  $k = 4, n = 25, m = 10$  and  $20$ , and  $c = -0.5$  and  $0.5$ .

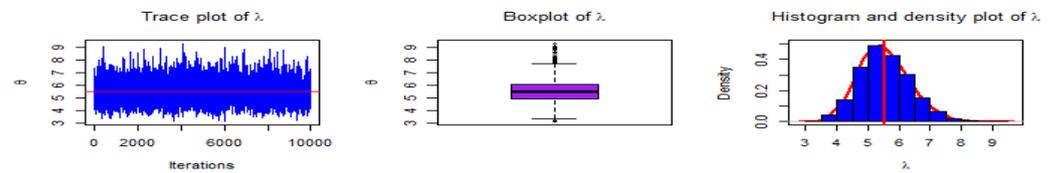
Interval estimates for parameter $\lambda$						
$(k, n, m)$	[CS]	Asym	Bootstrap		HPD	
			boot-p	boot-t	$c = -0.5$	$c = 0.5$
(4,25,10)	[CS1]	(4.973, 13.606)	(6.636, 12.366)	(6.213, 11.944)	(6.357, 11.872)	(6.357, 11.872)
	[CS2]	(5.804, 15.535)	(7.816, 13.602)	(7.736, 13.523)	(7.365, 13.587)	(7.365, 13.587)
	[CS3]	(3.157, 8.178)	(3.880, 8.365)	(2.970, 7.454)	(3.967, 7.187)	(3.967, 7.187)
(4,25,20)	[CS4]	(4.478, 8.883)	(5.006, 8.771)	(4.590, 8.355)	(5.125, 8.058)	(5.125, 8.058)
	[CS5]	(4.538, 8.989)	(5.044, 8.824)	(4.703, 8.483)	(5.192, 8.156)	(5.192, 8.156)
	[CS6]	(3.893, 7.634)	(4.064, 7.962)	(3.565, 7.463)	(4.443, 6.937)	(4.443, 6.938)
Interval estimates for Shannon’s entropy $H(\lambda)$						
$(k, n, m)$	[CS]	Asym	Bootstrap		HPD	
			boot-p	boot-t	$c = -0.5$	$c = 0.5$
(4,25,10)	[CS1]	(1.532, 1.996)	(1.596, 1.907)	(1.621, 1.932)	(1.591, 1.900)	(1.591, 1.900)
	[CS2]	(1.605, 2.061)	(1.678, 1.955)	(1.712, 1.989)	(1.664, 1.968)	(1.664, 1.968)
	[CS3]	(1.295, 1.738)	(1.327, 1.712)	(1.322, 1.706)	(1.353, 1.648)	(1.353, 1.648)
(4,25,20)	[CS4]	(1.434, 1.764)	(1.455, 1.735)	(1.463, 1.743)	(1.472, 1.698)	(1.472, 1.698)
	[CS5]	(1.441, 1.770)	(1.459, 1.738)	(1.472, 1.752)	(1.478, 1.704)	(1.478, 1.704)
	[CS6]	(1.363, 1.688)	(1.351, 1.687)	(1.364, 1.700)	(1.400, 1.623)	(1.400, 1.623)



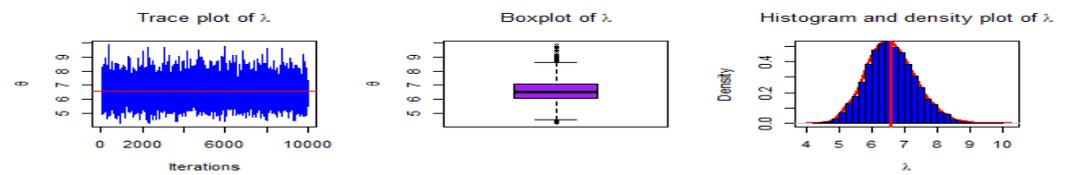
(a) Censoring scheme [CS1].



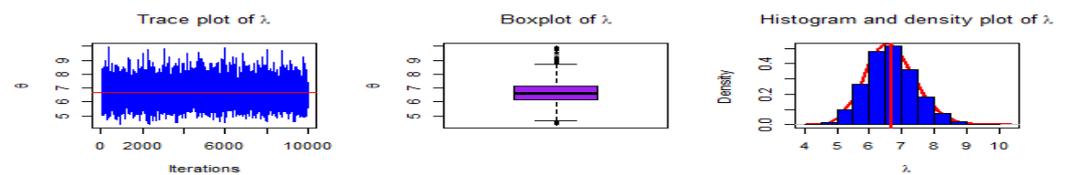
(b) Censoring scheme [CS2].



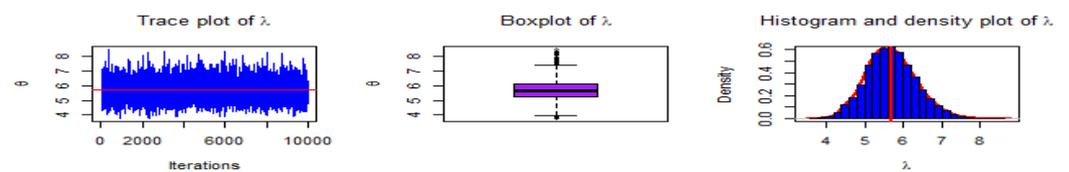
(c) Censoring scheme [CS3].



(d) Censoring scheme [CS4].

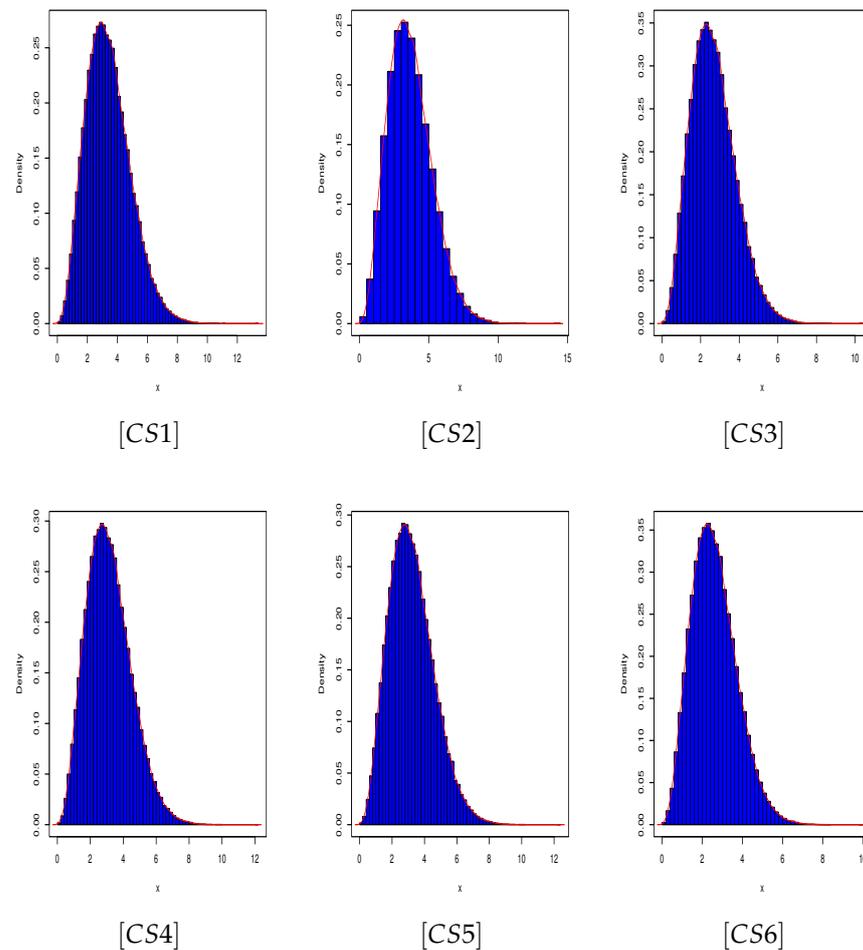


(e) Censoring scheme [CS5].



(f) Censoring scheme [CS6].

**Figure 4.** MCMC diagnostic plots for different PFFC samples in Table 8.



**Figure 5.** Simulated posterior predictive densities for the Bayesian estimates based on different PFFC samples in Table 8.

## 6. Concluding Remarks

In this study, we developed statistical inference for the associated parameter and Shannon's entropy of MWD based on PFFC data using frequentist and Bayesian approaches. For frequentist estimation procedures, we applied the EM algorithm to compute the MLEs. In addition, we obtained asymptotic, percentile bootstrap, and bootstrap- $t$  confidence intervals for the parameter and Shannon's entropy of MWD. For Bayesian estimation procedures, we applied two approximation techniques—the T-K approximation and the M-H algorithm—to obtain the Bayes estimates under the LINEX loss function. Moreover, we use the M-H algorithm to obtain the HPD credible intervals for the parameter and Shannon's entropy of MWD. A Monte Carlo simulation study is used to evaluate the performance of the estimation procedures and a real data analysis is used to illustrate the methodologies. Based on the simulation results, we recommend using the Bayesian point and interval estimation of the parameter and Shannon's entropy based on the MCMC method for the MWD when prior information about the model parameter is available. If the prior information about the model parameter is unavailable, then the MLEs are recommended.

For future research, it will be interesting to study the optimal censoring plan (i.e., to determine the values of  $n$ ,  $k$ , and the censoring scheme  $\mathcal{G} = (G_1, G_2, \dots, G_m)$ ) for specific effective sample size  $m$  and the total sample size available for the experiment (i.e.,  $nk$ ) to maximize the information about the Shannon's entropy can be obtained. On the other hand, we consider MWD due to its simplicity. Considering the trade-off between model

performance/usability and complexity, if one is willing to use a more complex model with higher flexibility, the current work can be extended to some generalized distributions.

**Author Contributions:** Conceptualization, Methodology and Original Draft Preparation: K.K. and I.K.; Investigation, Writing, Review, and Editing: H.K.T.N. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Acknowledgments:** The authors would like to thank the two anonymous referees for their valuable comments, which helped improve the quality of this article.

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

## References

1. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
2. Saraiva, P. On Shannon entropy and its applications. *Kuwait J. Sci.* **2023**, *50*, 194–199. [CrossRef]
3. Cover, T.M. *Elements of Information Theory*; John Wiley & Sons: Hoboken, NJ, USA, 1999.
4. Kayal, S.; Kumar, S. Estimation of the Shannon's entropy of several shifted exponential populations. *Stat. Probab. Lett.* **2013**, *83*, 1127–1135. [CrossRef]
5. Cho, Y.; Sun, H.; Lee, K. An estimation of the entropy for a Rayleigh distribution based on doubly-generalized Type-II hybrid censored samples. *Entropy* **2014**, *16*, 3655–3669. [CrossRef]
6. Du, Y.; Guo, Y.; Gui, W. Statistical Inference for the Information Entropy of the Log-Logistic Distribution under Progressive Type-I Interval Censoring Schemes. *Symmetry* **2018**, *10*, 445. [CrossRef]
7. Liu, S.; Gui, W. Estimating the entropy for Lomax distribution based on generalized progressively hybrid censoring. *Symmetry* **2019**, *11*, 1219. [CrossRef]
8. Yu, J.; Gui, W.; Shan, Y. Statistical Inference on the Shannon Entropy of Inverse Weibull Distribution under the Progressive First-Failure Censoring. *Entropy* **2019**, *21*, 1209. [CrossRef]
9. Bekker, A.; Roux, J. Reliability characteristics of the Maxwell distribution: A Bayes estimation study. *Commun.-Stat.-Theory Methods* **2005**, *34*, 2169–2178. [CrossRef]
10. Pozo, A.; Zhang, S.; Stecklein, G.; Garcia, R.; Glaser, J.; Tang, Z.; Strittmatter, R. *GaN Reliability and Lifetime Projections: Phase 14*; Technical report; Reliability Report; EPC Corp.: El Segundo, CA, USA, 2002.
11. Woolley, R.v.R. A note on the MVU estimation of reliability for the Maxwell failure distribution. *Estadística* **1989**, *41*, 73–79.
12. Harter, H.L. *Order Statistics and their Use in Testing and Estimation, Volume 2*; Technical report; Aerospace Research Laboratories, Office of Aerospace Research, United States Air Force. 1970. Available online: <https://www.amazon.com/Order-statistics-testing-estimation-Volumes/dp/B0006C3I0Q> (accessed on 13 December 2023).
13. Krishna, H.; Malik, M. Reliability estimation in Maxwell distribution with progressively Type-II censored data. *J. Stat. Comput. Simul.* **2012**, *82*, 623–641. [CrossRef]
14. Krishna, H.; Vivekanand; Kumar, K. Estimation in Maxwell distribution with randomly censored data. *J. Stat. Comput. Simul.* **2015**, *85*, 3560–3578. [CrossRef]
15. Tomer, S.K.; Panwar, M. Estimation procedures for Maxwell distribution under Type-I progressive hybrid censoring scheme. *J. Stat. Comput. Simul.* **2015**, *85*, 339–356. [CrossRef]
16. Panwar, M.; Tomer, S.K. Robust Bayesian Analysis of Lifetime Data from Maxwell Distribution. *Austrian J. Stat.* **2019**, *48*, 38–55. [CrossRef]
17. Kumari, A.; Kumar, K.; Kumar, I. Bayesian and classical inference in Maxwell distribution under adaptive progressively Type-II censored data. *Int. J. Syst. Assur. Eng. Manag.* **2023**, 1–22. [CrossRef]
18. Stacy, E.W. A Generalization of the Gamma Distribution. *Ann. Math. Stat.* **1962**, *33*, 1187–1192. [CrossRef]
19. Johnson, N.L.; Kotz, S.; Balakrishnan, N. *Continuous Univariate Distributions*, 2nd ed.; John Wiley & Sons: New York, NY, USA, 1994; Volume 1.
20. Lai, C.D. *Generalized Weibull Distributions*; Springer: New York, NY, USA, 1994.
21. Balasooriya, U. Failure-censored reliability sampling plans for the exponential distribution. *J. Stat. Comput. Simul.* **1995**, *52*, 337–349. [CrossRef]
22. Cohen, A.C. Progressively Censored Samples in Life Testing. *Technometrics* **1963**, *5*, 327–339. [CrossRef]
23. Wu, S.J.; Kuş, C. On estimation based on progressive first-failure-censored sampling. *Comput. Stat. Data Anal.* **2009**, *53*, 3659–3670. [CrossRef]

24. Dube, M.; Garg, R.; Krishna, H. On progressively first failure censored Lindley distribution. *Comput. Stat.* **2016**, *31*, 139–163. [[CrossRef](#)]
25. Mohammed, H.S.; Ateya, S.F.; AL-Hussaini, E.K. Estimation based on progressive first-failure censoring from exponentiated exponential distribution. *J. Appl. Stat.* **2017**, *44*, 1479–1494. [[CrossRef](#)]
26. Krishna, H.; Dube, M.; Garg, R. Estimation of  $P(Y < X)$  for progressively first-failure-censored generalized inverted exponential distribution. *J. Stat. Comput. Simul.* **2017**, *87*, 2274–2289.
27. Kayal, T.; Tripathi, Y.M.; Wang, L. Inference for the Chen Distribution Under Progressive First-Failure Censoring. *J. Stat. Theory Pract.* **2019**, *13*, 52. [[CrossRef](#)]
28. Saini, S.; Chaturvedi, A.; Garg, R. Estimation of stress–strength reliability for generalized Maxwell failure distribution under progressive first failure censoring. *J. Stat. Comput. Simul.* **2021**, *91*, 1366–1393. [[CrossRef](#)]
29. Kumar, I.; Kumar, K.; Ghosh, I. Reliability estimation in inverse Pareto distribution using progressively first failure censored data. *Am. J. Math. Manag. Sci.* **2023**, *42*, 126–147. [[CrossRef](#)]
30. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Statist. Soc. Ser. B* **1977**, *39*, 1–38. [[CrossRef](#)]
31. McLachlan, G.J.; Krishnan, T. *The EM Algorithm and Extensions*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2008.
32. Casella, G.; Berger, R.L. *Statistical Inference*, 2nd ed.; Duxbury Press: Pacific Grove, CA, USA, 2002.
33. Tanner, M.A. *Tools for Statistical Inference*, 3rd ed.; Springer: New York, NY, USA, 1996.
34. Louis, T.A. Finding the observed information matrix when using the EM algorithm. *J. R. Stat. Soc. Ser. B (Methodol.)* **1982**, *44*, 226–233. [[CrossRef](#)]
35. Krishnamoorthy, K.; Lin, Y. Confidence limits for stress—Strength reliability involving Weibull models. *J. Stat. Plan. Inference* **2010**, *140*, 1754–1764. [[CrossRef](#)]
36. Efron, B. Bootstrap Methods: Another Look at the Jackknife. *Ann. Stat.* **1979**, *7*, 1–26. [[CrossRef](#)]
37. Efron, B.; Tibshirani, R.J. An Introduction to The Bootstrap. *Monogr. Stat. Appl. Probab.* **1993**, *57*, 1–436.
38. Efron, B. *The Jackknife, the Bootstrap and Other Resampling Plans*; SIAM: Philadelphia, PA, USA, 1982.
39. Hall, P. Theoretical comparison of bootstrap confidence intervals. *Ann. Stat.* **1988**, *16*, 927–953. [[CrossRef](#)]
40. Box, G.E.P.; Tiao, G.C. *Bayesian Inference Statistical Analysis*; John Wiley & Sons: New York, NY, USA, 1992.
41. Gelman, A.; Carlin, J.B.; Stern, H.S.; Dunson, D.B.; Vehtari, A.; Rubin, D.B. *Bayesian Data Analysis*, 3rd ed.; CRC Press: Boca Raton, FL, USA, 2014.
42. Tian, Q.; Lewis-Beck, C.; Niemi, J.B.; Meeker, W.Q. Specifying prior distributions in reliability applications. In *Applied Stochastic Models in Business and Industry*; 2023. Available online: <https://onlinelibrary.wiley.com/doi/10.1002/asmb.2752> (accessed on 13 December 2023).
43. Chaudhary, S.; Kumar, J.; Tomer, S.K. Estimation of  $P[Y < X]$  for Maxwell distribution. *J. Stat. Manag. Syst.* **2017**, *20*, 467–481.
44. Varian, H.R. A Bayesian approach to real estate assessment. *Stud. Bayesian Econom. Stat. Honor. Leonard J. Savage* **1975**, *4*, 195–208.
45. Tierney, L.; Kadane, J.B. Accurate approximations for posterior moments and marginal densities. *J. Am. Stat. Assoc.* **1986**, *81*, 82–86. [[CrossRef](#)]
46. Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087–1091. [[CrossRef](#)]
47. Hastings, W.K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **1970**, *57*, 97–109. [[CrossRef](#)]
48. Chen, M.H.; Shao, Q.M. Monte Carlo estimation of Bayesian credible and HPD intervals. *J. Comput. Graph. Stat.* **1999**, *8*, 69–92.
49. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2023.
50. Nichols, M.D.; Padgett, W. A bootstrap control chart for Weibull percentiles. *Qual. Reliab. Eng. Int.* **2006**, *22*, 141–151. [[CrossRef](#)]
51. Xie, Y.; Gui, W. Statistical Inference of the Lifetime Performance Index with the Log-Logistic Distribution Based on Progressive First-Failure-Censored Data. *Symmetry* **2020**, *12*, 937. [[CrossRef](#)]
52. Mudholkar, G.S.; Srivastava, D.K.; Kollia, G.D. A generalization of the Weibull distribution with application to the analysis of survival data. *J. Am. Stat. Assoc.* **1996**, *91*, 1575–1583. [[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.