axioms

MDPI

Article

Estimation of Entropy for Generalized Rayleigh Distribution under Progressively Type-II Censored Samples

Haiping Ren [1,*], Qin Gong [2] and Xue Hu [2]

[1] Teaching Department of Basic Subjects, Jiangxi University of Science and Technology, Nanchang 330013, China
[2] College of Science, Jiangxi University of Science and Technology, Ganzhou 341000, China; gongqin0804@163.com (Q.G.); 6120210110@mail.jxust.edu.cn (X.H.)
* Correspondence: 9520060004@jxust.edu.cn; Tel.: +86-15979014223

Abstract: This paper investigates the problem of entropy estimation for the generalized Rayleigh distribution under progressively type-II censored samples. Based on progressively type-II censored samples, we first discuss the maximum likelihood estimation and interval estimation of Shannon entropy for the generalized Rayleigh distribution. Then, we explore the Bayesian estimation problem of entropy under three types of loss functions: K-loss function, weighted squared error loss function, and precautionary loss function. Due to the complexity of Bayesian estimation computation, we use the Lindley approximation and MCMC method for calculating Bayesian estimates. Finally, using a Monte Carlo statistical simulation, we compare the mean square errors to examine the superiority of maximum likelihood estimation and Bayesian estimation under different loss functions. An actual example is provided to verify the feasibility and practicality of various estimations.

Keywords: Bayes estimation; Shannon entropy; loss function; generalized Rayleigh distribution

MSC: 62F10; 62F15

check for updates

Citation: Ren, H.; Gong, Q.; Hu, X. Estimation of Entropy for Generalized Rayleigh Distribution under Progressively Type-II Censored Samples. *Axioms* **2023**, *12*, 776. https://doi.org/10.3390/axioms12080776

Academic Editor: Jong-Min Kim

Received: 11 July 2023
Revised: 6 August 2023
Accepted: 9 August 2023
Published: 10 August 2023

Copyright: © 2023 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

The Rayleigh distribution has many practical applications in reliability and life testing, and typically used in physical modeling processes in different fields such as wave heights in the ocean, sound and light radiation, wireless signals and wind energy, and ultrasound imaging modeling [1]. It is also used for modeling resistors, networks, crystals, relays, and capacitors in aircraft radar assemblies with lifetimes of only a few hours. In all of these applications, it is generally assumed that the lifetime of a particular device depends on its age. Because generalized models based on lifetime distributions have more intuitive applicability and attractiveness in the modeling process, Surles and Padgett [2] proposed the two-parameter Burr Type X distribution, also known as the generalized Rayleigh distribution. They studied the statistical characteristics and some properties of the generalized Rayleigh distribution, and pointed out that it can be used as a lifetime distribution and to characterize intensity data. Surles et al. [3] and Al-khedhari et al. [4] discussed parameter estimation for this distribution using different techniques.

The generalized Rayleigh distribution is essentially a two-parameter Burr Type X distribution [1], which was first introduced by Surles and Padgett in 2001 [2]. Burr Type X and Burr Type XII distributions are important distributions in survival analysis problems. They were two of the twelve types of lifetime distributions proposed by Burr in 1942, and since their proposal, their applications and statistical inference theory have received widespread attention and research [3,4]. In recent years, with the deepening of research, the generalized Rayleigh distribution has been widely applied in fields such as wireless communication, radar signal processing, image processing, seismology, astronomy, and ecology [5–7]. It not only meets the needs for different types of data, such as the distribution of radio signal

strength, power load, and seismic wave energy, but also provides more accurate data analysis methods and tools to help us better understand and interpret data in various practical problems. In addition, the generalized Rayleigh distribution is an extension of the Rayleigh distribution. Compared with the Rayleigh distribution, the generalized Rayleigh distribution has more superior performance. It can better fit the data in some practical problems by introducing additional parameters and variables. The Rayleigh distribution is usually used to describe amplitude distribution, but in some cases, different types of data distribution such as amplitude, power, and energy may require more complex models. Therefore, the generalized Rayleigh distribution has played a role in improving existing models and providing more accurate and comprehensive descriptions and analysis [8]. Based on ranked set sampling (RSS), Shen et al. [9] compared the performance of RSS with respect to simple random sampling by observing modified unbiased and modified best linear unbiased estimators from generalized Rayleigh distribution. They found that the prementioned estimators based on RSS were significantly more efficient than the ones in SRS. Shen et al. [10] applied the generalized Rayleigh distribution to real-life scenarios, using its heavy-tailed property for Reddit ads and breast cancer datasets, and compared it with other generalized Rayleigh distributions. Rabie et al. [11] studied the lifetime analysis of devices following the generalized Rayleigh distribution under a type-II mixed censored scheme, obtaining Bayes estimates and E-Bayes estimates using a squared error loss function and linear loss function. It can be seen that the generalized Rayleigh distribution not only helps to expand the scope of application of the Rayleigh distribution and improve the accuracy and reliability of data analysis, but also provides theoretical support for the modeling and prediction of practical problems. Therefore, it is meaningful to study the generalized Rayleigh distribution in this paper.

Let X be a random variable that follows a generalized Rayleigh distribution with parameters β and σ, denoted by $GRD(\sigma, \beta)$. The probability density function (PDF) and cumulative distribution function are, respectively, as follows:

$$f(x|\sigma,\beta) = 2\sigma\beta^2 x e^{-\beta^2 x^2}\left(1-e^{-\beta^2 x^2}\right)^{\sigma-1},\ x>0,\ \sigma,\ \beta>0, \tag{1}$$

$$F(x|\sigma,\beta) = \left(1-e^{-\beta^2 x^2}\right)^{\sigma},\ x>0,\ \sigma,\ \beta>0. \tag{2}$$

Here, σ and β are the shape and scale parameters of the $GRD(\sigma, \beta)$.

The survival function and hazard function (HF) of $GRD(\sigma, \beta)$ are, respectively, as follows:

$$r(x|\sigma,\beta) = 1-\left(1-e^{-\beta^2 x^2}\right)^{\sigma},\ x>0,\ \sigma,\ \beta>0,$$

$$h(x|\sigma,\beta) = \frac{2\sigma\beta^2 x e^{-\beta^2 x^2}\left(1-e^{-\beta^2 x^2}\right)^{\sigma-1}}{1-\left(1-e^{-\beta^2 x^2}\right)^{\sigma}},\ x>0,\ \sigma,\ \beta>0.$$

The PDF and HF of the $GRD(\sigma, \beta)$ are closely related to the parameters σ and β. Raqab and Kundu [12] pointed out that when $\sigma = 1$, the distribution is a standard Rayleigh distribution. When $\sigma \neq 1$, σ serves as the shape parameter of the distribution. When $\sigma > 1/2$, the PDF is a decreasing function, and the graph of HF exhibits a bathtub shape. When $\sigma \leq 1/2$, the PDF is a right-skewed unimodal function, and the graph of HF shows an increasing shape. The image of the PDF and HF of the $GRD(\sigma, \beta)$ is shown in Figure 1.

This paper considers the estimation problem of the $GRD(\sigma, \beta)$. In Section 2, the Shannon entropy of $GRD(\sigma, \beta)$ is firstly derived and the maximum likelihood(ML) estimation and asymptotic confidence interval (ACI) estimation of Shannon entropy are discussed. Section 3 performs the Bayes estimators (BEs) under the weighted squared error loss function (WESLF), precautionary loss function (PLF) and K-loss function (KLF), and numerically calculates these estimators using Lindley's approximation and the MCMC method. Section 4 observes the performance of each estimation method using a Monte

Carlo simulation. In Section 5, the superiority of the estimation methods used is compared using real data. Finally, conclusions are given in Section 6.

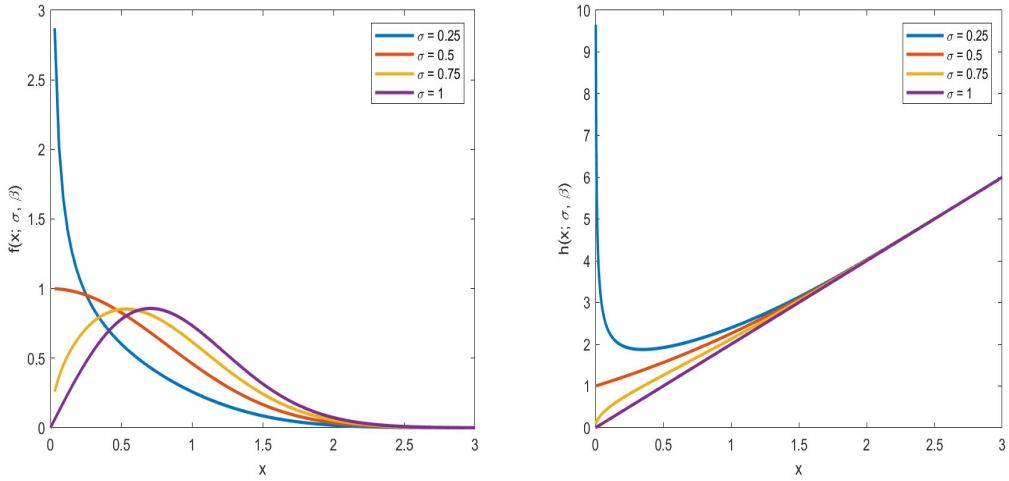


Figure 1. Curves of PDF and HF of the $GRD(\sigma, \beta)$.

2. Preliminary Knowledge

2.1. Shannon Entropy of $GRD(\sigma, \beta)$

Shannon entropy is an important concept in information theory used to measure the uncertainty or randomness of information. It was introduced by Shannon in 1948 and is considered one of the foundations of information theory. Shannon entropy is defined as follows:

$$H(f) = -\int_0^\infty f(x) \ln f(x) dx \tag{3}$$

where $f(x)$ is the PDF of a continuous random variable X.

In the existing literature, there have been significant advancements in the estimation of entropy functions for different distributions. Xu and Gui [13] discussed the estimation problem of entropy for the two-parameter inverse Weibull distribution under an adaptive Type-II progressive hybrid censoring scheme. Chacko and Asha [14] discussed the entropy estimation problem for the generalized exponential distribution based on record values. Shrahili et al. [15] discussed the estimation of logarithmic logistic distribution entropy under asymptotic testing. Wang and Gui [16] used the ML estimation and Bayesian estimation methods to obtain an estimate of the entropy of the two-parameter Burr type distribution under progressively censoring data. Al-Babtain et al. [17] analyzed the estimation of entropy for the Kumaraswamy distribution using ML estimation. Bantan et al. [18] discussed the estimation of entropy for the inverse Lomax distribution under multiple censoring data. Shi et al. [19] discussed the estimation of entropy for the generalized Bilal distribution under adaptive Type-II hybrid censored data and obtained ML estimates of entropy and parameters using Newton's iteration method. Liu and Gui [20] discussed the estimation problem of entropy for the two-parameter Lomax distribution under generalized asymptotic hybrid censored data. They estimated entropy by deriving the ML estimation of unknown parameters and analyzed BEs under different loss functions, obtaining Bayesian estimates using the Lindley approximation method and MCMC method.

Theorem 1. *Let X be a random variable whose PDF is $GRD(\sigma, \beta)$ in Equation (1). Then, the Shannon entropy of the $GRD(\sigma, \beta)$ is*

$$\begin{aligned} H(f) &= -\ln 2\sigma\beta^2 - \sigma \sum_{i=0}^{\infty} \binom{\sigma - 1}{i} (-1)^i \frac{\frac{1}{2}\Gamma'(1) - \Gamma(1) \ln\left[\beta(1+i)^{\frac{1}{2}}\right]}{1+i} \\ &+ [\varphi(\sigma + 1) - \varphi(1)] - \tfrac{1}{\sigma} + 1. \end{aligned} \tag{4}$$

here, $\varphi(.)$ represents the digamma function, and $\Gamma(.)$ represents the gamma function.

Proof. See Appendix A. □

2.2. Maximum Likelihood Estimation

At present, most of the research on the Rayleigh distribution is based on the analysis of sample data under complete or censored data. However, because there are some irresistible factors in the process of the test, such as time factors, test funds, etc., we chose to use the censored test method for the life test. The commonly used censored life tests include timed (Type-I) censored tests and fixed number (Type-II) censored tests [21]. Due to the fact that the timed and fixed number censored tests terminate the test after reaching the censored time and a certain sample size, and do not remove any products, it is easy to cause certain errors in the conclusions obtained from the test in both cases [3]. Therefore, the gradually censoring test is applicable. Generally, there are progressive Type-I censored (PC-I) tests and progressive Type-II censored (PC-II) tests, both of which are censored test methods that can be terminated in advance, thereby saving test costs and time [22]. This paper discusses the estimation of the entropy of the $GRD(\sigma, \beta)$ under the Type-II censoring test. The advantage of the stepwise Type-II consored test is that it can remove some of the products that have not failed during different specified testing stages to analyze the performance of the products. Using this test method can flexibly adjust the sample size and consored time, more effectively utilize time and resources, and thus improve the efficiency of data collection. From a practical perspective, compared to the traditional Type-II censored scheme, this censored scheme can determine the sample size based on the actual needs and currently available resources. By gradually increasing the sample size, resources are maximized while ensuring product quality, avoiding resource waste while ensuring the accuracy of the obtained data. At the same time, the censored plan can also set the censored time according to the actual situation. By gradually increasing the censored time, it can save time to the maximum extent while ensuring product quality, thereby avoiding time consumption and improving the reliability of data analysis.

Assuming a life test is conducted on n products, each time a failed product is observed, a certain number of products are randomly removed from the remaining non-failed products [23]. That is to say, when the failure time of the first product is observed to be $X_{1:m:n}$, P_1 products are removed from the remaining $n-1$ non-failed products, and $n-1-P_1$ products remain to be observed. When the expiration time of the second product is observed to be $X_{2:m:n}$, P_2 products are removed from the remaining non-expired products, leaving $n-2-P_1-P_2$ products; continuing this way, when the failure time of the m-th product is observed to be $X_{m:m:n}$, the experiment is stopped and P_m products are removed, $P_m = n-m-P_1-P_2-\ldots-P_{m-1}$.

Let $(X_{1:m:n}, X_{2:m:n}, \ldots, X_{m:m:n})$ be the m PC-II samples observed from n total test samples, and each $X_{i:m:n}$ $(0 < i \leq m)$ follows the $GRD(\sigma, \beta)$ defined by Equation (1). According to the literature [24], based on the PC-II sample, the likelihood function (LF) can be defined as follows:

$$L(\sigma, \beta | x^*) = q \prod_{i=1}^{m} f(x_{i:m:n}) [1 - F(x_{i:m:n})]^{P_i}, \tag{5}$$

where $q = n(n-P_1-1)(n-P_1-P_2)\ldots\left(n-\sum_{i=1}^{m-1}(P_i+1)\right)$, and $x_{i:m:n}$ represents the observation values of the $X_{i:m:n}$ sample. For convenience, we briefly write $x_{i:m:n}$ as x_i, and $x^* = (x_1, x_2, \ldots, x_m)$. Substituting Equations (1) and (2) into Equation (5), we can obtain

$$L(\sigma, \beta | x^*) = 2^m \sigma^m \beta^{2m} q \prod_{i=1}^{m} \left(x_i (1-u_i) {u_i}^{\sigma-1} (1-{u_i}^{\sigma})^{P_i} \right), \tag{6}$$

where $u(\beta, x) = 1 - e^{-\beta^2 x^2}$.

The corresponding log-LF is

$$\begin{aligned} l(\sigma,\beta) &= \ln L(\sigma,\beta|x^*) \\ &= m\ln 2 + m\ln\sigma + 2m\ln\beta + \ln q + \sum_{i=1}^{m}\ln x_i + \sum_{i=1}^{m}\ln(1-u_i) \\ &+(\sigma-1)\sum_{i=1}^{m}\ln u_i + \sum_{i=1}^{m} P_i \ln(1-{u_i}^{\sigma}). \end{aligned} \tag{7}$$

The likelihood equations are as follows:

$$\frac{\partial l(\sigma,\beta)}{\partial\sigma} = \frac{m}{\sigma} + \sum_{i=1}^{m}\ln u_i - \sum_{i=1}^{m} P_i \frac{{u_i}^{\sigma}\ln u_i}{1-{u_i}^{\sigma}} = 0, \tag{8}$$

$$\frac{\partial l(\sigma,\beta)}{\partial\beta} = \frac{2m}{\beta} - 2\beta\sum_{i=1}^{m}{x_i}^2 + 2\beta\sum_{i=1}^{m}\left[\frac{{x_i}^2(1-u_i)(\sigma-1)}{u_i} - \frac{\sigma P_i {x_i}^2(1-u_i){u_i}^{\sigma-1}}{1-{u_i}^{\sigma}}\right] = 0. \tag{9}$$

By solving Equations (8) and (9), σ and β can be obtained. It should be noted that the ML estimators of these equations cannot be analytically determined, so we need to use numerical calculation methods to calculate σ and β. Now, we introduce the following algorithm to solve the ML estimators of σ and β, the steps of which are given as follows:

(i) Determine the LF and take the logarithm to obtain the log-LF.
(ii) Calculate the first derivative and the second derivative of the log-LF.
(iii) Initialize parameter values, usually using sample mean or other empirical values as initial values.
(iv) Calculate the first and second derivatives using the current parameter values.
(v) Update the parameter values using Newton's iterative formula:

$$\hat{\theta} = \theta - \frac{f'(\theta)}{f''(\theta)},$$

where $\hat{\theta}$ represents the new parameter values and θ represents the initial parameter values.

(vi) Repeat steps (iv) and (v) until the parameter values converge or reach the preset number of iterations.
(vii) The final parameter values obtained are the ML estimates.

Due to the invariance of ML estimation, the parameter estimates obtained through iteration are not significantly different. Therefore, we can use the estimated value $\hat{\sigma},\hat{\beta}$ instead of σ,β in Equations (8) and (9). Then, we can obtain the ML estimators of Shannon entropy, which are expressed as

$$\begin{aligned} \hat{H}(f) &= -\ln 2\hat{\sigma}\hat{\beta}^2 - \hat{\sigma}\sum_{i=0}^{\infty}\binom{\hat{\sigma}-1}{i}(-1)^i \frac{\frac{1}{2}\Gamma'(1)-\Gamma(1)\ln\left[\hat{\beta}(1+i)^{\frac{1}{2}}\right]}{1+i} \\ &+[\varphi(\hat{\sigma}+1)-\varphi(1)] - \tfrac{1}{\hat{\sigma}} + 1. \end{aligned} \tag{10}$$

In addition to ML estimation, which is a non-Bayesian method to estimate parameters, there is also a non-Bayesian generalized reasoning method to estimate parameters. This method can perform efficient calculations on large-scale data, especially for big data analysis and machine learning tasks. See [25] for details.

2.3. Asymptotic Confidence Interval of Shannon Entropy

To obtain the ACI of Shannon entropy, we can use the Delta method [26] to derive. According to the Delta method, we need to centralize and standardize the estimators of Shannon entropy to obtain a new random variable. Next is the step of finding the ACI of information entropy:

Firstly, the Fisher information matrix is

$$I(\hat{\sigma},\hat{\beta}) = \begin{pmatrix} -\frac{\partial^2 l(\sigma,\beta)}{\partial\sigma^2} & -\frac{\partial^2 l(\sigma,\beta)}{\partial\sigma\partial\beta} \\ -\frac{\partial^2 l(\sigma,\beta)}{\partial\sigma\partial\beta} & -\frac{\partial^2 l(\sigma,\beta)}{\partial\beta^2} \end{pmatrix}_{\sigma=\hat{\sigma},\beta=\hat{\beta}}, \tag{11}$$

where

$$\frac{\partial^2 l(\sigma,\beta)}{\partial\sigma^2} = -\frac{m}{\sigma^2} - \sum_{i=1}^{m} P_i \frac{{u_i}^{\sigma}(\ln u_i)^2}{(1-{u_i}^{\sigma})^2},$$

$$\frac{\partial^2 l(\sigma,\beta)}{\partial\sigma\partial\beta} = 2\beta\sum_{i=1}^{m}\frac{{x_i}^2(1-u_i)}{u_i} - 2\beta\sum_{i=1}^{m} P_i \frac{{x_i}^2 {u_i}^{\sigma-1}(1-u_i)(\sigma\ln u_i - {u_i}^{\sigma}+1)}{(1-{u_i}^{\sigma})^2},$$

$$\begin{aligned}\frac{\partial^2 l(\sigma,\beta)}{\partial\beta^2} &= -\frac{2m}{\beta^2} - 2\sum_{i=1}^{m}{x_i}^2 + 2(\sigma-1)\sum_{i=1}^{m}\frac{{x_i}^2(1-u_i)}{{u_i}^2} - 4\beta^2(\sigma-1)\sum_{i=1}^{m}\frac{{x_i}^4(1-u_i)}{{u_i}^2} \\ &\quad + 4\beta^2\sigma\sum_{i=1}^{m} P_i \frac{{x_i}^4(1-u_i){u_i}^{\sigma-2}\left(\sigma u_i - u_i^{\sigma} - \sigma + 1\right)}{\left(1-u_i^{\sigma}\right)^2} \\ &\quad - 2\sigma\sum_{i=1}^{m} P_i \frac{{x_i}^2(1-u_i){u_i}^{\sigma-1}}{1-u_i^{\sigma}},\end{aligned}$$

$$\frac{\partial^2 l(\sigma,\beta)}{\partial\beta\partial\sigma} = 2\beta\sum_{i=1}^{m}\frac{{x_i}^2(1-u_i)}{u_i} - 2\beta\sum_{i=1}^{m} P_i \frac{{x_i}^2 {u_i}^{\sigma-1}(1-u_i)(\sigma\ln u_i - {u_i}^{\sigma}+1)}{(1-{u_i}^{\sigma})^2}.$$

Thus, the inverse $I^{-1}(\sigma,\beta)$ of Fisher's information moment $I(\sigma,\beta)$ is obtained.

Let $K = \left(\frac{\partial H(X)}{\partial\sigma}, \frac{\partial H(X)}{\partial\beta}\right)$. Due to the presence of combinatorial numbers in Equation (4), it is very complex to derive. Therefore, for convenience, we use the following equivalent form of Equation (4):

$$H(f) = -\ln 2\sigma\beta^2 - 2\sigma\beta^2\int_0^{\infty}(1-u)u^{\sigma-1}x\ln x dx + [\varphi(\sigma+1) - \varphi(1)] - \frac{1}{\sigma} + 1,$$

where $u = 1 - e^{-\beta^2 x^2}$. Hence, we can obtain

$$\frac{\partial H(f)}{\partial\sigma} = -\frac{1}{\sigma} - 2\beta^2\int_0^{\infty}(1-u)u^{\sigma-1}x\ln x dx - 2\sigma\beta^2\int_0^{\infty}(1-u)u^{\sigma-1}x\ln u\ln x dx + \varphi\prime(\sigma+1) + \frac{1}{\sigma^2},$$

$$\frac{\partial H(f)}{\partial\beta} = 4\sigma\beta^3\int_0^{\infty}(\sigma u - \sigma + 1)(1-u)u^{\sigma-2}x^3\ln x dx - \frac{2}{\beta} - 4\sigma\beta\int_0^{\infty}(1-u)u^{\sigma-1}x\ln x dx.$$

Then, the variance of entropy is

$$Var\left(\hat{H}(f)\right) = K_s I^{-1}(\sigma,\beta)K_s^T\Big|_{\beta=\hat{\beta},\sigma=\hat{\sigma}},$$

where K^T represents the transposition of K. Thus, the $100(1-\alpha)\%$ ACI of Shannon entropy is

$$\hat{H}(f) - T_{\alpha/2}\sqrt{Var\left(\hat{H}(f)\right)}, \hat{H}(f) + T_{\alpha/2}\sqrt{Var\left(\hat{H}(f)\right)}, \tag{12}$$

where $T_{\alpha/2}$ is the upper quantile of the standard normal distribution.

3. Bayes Estimation

Bayesian estimation is a statistical method used to estimate unknown parameters given a prior probability distribution and observation data. It is derived from Bayes' theorem. The development of Bayesian estimation can be traced back to the 18th century, but due to insufficient computer technology and data collection capabilities during that era, Bayesian estimation was not widely used. With the continuous development of computer technology, Bayesian estimation began to receive widespread attention from scholars in

the late 20th century and developed rapidly. Nowadays, Bayesian estimation is widely applied in fields such as machine learning, statistics, and data mining. Variational Bayes is a type of technique used for approximating complex integrals in the fields of Bayesian estimation and machine learning. Currently, Bayesian inference methods such as variational Bayes are also widely used, as detailed in [27,28]. Due to the excellent performance of Bayesian estimation, a large number of scholars currently use Bayesian estimation methods to estimate parameters and correlation functions.

In Bayesian estimation, we need to introduce a loss function to measure the risks brought by different decisions [29,30]. This paper studies the Bayesian estimation of Shannon entropy under the WSELF, PLF, and KLF.

3.1. Prior and Posterior Distributions

In this paper, we only consider an informative prior. The principle of selecting an informative prior is to estimate parameters through previous knowledge or experience combined with observation data. This previous knowledge or experience can be derived from the truth summarized by previous explorations or related historical data. An informative prior can help us improve the shortcomings of the data and obtain more accurate parameter estimates. When addressing considerations related to previous introductions, sensitivity analysis, or robustness, we need to consider selecting informative priors that are obtained while ensuring the reliability and accuracy of prior knowledge, which can be obtained through reading a large amount of research from the literature or conducting field visits and investigations; secondly, different prior distributions should be selected based on the essence of the problem, and prior knowledge should be applied to them; after selecting an informative prior, prior sensitivity analysis can be conducted to evaluate the impact of the prior on the results, and adjustments can be made appropriately based on the differences in results to optimize prior selection; and finally, a prior robustness analysis is also needed to evaluate some outlier in the data and make appropriate adjustments according to the difference in results to optimize prior selection so as to improve the accuracy and reliability of the analysis data. Assume that σ and β are independent and follow the gamma distribution $\Gamma(a_1, b_1)$ and $\Gamma(a_2, b_2)$, respectively; that is, the PDFs of σ and β are

$$\pi_1(\sigma|a_1, b_1) = \frac{{b_1}^{a_1}}{\Gamma(a_1)}\sigma^{a_1-1}e^{-b_1\sigma};\ a_1 > 0,\ b_1 > 0,$$

$$\pi_2(\beta|a_2, b_2) = \frac{{b_2}^{a_2}}{\Gamma(a_2)}\beta^{a_2-1}e^{-b_2\beta};\ a_2 > 0,\ b_2 > 0.$$

Then, the joint prior distribution of σ and β is

$$\pi(\sigma, \beta) = \frac{{b_1}^{a_1}{b_2}^{a_2}}{\Gamma(a_1)\Gamma(a_2)}\sigma^{a_1-1}\beta^{a_2-1}e^{-b_1\sigma-b_2\beta}.$$

According to the Bayesian theorem, the joint posterior density of σ and β is:

$$\pi(\sigma, \beta \mid x^*) = \frac{L(\sigma, \beta \mid x^*)\pi(\sigma, \beta)}{\int_0^\infty \int_0^\infty L(\sigma, \beta \mid x^*)\pi(\sigma, \beta)d\sigma d\beta}. \tag{13}$$

Next, we analyze the BEs of entropy under the following three loss functions.

(i) WSELF [31]:

$$L_W\big(H(f), \hat{H}(f)\big) = \frac{\big(\hat{H}(f) - H(f)\big)^2}{H(f)}.$$

Then, the BE under the WSELF is

$$\hat{H}(f)_{L_{WB}} = \Big[E\Big(H^{-1}(f)|x\Big)\Big]^{-1} = \left(\int_0^\infty \int_0^\infty H^{-1}(f)\pi(\sigma, \beta|x^*)d\sigma d\beta\right)^{-1}. \tag{14}$$

(ii) PLF [32]:

$$L_P\big(H(f),\hat{H}(f)\big)=\frac{\big(\hat{H}(f)-H(f)\big)^2}{\hat{H}(f)}.$$

Then, the BE under the PLF is

$$\hat{H}(f)_{L_{PB}}=\sqrt{E(H^2(f)|x)}=\sqrt{\int_0^\infty\int_0^\infty H^2(f)\pi(\sigma,\beta|x^*)d\sigma d\beta}. \tag{15}$$

(iii) KLF [33]:

$$L_K\big(H(f),\hat{H}(f)\big)=\left(\sqrt{\frac{\hat{H}(f)}{H(f)}}-\sqrt{\frac{H(f)}{\hat{H}(f)}}\right)^2.$$

Then, the BE under KLF is

$$\hat{H}(f)_{L_{KB}}=\sqrt{\frac{E(H(f)|x)}{E(H^{-1}(f)|x)}}=\left(\frac{\int_0^\infty\int_0^\infty H(f)\pi(\sigma,\beta|x^*)d\sigma d\beta}{\int_0^\infty\int_0^\infty H^{-1}(f)\pi(\sigma,\beta|x^*)d\sigma d\beta}\right)^{-\frac{1}{2}}. \tag{16}$$

where L_W, L_P, L_K represent the WSELF, PLF and KLF, respectively; $\hat{H}(f)_{L_{WB}}$, $\hat{H}(f)_{L_{PB}}$, $\hat{H}(f)_{L_{KB}}$ represent the BEs of entropy for the WSELF, PLF, and KLF, respectively; $E(H(f)|x)$ represents the posterior expectation of $H(f)$; $E\big(H^{-1}(f)|x\big)$ represents the posterior expectation of H^{-1}; and $E\big(H^2(f)|x\big)$ represents the posterior expectation of H^2. In order to calculate the Bayesian estimates in Equations (14)–(16), we need to obtain $E(H(f)|x)$, $E\big(H^{-1}(f)|x\big)$, and $E\big(H^2(f)|x\big)$. Since the BEs under the above three loss functions are in the form of two integral ratios, which are more complex to calculate, we choose Lindley's approximation to obtain the Bayesian estimates of entropy under the loss function.

3.2. Lindley's Approximation

Referring to Lindley's approximation [34], we provide the definition of $I(x^*)$:

$$I(x^*)=E[U(\sigma,\beta)|x^*]=\frac{\int_0^\infty\int_0^\infty U(\sigma,\beta)e^{P(\sigma,\beta)+l(\sigma,\beta|x^*)}d\sigma d\beta}{\int_0^\infty\int_0^\infty e^{P(\sigma,\beta)+l(\sigma,\beta|x^*)}d\sigma d\beta}. \tag{17}$$

Among them, $U(\sigma,\beta)$ is the function of σ and β; $P(\sigma,\beta)$ is the logarithmic joint prior distribution of σ and β, i.e., $P(\sigma,\beta)=\ln\pi(\sigma,\beta)$; and $l(\sigma,\beta|x^*)$ is the log-LF in Equation (7). If the sample is large enough, Equation (17) can be further reduced to

$$\begin{aligned}I(x^*) &= U(\hat{\sigma},\hat{\beta})+\tfrac{1}{2}[(\hat{U}_{\sigma\sigma}+2\hat{U}_\sigma\hat{P}_\sigma)\hat{\tau}_{\sigma\sigma}+(\hat{U}_{\beta\sigma}+2\hat{U}_\beta\hat{P}_\sigma)\hat{\tau}_{\beta\sigma}\\ &+(\hat{U}_{\sigma\beta}+2\hat{U}_\sigma\hat{P}_\beta)\hat{\tau}_{\sigma\beta}+(\hat{U}_{\beta\beta}+2\hat{U}_\beta\hat{P}_\beta)\hat{\tau}_{\beta\beta}]\\ &+\tfrac{1}{2}[(\hat{U}_\sigma\hat{\tau}_{\sigma\sigma}+\hat{U}_\beta\hat{\tau}_{\sigma\beta})(\hat{l}_{\sigma\sigma\sigma}\hat{\tau}_{\sigma\sigma}+\hat{l}_{\sigma\beta\sigma}\hat{\tau}_{\sigma\beta}+\hat{l}_{\beta\sigma\sigma}\hat{\tau}_{\beta\sigma}+\hat{l}_{\beta\beta\sigma}\hat{\tau}_{\beta\beta})\\ &+(\hat{U}_\sigma\hat{\tau}_{\beta\sigma}+\hat{U}_\beta\hat{\tau}_{\beta\beta})(\hat{l}_{\beta\sigma\sigma}\hat{\tau}_{\sigma\sigma}+\hat{l}_{\sigma\beta\beta}\hat{\tau}_{\sigma\beta}+\hat{l}_{\beta\sigma\beta}\hat{\tau}_{\beta\sigma}+\hat{l}_{\beta\beta\beta}\hat{\tau}_{\beta\beta})].\end{aligned}$$

here, $\hat{\sigma}$ and $\hat{\beta}$ are the ML estimates of σ and β. $U_{\sigma\sigma}$ is the second derivative of σ in $U(\sigma,\beta)$, U_σ is the first derivative of σ in $U(\sigma,\beta)$, P_σ is the first derivative of σ in $P(\sigma,\beta)$, $\hat{l}_{\sigma\sigma\sigma}$ is the third derivative of σ in $l(\sigma,\beta|x^*)$, and so on.

Furthermore, we have

$$\hat{P}_\sigma=\frac{m-1}{\hat{\sigma}}-n,$$

$$\hat{l}_{\sigma\sigma}=\frac{\partial^2 l(\sigma,\beta)}{\partial\sigma^2}|_{\sigma=\hat{\sigma}}=-\frac{m}{\hat{\sigma}^2}-\sum_{i=1}^m P_i\frac{u_i^{\hat{\sigma}}(\ln u_i)^2}{\big(1-u_i^{\hat{\sigma}}\big)^2},$$

$$\hat{l}_{\beta\beta} = \frac{\partial^2 l(\sigma,\beta)}{\partial\beta^2}\Big|_{\beta=\hat{\beta},\sigma=\hat{\sigma}} = -\frac{2m}{\hat{\beta}^2} - 2\sum_{i=1}^{m} x_i^2 + 2(\hat{\sigma}-1)\sum_{i=1}^{m}\frac{x_i^2}{u_i}$$
$$-4\hat{\beta}^2(\hat{\sigma}-1)\sum_{i=1}^{m}\frac{x_i^4(1-u_i)}{u_i^2} + 4\hat{\beta}^2\hat{\sigma}\sum_{i=1}^{m} P_i\frac{x_i^4(1-u_i)u_i^{\hat{\sigma}-2}\left(\hat{\sigma}u_i-\hat{\sigma}-u_i^{\hat{\sigma}}+1\right)}{\left(u_i^{\hat{\sigma}}-1\right)^2},$$

$$\hat{l}_{\sigma\beta} = \hat{l}_{\beta\sigma} = \frac{\partial^2 l(\sigma,\beta)}{\partial\sigma\partial\beta}\Big|_{\beta=\hat{\beta},\sigma=\hat{\sigma}} = 2\hat{\beta}\sum_{i=1}^{m}\frac{x_i^2(1-u_i)}{u_i} - 2\hat{\beta}\sum_{i=1}^{m} P_i\frac{x_i^2u_i^{\hat{\sigma}-1}(1-u_i)\left(\hat{\sigma}\ln u_i - u_i^{\hat{\sigma}}+1\right)}{\left(1-u_i^{\hat{\sigma}}\right)^2},$$

$$\hat{l}_{\sigma\sigma\sigma} = \frac{\partial^3 l(\sigma,\beta)}{\partial\sigma^3}|_{\sigma=\hat{\sigma}} = \sum_{i=1}^{m}\frac{2m}{\hat{\sigma}^3} - \sum_{i=1}^{m} P_i\frac{u_i^{\hat{\sigma}}(\ln u_i)^3\left(1+u_i^{\hat{\sigma}}\right)}{\left(1-u_i^{\hat{\sigma}}\right)^3},$$

$$\hat{l}_{\sigma\beta\sigma} = \frac{\partial^3 l(\sigma,\beta)}{\partial\sigma\partial\beta\partial\sigma}\Big|_{\beta=\hat{\beta},\sigma=\hat{\sigma}} = 2\hat{\beta}\sum_{i=1}^{m} P_i\frac{x_i^2(1-u_i)u_i^{\hat{\sigma}}\ln u_i\left(\hat{\sigma}\ln u_i - 2u_i^{\hat{\sigma}} + \hat{\sigma}u_i^{\hat{\sigma}}\ln u_i + 2\right)}{\left(u_i^{\hat{\sigma}}-1\right)^3 u_i},$$

$$\hat{l}_{\beta\beta\sigma} = \frac{\partial l(\sigma,\beta)}{\partial\beta^2\partial\sigma}\Big|_{\beta=\hat{\beta},\sigma=\hat{\sigma}} = 2\sum_{i=1}^{m}\frac{x_i^2(1-u_i)}{u_i} - 4\hat{\beta}^2\sum_{i=1}^{m}\frac{x_i^4(1-u_i)}{u_i^2} - 2\sum_{i=1}^{m} P_i\frac{x_i^2u_i^{\hat{\sigma}-1}(1-u_i)}{1-u_i^{\hat{\sigma}}}$$
$$-2\sum_{i=1}^{m} P_i\frac{x_i^2u_i^{\hat{\sigma}-1}(1-u_i)\ln u_i}{\left(1-u_i^{\hat{\sigma}}\right)^2} + 4\hat{\beta}^2\sum_{i=1}^{m} P_i\frac{x_i^4(1-u_i)u_i^{\hat{\sigma}-2}(\hat{\sigma}u_i-\hat{\sigma}+1)}{\left(u_i^{\hat{\sigma}}-1\right)^2}$$
$$+4\hat{\beta}^2\hat{\sigma}\sum_{i=1}^{m} P_i\frac{x_i^4(1-u_i)\left[\left(2\hat{\sigma}u_i^{\hat{\sigma}}\ln u_i\right)(1-u_i)-\left(4u_i^{\hat{\sigma}}-\hat{\sigma}u_i-2u_i^{2\hat{\sigma}}-1\right)\ln u_i\right]}{\left(u_i^{\hat{\sigma}}-1\right)^3},$$

$$\hat{l}_{\beta\beta\beta} = \frac{\partial(\sigma,\beta)}{\partial\beta^3}\Big|_{\beta=\hat{\beta},\sigma=\hat{\sigma}} = \frac{4m}{\hat{\beta}^3} - 12\hat{\beta}(\hat{\sigma}-1)\sum_{i=1}^{m}\frac{x_i^4(1-u_i)}{u_i^2} + 8\hat{\beta}^3(\hat{\sigma}-1)\sum_{i=1}^{m}\frac{x_i^6(1-u_i)(2-u_i)}{u_i^3}$$
$$+4\hat{\beta}\hat{\sigma}\sum_{i=1}^{m} P_i\frac{x_i^4(1-u_i)u_i^{\hat{\sigma}-2}\left[(\hat{\sigma}u_i-\hat{\sigma}+1)\left(1-u_i^{\hat{\sigma}}\right)-(1-u_i)u_i^{\hat{\sigma}}\right]}{\left(1-u_i^{\hat{\sigma}}\right)^2}$$
$$-16\hat{\beta}^2\hat{\sigma}\sum_{i=1}^{m} P_i\frac{x_i^6(1-u_i)u_i^{\hat{\sigma}-2}\left[\left(\hat{\sigma}u_i-\hat{\sigma}-u_i^{\hat{\sigma}}+1\right)(\hat{\sigma}u_i-\hat{\sigma}-u_i+2)u_i^{-1}-(1-u_i)\left(1-u_i^{\hat{\sigma}-1}\right)\right]}{\left(u_i^{\hat{\sigma}}-1\right)^2}$$
$$+32\hat{\beta}^2\hat{\sigma}^2\sum_{i=1}^{m} P_i\frac{x_i^6(1-u_i)^2u_i^{2\hat{\sigma}-3}\left[\hat{\sigma}u_i-\hat{\sigma}-u_i^{\hat{\sigma}}+1\right]}{\left(u_i^{\hat{\sigma}}-1\right)^3}$$

where $\tau(i,j)$ represents the (i,j)th element of $\left[-\frac{\partial^2 l(\sigma,\beta)}{\partial\sigma\partial\beta}\right]^{-1}$, $i,j=1,2$. Using the above expression, the BEs of entropy in Equations (14)–(16) can be specifically expressed as follows:

(i) BE under WSELF.

In this case,

$$U_w(\sigma,\beta) = H^{-1}(f)\left[-\ln 2\sigma\beta^2 - 2\sigma\beta^2\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + [\phi(\sigma+1)-\phi(1)] - \frac{1}{\sigma}+1\right]^{-1},$$

$$U_{w\sigma} = -H^{-2}(f)\left[-\frac{1}{\sigma} + \frac{1}{\sigma^2} + \phi'(\sigma+1) - 2\beta^2\int_0^\infty (1-u)u^{\sigma-1}(1+\sigma\ln u)x\ln x dx\right],$$

$$U_{w\sigma\sigma} = 2H^{-3}(f)\left[-\tfrac{1}{\sigma}+\tfrac{1}{\sigma^2}+\phi'(\sigma+1) - 2\beta^2\textstyle\int_0^\infty (1-u)u^{\sigma-1}(1+\sigma\ln u)x\ln x dx\right]^2$$
$$-H^{-2}(f)\left[\tfrac{1}{\sigma^2}-\tfrac{2}{\sigma^3}+\phi''(\sigma+1) - 2\beta^2\textstyle\int_0^\infty (1-u)u^{\sigma-1}(2+\sigma\ln u)x\ln u\ln x dx\right],$$

$$U_{w\beta} = -H^{-2}(f)\left[-\frac{2}{\beta} - 4\sigma\beta\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + 4\sigma\beta^3\int_0^\infty (\sigma u-\sigma+1)(1-u)u^{\sigma-2}x^3\ln x dx\right]$$

$$U_{w\beta\beta} = 2H^{-3}(f)[-\tfrac{2}{\beta} - 4\sigma\beta\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + 4\sigma\beta^3\int_0^\infty (\sigma u-\sigma+1)(1-u)u^{\sigma-2}x^3\ln x dx]^2$$
$$+H^{-2}(f)\{\tfrac{2}{\beta^2} - 4\sigma\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + 20\sigma\beta^2\int_0^\infty (\sigma u-\sigma+1)(1-u)u^{\sigma-2}x^3\ln x dx$$
$$+8\sigma\beta^4\int_0^\infty [\sigma(1-u)-(\sigma u-\sigma+1)(\sigma u-\sigma-u+2)u^{-1}](1-u)u^{\sigma-2}x^5\ln x dx\},$$

$$\begin{aligned} U_{w\sigma\beta} &= 2H^{-3}(f)[-\tfrac{2}{\beta} - 4\sigma\beta\textstyle\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + 4\sigma\beta^3\int_0^\infty (\sigma u-\sigma+1)(1-u)u^{\sigma-2}x^3\ln x dx] \\ &\times[-\tfrac{1}{\sigma} + \tfrac{1}{\sigma^2} + \phi'(\sigma+1) - 2\beta^2\textstyle\int_0^\infty (1-u)u^{\sigma-1}(1+\sigma\ln u)x\ln x dx] \\ &-H^{-2}(f)\{-4\beta\textstyle\int_0^\infty (1-u)u^{\sigma-1}(1+\sigma\ln u)x\ln x dx \\ &+4\beta^3\textstyle\int_0^\infty x^3(1-u)u^{\sigma-2}[(1+\sigma\ln u)(\sigma u-\sigma+1)-\sigma(1-u)]\ln x dx\} = U_{w\beta\sigma}. \end{aligned}$$

Therefore,

$$\begin{aligned} E\big(H^{-1}(f)|x\big) &= U_w(\hat\sigma,\hat\beta) + \tfrac{1}{2}[(\hat U_{w\sigma\sigma} + 2\hat U_{w\sigma}\hat P_\sigma)\hat\tau_{\sigma\sigma} + (\hat U_{w\beta\sigma} + 2\hat U_{w\beta}\hat P_\sigma)\hat\tau_{\beta\sigma} \\ &+(\hat U_{w\sigma\beta} + 2\hat U_{w\sigma}\hat P_\beta)\hat\tau_{\sigma\beta} + (\hat U_{w\beta\beta} + 2\hat U_{w\beta}\hat P_\beta)\hat\tau_{\beta\beta}] \\ &+\tfrac{1}{2}[(\hat U_{w\sigma}\hat\tau_{\sigma\sigma} + \hat U_{w\beta}\hat\tau_{\sigma\beta})(\hat l_{\sigma\sigma\sigma}\hat\tau_{\sigma\sigma} + \hat l_{\sigma\beta\sigma}\hat\tau_{\sigma\beta} + \hat l_{\beta\sigma\sigma}\hat\tau_{\beta\sigma} + \hat l_{\beta\beta\sigma}\hat\tau_{\beta\beta}) \\ &+(\hat U_{w\sigma}\hat\tau_{\beta\sigma} + \hat U_{w\beta}\hat\tau_{\beta\beta})(\hat l_{\beta\sigma\sigma}\hat\tau_{\sigma\sigma} + \hat l_{\sigma\beta\beta}\hat\tau_{\sigma\beta} + \hat l_{\beta\sigma\beta}\hat\tau_{\beta\sigma} + \hat l_{\beta\beta\beta}\hat\tau_{\beta\beta})]. \end{aligned}$$

Then, from the expression $\hat H(f)_{L_{WB}} = \big[E\big(H^{-1}(f)|x\big)\big]^{-1}$, we can obtain the BE under the WSELF.

(ii) BE under PLF.

In this case,

$$U_p(\sigma,\beta) = H^2(f) = \left[-\ln 2\sigma\beta^2 - 2\sigma\beta^2\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + [\phi(\sigma+1)-\phi(1)] - \frac{1}{\sigma} + 1\right]^2,$$

$$U_{p\sigma} = 2H(f)\left[-\frac{1}{\sigma} + \frac{1}{\sigma^2} + \phi'(\sigma+1) - 2\beta^2\int_0^\infty (1-u)u^{\sigma-1}(1+\sigma\ln u)x\ln x dx\right],$$

$$\begin{aligned} U_{p\sigma\sigma} &= 2\Big[-\tfrac{1}{\sigma} + \tfrac{1}{\sigma^2} + \phi'(\sigma+1) - 2\beta^2\textstyle\int_0^\infty (1-u)u^{\sigma-1}(1+\sigma\ln u)x\ln x dx\Big]^2 \\ &+2H(f)\Big[\tfrac{1}{\sigma^2} - \tfrac{2}{\sigma^3} + \phi''(\sigma+1) - 2\beta^2\textstyle\int_0^\infty (1-u)u^{\sigma-1}(2+\sigma\ln u)x\ln u\ln x dx\Big], \end{aligned}$$

$$U_{p\beta} = 2H(f)\left[-\frac{2}{\beta} - 4\sigma\beta\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + 4\sigma\beta^3\int_0^\infty (\sigma u-\sigma+1)(1-u)u^{\sigma-2}x^3\ln x dx\right],$$

$$\begin{aligned} U_{p\beta\beta} &= 2\Big[-\tfrac{2}{\beta} - 4\sigma\beta\textstyle\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + 4\sigma\beta^3\int_0^\infty (\sigma u-\sigma+1)(1-u)u^{\sigma-2}x^3\ln x dx\Big]^2 \\ &+2H(f)\{\tfrac{2}{\beta^2} - 4\sigma\textstyle\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + 20\sigma\beta^2\int_0^\infty (\sigma u-\sigma+1)(1-u)u^{\sigma-2}x^3\ln x dx \\ &+8\sigma\beta^4\textstyle\int_0^\infty \big[\sigma(1-u)-(\sigma u-\sigma+1)(\sigma u-\sigma-u+2)u^{-1}\big](1-u)u^{\sigma-2}x^5\ln x dx\}, \\ U_{p\sigma\beta} &= 2\Big[-\tfrac{2}{\beta} - 4\sigma\beta\textstyle\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + 4\sigma\beta^3\int_0^\infty (\sigma u-\sigma+1)(1-u)u^{\sigma-2}x^3\ln x dx\Big] \\ &\times\Big[-\tfrac{1}{\sigma} + \tfrac{1}{\sigma^2} + \phi'(\sigma+1) - 2\beta^2\textstyle\int_0^\infty (1-u)u^{\sigma-1}(1+\sigma\ln u)x\ln x dx\Big] \\ &+2H(f)\{-4\beta\textstyle\int_0^\infty (1-u)u^{\sigma-1}(1+\sigma\ln u)x\ln x dx \\ &+4\beta^3\textstyle\int_0^\infty x^3(1-u)u^{\sigma-2}[(1+\sigma\ln u)(\sigma u-\sigma+1)-\sigma(1-u)]\ln x dx\} = U_{p\beta\sigma}. \end{aligned}$$

Therefore,

$$\begin{aligned} E\left(H^2(f)|x\right) &= U_p(\hat{\sigma},\hat{\beta}) + \tfrac{1}{2}[(\hat{U}_{p\sigma\sigma} + 2\hat{U}_{p\sigma}\hat{P}_\sigma)\hat{\tau}_{\sigma\sigma} + (\hat{U}_{p\beta\sigma} + 2\hat{U}_{p\beta}\hat{P}_\sigma)\hat{\tau}_{\beta\sigma} \\ &\quad + (\hat{U}_{p\sigma\beta} + 2\hat{U}_{p\sigma}\hat{P}_\beta)\hat{\tau}_{\sigma\beta} + (\hat{U}_{p\beta\beta} + 2\hat{U}_{p\beta}\hat{P}_\beta)\hat{\tau}_{\beta\beta}] \\ &\quad + \tfrac{1}{2}\Big[(\hat{U}_{p\sigma}\hat{\tau}_{\sigma\sigma} + \hat{U}_{p\beta}\hat{\tau}_{\sigma\beta})\left(\hat{l}_{\sigma\sigma\sigma}\hat{\tau}_{\sigma\sigma} + \hat{l}_{\sigma\beta\sigma}\hat{\tau}_{\sigma\beta} + \hat{l}_{\beta\sigma\sigma}\hat{\tau}_{\beta\sigma} + \hat{l}_{\beta\beta\sigma}\hat{\tau}_{\beta\beta}\right) \\ &\quad + (\hat{U}_{p\sigma}\hat{\tau}_{\beta\sigma} + \hat{U}_{p\beta}\hat{\tau}_{\beta\beta})\left(\hat{l}_{\beta\sigma\sigma}\hat{\tau}_{\sigma\sigma} + \hat{l}_{\sigma\beta\beta}\hat{\tau}_{\sigma\beta} + \hat{l}_{\beta\sigma\beta}\hat{\tau}_{\beta\sigma} + \hat{l}_{\beta\beta\beta}\hat{\tau}_{\beta\beta}\right)\Big]. \end{aligned}$$

Then, from the expression $\hat{H}(f)_{L_{PB}} = \sqrt{E(H^2(f)|x)}$, we can obtain the BE under the PLF.

(iii) BE under KLF.

In this case, we first let

$$U_k(\sigma,\beta) = \frac{H(f)}{H^{-1}(f)},$$

$$U_k(\sigma,\beta) = \frac{-\ln 2\sigma\beta^2 - 2\sigma\beta^2\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + [\phi(\sigma+1)-\phi(1)] - \frac{1}{\sigma} + 1}{\left[-\ln 2\sigma\beta^2 - 2\sigma\beta^2\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + [\phi(\sigma+1)-\phi(1)] - \frac{1}{\sigma} + 1\right]^{-1}},$$

It can be seen from (i) that

$$\begin{aligned} E\left(H^{-1}(f)|x\right) &= U_w(\hat{\sigma},\hat{\beta}) + \tfrac{1}{2}[(\hat{U}_{w\sigma\sigma} + 2\hat{U}_{w\sigma}\hat{P}_\sigma)\hat{\tau}_{\sigma\sigma} + (\hat{U}_{w\beta\sigma} + 2\hat{U}_{w\beta}\hat{P}_\sigma)\hat{\tau}_{\beta\sigma} \\ &\quad + (\hat{U}_{w\sigma\beta} + 2\hat{U}_{w\sigma}\hat{P}_\beta)\hat{\tau}_{\sigma\beta} + (\hat{U}_{w\beta\beta} + 2\hat{U}_{w\beta}\hat{P}_\beta)\hat{\tau}_{\beta\beta}] \\ &\quad + \tfrac{1}{2}\Big[(\hat{U}_{w\sigma}\hat{\tau}_{\sigma\sigma} + \hat{U}_{w\beta}\hat{\tau}_{\sigma\beta})\left(\hat{l}_{\sigma\sigma\sigma}\hat{\tau}_{\sigma\sigma} + \hat{l}_{\sigma\beta\sigma}\hat{\tau}_{\sigma\beta} + \hat{l}_{\beta\sigma\sigma}\hat{\tau}_{\beta\sigma} + \hat{l}_{\beta\beta\sigma}\hat{\tau}_{\beta\beta}\right) \\ &\quad + (\hat{U}_{w\sigma}\hat{\tau}_{\beta\sigma} + \hat{U}_{w\beta}\hat{\tau}_{\beta\beta})\left(\hat{l}_{\beta\sigma\sigma}\hat{\tau}_{\sigma\sigma} + \hat{l}_{\sigma\beta\beta}\hat{\tau}_{\sigma\beta} + \hat{l}_{\beta\sigma\beta}\hat{\tau}_{\beta\sigma} + \hat{l}_{\beta\beta\beta}\hat{\tau}_{\beta\beta}\right)\Big]. \end{aligned}$$

So, next, we need to find $E(H(f)|x)$:

$$U(\sigma,\beta) = H(f) = -\ln 2\sigma\beta^2 - 2\sigma\beta^2\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + [\phi(\sigma+1)-\phi(1)] - \frac{1}{\sigma} + 1,$$

$$U_\sigma = -\frac{1}{\sigma} + \frac{1}{\sigma^2} + \phi'(\sigma+1) - 2\beta^2\int_0^\infty (1-u)u^{\sigma-1}(1+\sigma\ln u)x\ln x dx,$$

$$U_{\sigma\sigma} = \frac{1}{\sigma^2} - \frac{2}{\sigma^3} + \phi''(\sigma+1) - 2\beta^2\int_0^\infty (1-u)u^{\sigma-1}(2+\sigma\ln u)x\ln u\ln x dx,$$

$$U_\beta = -\frac{2}{\beta} - 4\sigma\beta\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + 4\sigma\beta^3\int_0^\infty (\sigma u-\sigma+1)(1-u)u^{\sigma-2}x^3\ln x dx,$$

$$\begin{aligned} U_{\beta\beta} &= \tfrac{2}{\beta^2} - 4\sigma\textstyle\int_0^\infty (1-u)u^{\sigma-1}x\ln x dx + 20\sigma\beta^2\int_0^\infty (\sigma u-\sigma+1)(1-u)u^{\sigma-2}x^3\ln x dx \\ &\quad + 8\sigma\beta^4\textstyle\int_0^\infty \left[\sigma(1-u) - (\sigma u-\sigma+1)(\sigma u-\sigma-u+2)u^{-1}\right](1-u)u^{\sigma-2}x^5\ln x dx, \end{aligned}$$

$$\begin{aligned} U_{\sigma\beta} &= -4\beta\textstyle\int_0^\infty (1-u)u^{\sigma-1}(1+\sigma\ln u)x\ln x dx \\ &\quad + 4\beta^3\textstyle\int_0^\infty x^3(1-u)u^{\sigma-2}[(1+\sigma\ln u)(\sigma u-\sigma+1) - \sigma(1-u)]\ln x dx = U_{\beta\sigma}. \end{aligned}$$

Therefore,

$$\begin{aligned} E(H(f)|x) &= U(\hat{\sigma},\hat{\beta}) + \tfrac{1}{2}[(\hat{U}_{\sigma\sigma} + 2\hat{U}_{\sigma}\hat{P}_{\sigma})\hat{\tau}_{\sigma\sigma} + (\hat{U}_{\beta\sigma} + 2\hat{U}_{\beta}\hat{P}_{\sigma})\hat{\tau}_{\beta\sigma} \\ &\quad + (\hat{U}_{\sigma\beta} + 2\hat{U}_{\sigma}\hat{P}_{\beta})\hat{\tau}_{\sigma\beta} + (\hat{U}_{\beta\beta} + 2\hat{U}_{\beta}\hat{P}_{\beta})\hat{\tau}_{\beta\beta}] \\ &\quad + \tfrac{1}{2}\Big[(\hat{U}_{\sigma}\hat{\tau}_{\sigma\sigma} + \hat{U}_{\beta}\hat{\tau}_{\sigma\beta})\Big(\hat{l}_{\sigma\sigma\sigma}\hat{\tau}_{\sigma\sigma} + \hat{l}_{\sigma\beta\sigma}\hat{\tau}_{\sigma\beta} + \hat{l}_{\beta\sigma\sigma}\hat{\tau}_{\beta\sigma} + \hat{l}_{\beta\beta\sigma}\hat{\tau}_{\beta\beta}\Big) \\ &\quad + (\hat{U}_{\sigma}\hat{\tau}_{\beta\sigma} + \hat{U}_{\beta}\hat{\tau}_{\beta\beta})\Big(\hat{l}_{\beta\sigma\sigma}\hat{\tau}_{\sigma\sigma} + \hat{l}_{\sigma\beta\beta}\hat{\tau}_{\sigma\beta} + \hat{l}_{\beta\sigma\beta}\hat{\tau}_{\beta\sigma} + \hat{l}_{\beta\beta\beta}\hat{\tau}_{\beta\beta}\Big)\Big]. \end{aligned}$$

Then, from the expression $\hat{H}(f)_{L_{KB}} = \sqrt{\frac{E(H(f)|x)}{E\left(H^{-1}(f)|x\right)}}$, we can obtain the BE under KLF.

3.3. MCMC Method for Mixed Gibbs Sampling

In addition to using Lindley's approximation method to calculate Bayesian estimates, we can also use the MCMC method to calculate Bayesian estimates. We know that calculating Bayesian estimates generally boils down to a posterior distribution $\pi(x)$. The integral calculation is often complex and inconvenient, so using the MCMC method can effectively solve such problems. The essence of the MCMC method is to establish a stationary distribution of $\pi(x)$ to generate a posterior distribution through the Markov chain of $\pi(x)$ based on the generated samples; then, a further statistical inference can be made. Nowadays, the MCMC method is widely used in practice, playing an important role in the processing of statistical signals and various model estimates. For details, please refer to [35–37]. The commonly used MCMC methods include Metropolis–Hastings (M-H) sampling and Gibbs sampling. M-H sampling is suitable for low-dimensional posterior distribution sampling, while Gibbs sampling is suitable for high-dimensional and non-standard posterior distribution sampling. This article introduces the mixed Gibbs sampling method, which combines Gibbs sampling with M-H sampling. In some Gibbs sampling steps, the M-H algorithm is used to extract random numbers, making sampling easier. According to Equation (13), the joint posterior distribution of parameters is

$$\pi(\sigma,\beta|x^*) \propto \sigma^{a_1+m-1}\beta^{a_2+2m-1}e^{-b_1\sigma}\prod_{i=1}^{m} x_i e^{-\beta^2 x_i^2}\left(1-e^{-\beta^2 x_i^2}\right)^{\sigma-1}\left[1-\left(1-e^{-\beta^2 x_i^2}\right)^{\sigma}\right]^{P_i}.$$

So, the posterior distributions (i.e., fully conditional distributions) of σ, β are:

$$\pi(\sigma|x^*,\beta) \propto \sigma^{a_1+m-1}e^{-b_1\sigma}\prod_{i=1}^{m} x_i e^{-\beta^2 x_i^2}\left(1-e^{-\beta^2 x_i^2}\right)^{\sigma-1}\left[1-\left(1-e^{-\beta^2 x_i^2}\right)^{\sigma}\right]^{P_i}, \quad (18)$$

$$\pi(\beta|x^*,\sigma) \propto \sigma^{a_2+2m-1}e^{-b_2\beta}\prod_{i=1}^{m} x_i e^{-\beta^2 x_i^2}\left(1-e^{-\beta^2 x_i^2}\right)^{\sigma-1}\left[1-\left(1-e^{-\beta^2 x_i^2}\right)^{\sigma}\right]^{P_i}, \quad (19)$$

As can be seen from Equations (18) and (19), the posterior distribution of σ, β is non-standard, so we use mixed Gibbs sampling to obtain samples of parameters. Here are the steps for Gibbs sampling:

(1) Using the M-H algorithm to extract $\sigma^{(i+1)}$ samples from $\pi\left(\sigma|x^*,\beta^{(i)}\right)$.

The recommended distribution $q(\sigma')$ of select σ is normal distribution $N\left(\sigma_k,\mu_\sigma^2\right)$, σ_k is the current state, and μ_σ^2 for σ is the variance.

(i) Extract samples σ' from $N(\sigma_k, \mu_\sigma^2)$. When $\sigma' \leq 0$, resample. Next, calculate the acceptance probability:

$$p(\sigma_k, \sigma') = \min\left\{1, \frac{\pi(\sigma'|x^*, \beta^{(i)})}{\pi(\sigma_k|x^*, \beta^{(i)})}\right\} = \min\left\{1, \frac{(\sigma')^{a_1+m-1}e^{-b_1\sigma'}\prod_{i=1}^{m}\left(1-e^{-(\beta^{(i)})^2x_i^2}\right)^{\sigma'-1}\left[1-\left(1-e^{-(\beta^{(i)})^2x_i^2}\right)^{\sigma'}\right]^{P_i}}{\sigma_k^{a_1+m-1}e^{-b_1\sigma_k}\prod_{i=1}^{m}\left(1-e^{-(\beta^{(i)})^2x_i^2}\right)^{\sigma_k-1}\left[1-\left(1-e^{-(\beta^{(i)})^2x_i^2}\right)^{\sigma_k}\right]^{P_i}}\right\}.$$

(ii) Extract sample λ_1 from uniformly distributed $U(0,1)$, set $\sigma_{k+1} = \begin{cases}\sigma', \lambda_1 \leq p(\sigma_k, \sigma') \\ \sigma_k, \lambda_1 > p(\sigma_k, \sigma')\end{cases}$

(iii) Let k = k + 1, and return to step (i).

(2) Using the M-H algorithm to extract $\beta^{(i+1)}$ samples from $\pi\left(\beta|x^*, \sigma^{(i+1)}\right)$.

The recommended distribution $q(\beta')$ of select β is normal distribution $N\left(\beta_k, \mu_\beta^2\right)$, β_k is the current state, and μ_β^2 for β is the variance.

(i) Extract samples β' from $N\left(\beta_k, \mu_\beta^2\right)$. When $\beta' \leq 0$, resample. Next, calculate the acceptance probability:

$$p(\beta_k, \beta') = \min\left\{1, \frac{\pi(\beta'|x^*, \sigma^{(i+1)})}{\pi(\beta_k|x^*, \sigma^{(i+1)})}\right\} = \min\left\{1, \frac{(\beta')^{a_1+2m-1}e^{-b_2\beta'}e^{-(\beta')^2\sum_{i=1}^{m}x_i^2}\prod_{i=1}^{m}\left(1-e^{-(\beta')^2x_i^2}\right)^{\sigma^{(i+1)}-1}\left[1-\left(1-e^{-(\beta')^2x_i^2}\right)^{\sigma^{(i+1)}}\right]^{P_i}}{\beta_k^{a_1+2m-1}e^{-b_2\beta_k}e^{-\beta_k^2\sum_{i=1}^{m}x_i^2}\prod_{i=1}^{m}\left(1-e^{-\beta_k^2x_i^2}\right)^{\sigma^{(i+1)}-1}\left[1-\left(1-e^{-\beta_k^2x_i^2}\right)^{\sigma^{(i+1)}}\right]^{P_i}}\right\}.$$

(ii) Extract sample λ_2 from uniformly distributed $U(0,1)$, and set $\beta_{k+1} = \begin{cases}\beta', \lambda_2 \leq p(\beta_k, \beta') \\ \beta_k, \lambda_2 > p(\beta_k, \beta')\end{cases}$

(iii) Let k = k + 1, and return to step (i).

4. Monte Carlo Simulation

The Monte Carlo method is a calculation method based on random numbers which is commonly used to solve complex problems. Its basic idea is to simulate the experimental process through random sampling to obtain approximate solutions. The core of the Monte Carlo method is the generation of random numbers which may follow a certain distribution. Through a large number of random sampling, approximate solutions can be obtained, and the accuracy of the approximate solutions will gradually improve with the increase in sampling times. In the following section, we will use this method to compare the performance of different estimates.

(1) First, set the PC-II schemes (See Table 1).

(2) According to the set PC-II scheme, the PC-II data can be obtained through the following algorithm (Algorithm 1):

Table 1. Censoring schemes.

Scheme	
a	$P_1 = P_2 = \cdots = P_{m-1} = 0, P_m = n - m$
b	$P_1 = P_2 = P_3 = \cdots = P_m = 1$
c	$P_1 = n - m, P_2 = \cdots = P_{m-1} = 0$

Algorithm 1. Generating PC-II data under $GRD(\sigma, \beta)$

1. Generate m observations $G_i (1 \leq i \leq m)$ that follow a (0,1) uniform distribution.
2. Assign corresponding values to different censored schemes, and set $H_i = G_i^{1/(i+\sum_{j=m-i+1}^{m} P_j)}$.
3. Generate $Z_i = 1 - H_m H_{m-1} \cdots H_{m-i+1} (i = 1, 2, \cdots, m)$ through H_i.
4. Given the initial parameters σ, β, set $x_i = F^{-1}(Z_i) = \frac{1}{\beta}[\ln(1 - Z_i^{\frac{1}{\sigma}})]^{\frac{1}{2}}$; then, x_i is the asymptotic Type-II censored data.
5. Calculate the ML estimates of σ, β using Equations (8) and (9), and calculate the ML estimates of entropy $H(f)$ and ACI using Equation (12).
6. Set the superparameters $a_1 = 1, b_1 = 1, a_2 = 1, b_2 = 1$. For different loss functions, use Lindley's approximation method to calculate the exact value of the entropy of Equations (14)–(16)
7. Repeat steps 1 to 6 a total of 1000 times, and calculate the ML estimates, MSE, coverage probability (CV), and variance; the calculation results are shown in the table below.

We set $\sigma = 2$, $\beta = 1$, and the superparameters under the gamma prior $a_1 = 1$, $b_1 = 1$, $a_2 = 1$, $b_2 = 1$. The ML estimates and Bayesian estimates of different censoring schemes (CSs) under the $GRD(\sigma, \beta)$, and their corresponding MSEs (in parentheses), are given in Table 2.

Table 2. The ML estimates and Bayesian estimates of different censoring schemes (CSs) under the $GRD(\sigma, \beta)$ and their corresponding MSEs (in parentheses).

			MLE	Lindley			MCMC		
n	m	CS	M_{ML}	M_w	M_P	M_K	M_w	M_P	M_K
30	10	a	0.4103 (0.0951)	0.3930 (0.0977)	0.4568 (0.0665)	0.3173 (0.1125)	0.7545 (0.0498)	0.8294 (0.0749)	0.7788 (0.0749)
		b	0.4354 (0.0721)	0.4144 (0.0753)	0.4698 (0.0514)	0.3252 (0.0923)	0.6278 (0.0052)	0.7333 (0.0315)	0.6621 (0.0113)
		c	0.4868 (0.0503)	0.4599 (0.0548)	0.5141 (0.0296)	0.3651 (0.0799)	0.7792 (0.0500)	0.8092 (0.0643)	0.7891 (0.0544)
	15	a	0.4825 (0.0466)	0.4688 (0.0490)	0.4999 (0.0419)	0.3603 (0.0804)	0.6821 (0.0160)	0.7243 (0.0284)	0.6961 (0.0197)
		b	0.4876 (0.0414)	0.4726 (0.0439)	0.5054 (0.0380)	0.3640 (0.0744)	0.6216 (0.0043)	0.6441 (0.0078)	0.6291 (0.0054)
		c	0.5043 (0.0325)	0.4852 (0.0350)	0.5205 (0.0292)	0.3729 (0.0635)	0.6807 (0.0156)	0.7122 (0.0245)	0.6914 (0.0184)
	20	a	0.5044 (0.0290)	0.4926 (0.0305)	0.5159 (0.0278)	0.3739 (0.0609)	0.6649 (0.0119)	0.6827 (0.0161)	0.6708 (0.0132)
		b	0.5100 (0.0313)	0.4987 (0.0327)	0.5214 (0.0298)	0.3783 (0.0623)	0.5896 (0.0011)	0.6142 (0.0034)	0.5976 (0.0018)
		c	0.5083 (0.0243)	0.4946 (0.0256)	0.5207 (0.0234)	0.3737 (0.0545)	0.5670 (0.0002)	0.6121 (0.0032)	0.5841 (0.0008)
40	10	a	0.4130 (0.1044)	0.3955 (0.1702)	0.4631 (0.0757)	0.3289 (0.1196)	0.6840 (0.0165)	0.7163 (0.0258)	0.6949 (0.0194)
		b	0.4483 (0.0722)	0.4267 (0.0751)	0.4866 (0.0533)	0.3413 (0.0971)	0.7307 (0.0306)	0.7932 (0.0564)	0.7524 (0.0387)
		c	0.4870 (0.0452)	0.4590 (0.0495)	0.5121 (0.0330)	0.3614 (0.0739)	0.6416 (0.0074)	0.6936 (0.0190)	0.6593 (0.0107)
	15	a	0.4702 (0.0565)	0.4574 (0.0583)	0.4910 (0.0495)	0.3533 (0.0881)	0.6292 (0.0054)	0.6587 (0.0106)	0.6391 (0.0069)
		b	0.4842 (0.0413)	0.4693 (0.0435)	0.5015 (0.0376)	0.3597 (0.0743)	0.6748 (0.0142)	0.6919 (0.0186)	0.6807 (0.0156)
		c	0.5060 (0.0328)	0.4873 (0.0351)	0.5223 (0.0303)	0.3731 (0.0633)	0.5874 (0.0010)	0.6208 (0.0042)	0.5984 (0.0018)
	20	a	0.4934 (0.0332)	0.4829 (0.0346)	0.5049 (0.0315)	0.3645 (0.0658)	0.6097 (0.0030)	0.6329 (0.0060)	0.6174 (0.0038)
		b	0.5077 (0.0313)	0.4964 (0.0330)	0.5193 (0.0299)	0.3759 (0.0662)	0.6580 (0.0105)	0.6880 (0.0175)	0.6680 (0.0126)
		c	0.5201 (0.0249)	0.5066 (0.0264)	0.5324 (0.0234)	0.3861 (0.0557)	0.5681 (0.0002)	0.5948 (0.0015)	0.5772 (0.0005)
50	15	a	0.4728 (0.0611)	0.4610 (0.0630)	0.4926 (0.0542)	0.3562 (0.0931)	0.6316 (0.0058)	0.6636 (0.0116)	0.6422 (0.0075)
		b	0.4860 (0.0470)	0.4709 (0.0494)	0.5042 (0.0403)	0.3629 (0.0782)	0.6450 (0.0080)	0.6739 (0.0140)	0.6543 (0.0097)
		c	0.5160 (0.0257)	0.4975 (0.0277)	0.5319 (0.0241)	0.3840 (0.0565)	0.6145 (0.0035)	0.6757 (0.0144)	0.6341 (0.0061)
	20	a	0.4929 (0.0391)	0.4832 (0.0403)	0.5056 (0.0372)	0.3687 (0.0707)	0.6039 (0.0023)	0.6260 (0.0049)	0.6112 (0.0031)
		b	0.5058 (0.0316)	0.4946 (0.0330)	0.5176 (0.0296)	0.3752 (0.0634)	0.6432 (0.0077)	0.6528 (0.0094)	0.6465 (0.0082)
		c	0.5205 (0.0217)	0.5069 (0.0232)	0.5329 (0.0200)	0.3865 (0.0515)	0.6048 (0.0024)	0.6559 (0.0100)	0.6215 (0.0043)
	30	a	0.5222 (0.0209)	0.5147 (0.0216)	0.5294 (0.0203)	0.3884 (0.0527)	0.5792 (0.0006)	0.5921 (0.0013)	0.5836 (0.0008)
		b	0.5236 (0.0175)	0.5162 (0.0180)	0.5310 (0.0172)	0.3876 (0.0456)	0.5778 (0.0005)	0.5910 (0.0012)	0.5822 (0.0007)
		c	0.5265 (0.0168)	0.5174 (0.0176)	0.5346 (0.0164)	0.3891 (0.0464)	0.5769 (0.0004)	0.5909 (0.0011)	0.5816 (0.0006)

Based on the above tables, the following conclusions can be drawn for these research results:

(1) For Shannon entropy, Bayesian estimation performs better overall than ML estimation.
(2) In Bayesian estimation, the overall performance of the MCMC method is superior to Lindley's approximation method.
(3) When the number of observed faults is fixed, MSE shows an upward trend as the sample size increases; when the sample size is fixed, MSE shows a decreasing trend as the number of observed faults increases.

(4) In Lindley's approximation, the Bayesian estimate under PLF is superior to the other two loss functions; in MCMC sampling, the performance of Bayesian estimates under WSELF and KLF is similar, both of which are better than those under PLF.
(5) As shown in Table 3, the coverage probability at that time tended to be more at a confidence level.

Table 3. Lower 100 $(1-\alpha)$% coverage and variance with different α.

n	m	CS	α = 0.1	α = 0.05	$V(\hat{H}(f))$
30	10	a	0.7480	0.7160	0.0219
		b	0.8240	0.7480	0.0095
		c	0.9110	0.8750	0.0149
	15	a	0.8110	0.7650	0.0063
		b	0.8560	0.7950	0.0076
		c	0.9170	0.8590	0.0135
	20	a	0.8570	0.7960	0.0064
		b	0.8880	0.8090	0.0069
		c	0.8940	0.8740	0.0089
40	10	a	0.7290	0.6970	0.0090
		b	0.8220	0.7510	0.0219
		c	0.9230	0.8820	0.0112
	15	a	0.7930	0.7240	0.0069
		b	0.8420	0.7700	0.0087
		c	0.9100	0.8770	0.0171
	20	a	0.8390	0.7690	0.0044
		b	0.8310	0.8270	0.0037
		c	0.9190	0.8780	0.0070
50	15	a	0.7620	0.7160	0.0030
		b	0.8510	0.8020	0.0107
		c	0.9410	0.8880	0.0099
	20	a	0.8060	0.7530	0.0040
		b	0.8590	0.8120	0.0055
		c	0.9200	0.8950	0.0061
	30	a	0.8520	0.7810	0.0050
		b	0.8720	0.8300	0.0028
		c	0.9270	0.8810	0.0052

5. Real Data Analysis

In this section, we consider using a real-world dataset to demonstrate the feasibility of the estimation method. Raqab [38] mentioned a set of data, which represents the total annual rainfall of Los Angeles in the past 25 years from 1985 to 2009. By analyzing the scaled TTT, the $GRD(\sigma, \beta)$ model can provide a reasonable fit for this set of data. This article provides the following dataset: 12.82, 17.86, 7.66, 12.48, 8.08, 7.35, 11.99, 21.00, 27.36, 8.11, 24.35, 12.44, 12.40, 31.01, 9.09, 11.57, 17.94, 4.42, 16.42, 9.25, 37.96, 13.19, 3.21, 13.53, 9.08.

Using the PC-II scheme a: $P_1 = P_2 = \cdots = P_{m-1} = 0, P_m = n - m$, taking m = 16, the censored data under this censoring scheme are:

12.82, 17.86, 7.66, 12.48, 8.08, 7.35, 11.99, 21.00, 27.36, 8.11, 24.35, 12.44, 12.40, 31.01, 9.09, 11.57. In order to further analyze whether the obtained censored data also fit $GRD(\sigma, \beta)$, we drew an empirical distribution function graph under the censored data, and compared it with a $GRD(\sigma, \beta)$ cumulative distribution function graph. Through image visualization, we can conclude that the censored data can reasonably fit the distribution (see Figure 2).

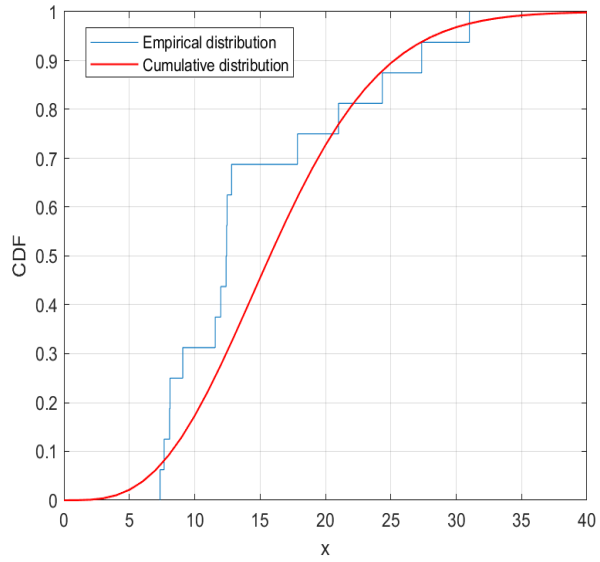


Figure 2. Empirical distribution under censored data and cumulative distribution function of $GRD(\sigma, \beta)$.

Before calculating real data, it is necessary to explain the existence and uniqueness of the ML estimation solution. Due to Equations (8) and (9) being nonlinear, proving them directly would be complex. Let us consider Equations (8) and (9); they are visually represented through the image shown in Figure 3, where $l(\sigma)$ represents $\frac{\partial l(\sigma,\beta)}{\partial \sigma}$ in Equation (8) and $l(\beta)$ represents in Equation (9).

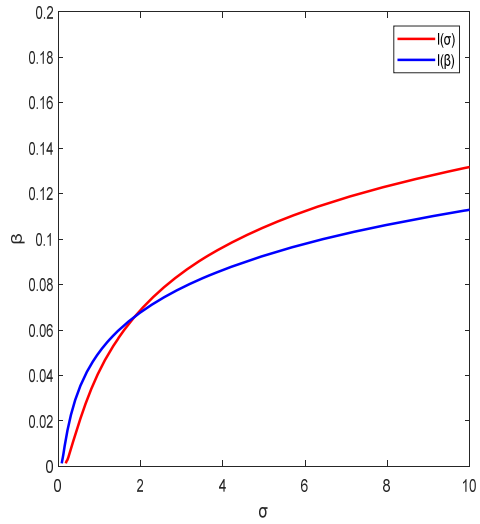


Figure 3. Partial derivative of log-LF.

From Figure 3, we find that the two curves only have one intersection point, so we conclude that the solution for the ML estimation exists and is unique.

The estimation method described in this article was used to perform ML estimation and Bayesian estimation on the real data mentioned above and obtain parameters $\sigma = 1.7051,\ \beta = 0.0665$ using real data; the specific results are shown in the Table 4.

Table 4. Estimates of Shannon entropy.

ML Estimates	Bayesian Estimates					
	Lindley			MCMC		
	under WSEKF	under PLF	under KLF	under WSEKF	under PLF	under KLF
3.2815	3.2704	3.2844	5.9348	3.1649	3.1737	3.1678

6. Conclusions

This paper first described the $GRD(\sigma, \beta)$ model and briefly outlined the characteristics of the model. Under the PC-II scheme, the Shannon entropy function obeying the $GRD(\sigma, \beta)$ was derived, and the ML estimates of parameters were derived using the numerical Newton iterative method. In order to understand the reliability and stability of the obtained entropy estimation, the Delta method was subsequently used to obtain the ACI of entropy. In order to better choose the estimation method, Bayesian estimation was used to continue estimating the Shannon entropy function. This paper discussed the BEs of three loss functions: the WSELF, PLF, and KLF. Because the BEs under the above three loss functions were in the form of two integral ratios, which are complex to calculate, we chose Lindley's approximation to obtain the Bayesian estimates of the entropy under the loss function. Finally, a Monte Carlo simulation was used to obtain the ML estimates and Bayesian estimate of MSE and CV. When the number of observed faults was fixed, MSE showed an upward trend as the sample size increased; when the sample size was fixed, as the number of observed faults increased, MSE showed a decreasing trend, and a higher CV indicated that the estimated results tended to be more realistic parameter values. Through the simulation in Section 5, it can be concluded that the performance of Bayesian estimation was superior to ML estimation, and the overall performance of the MCMC method was superior to Lindley's approximation method.

The $GRD(\sigma, \beta)$ model is one of the relatively new probability distribution models, and there is relatively little research on the estimation of the entropy of $GRD(\sigma, \beta)$. It was very meaningful to focus on the estimation of the entropy of $GRD(\sigma, \beta)$ in this paper, which not only provided new ideas and methods for the estimation of the entropy of $GRD(\sigma, \beta)$, but also provides a deep understanding of the characteristics of the entropy of $GRD(\sigma, \beta)$, and promotes the research and development of related theories.

Author Contributions: Conceptualization, H.R. and Q.G.; methodology, H.R.; software, Q.G. and X.H.; validation, H.R. and X.H.; writing—original draft preparation, Q.G.; writing—review and editing, H.R.; funding acquisition, H.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 71661012.

Data Availability Statement: Not applicable.

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

Appendix A

According to Equation (1), we can obtain

$$\begin{aligned} \ln f(x) &= \ln\left[2\sigma\beta^2 x e^{-\beta^2 x^2}\left(1-e^{-\beta^2 x^2}\right)^{\sigma-1}\right] \\ &= \ln 2\sigma\beta^2 + \ln x - \beta^2 x^2 + (\sigma-1)\ln\left(1-e^{-\beta^2 x^2}\right), \end{aligned}$$

$$H(f) = -\int_0^\infty f(x)\left[\ln 2\sigma\beta^2 + \ln x - \beta^2x^2 + (\sigma-1)\ln\left(1-e^{-\beta^2x^2}\right)\right]dx$$
$$= -\ln 2\sigma\beta^2\int_0^\infty f(x)dx - \int_0^\infty \ln x f(x)dx + \int_0^\infty \beta^2x^2 f(x)dx$$
$$-\int_0^\infty (\sigma-1)\ln\left(1-e^{-\beta^2x^2}\right)f(x)dx$$
$$= -\ln 2\sigma\beta^2 - E[\ln x] + E\left[\beta^2x^2\right] - (\sigma-1)E\left[\ln\left(1-e^{-\beta^2x^2}\right)\right].$$

Obviously, $\int_0^\infty f(x)dx = 1$; that is

$$\int_0^\infty 2\sigma\beta^2 x e^{-\beta^2x^2}\left(1-e^{-\beta^2x^2}\right)^{\sigma-1}dx = 1$$

At this point, both sides simultaneously align the σ derivative; we have

$$\int_0^\infty 2\beta^2 x e^{-\beta^2x^2}\left(1-e^{-\beta^2x^2}\right)^{\sigma-1}dx + \int_0^\infty 2\sigma\beta^2 x e^{-\beta^2x^2}\left(1-e^{-\beta^2x^2}\right)^{\sigma-1}\ln\left(1-e^{-\beta^2x^2}\right)dx = 0$$

Therefore,

$$\int_0^\infty 2\sigma\beta^2 x e^{-\beta^2x^2}\left(1-e^{-\beta^2x^2}\right)^{\sigma-1}\ln\left(1-e^{-\beta^2x^2}\right) = -\frac{1}{\sigma}$$

Hence,

$$E\left[\ln\left(1-e^{-\beta^2x^2}\right)\right] = -\frac{1}{\sigma}$$

Next, we can calculate that

$$E\left[\beta^2x^2\right] = \int_0^\infty \beta^2x^2\left(2\sigma\beta^2 x e^{-\beta^2x^2}\left(1-e^{-\beta^2x^2}\right)^{\sigma-1}\right)dx$$
$$= \int_0^\infty \sigma(\beta^2x^2)e^{-\beta^2x^2}\left(1-e^{-\beta^2x^2}\right)^{\sigma-1}d(\beta^2x^2).$$

Let $\beta^2x^2 = w$; then,

$$E[w] = \int_0^\infty \sigma w e^{-w}(1-e^{-w})^{\sigma-1}dw$$

We know that the PDF of the generalized exponential distribution is

$$f(v) = \frac{\beta}{\sigma}e^{-\frac{v}{\sigma}}(1-e^{-\frac{v}{\sigma}})^{\beta-1},$$

and know that [13] the following is true:

$$E(v) = \int_0^\infty \sigma\beta\tfrac{v}{\sigma}e^{-\frac{v}{\beta}}(1-e^{-\frac{v}{\beta}})\beta-1d\left(\tfrac{v}{\sigma}\right) = \sigma\int_0^\infty \beta\tfrac{v}{\sigma}e^{-\frac{v}{\beta}}(1-e^{-\frac{v}{\beta}})\beta-1d\left(\tfrac{v}{\sigma}\right)$$
$$= \sigma[\psi(\lambda+1)-\psi(1)],$$

that is to say

$$\int_0^\infty \beta\frac{v}{\sigma}e^{-\frac{v}{\beta}}(1-e^{-\frac{v}{\beta}})^{\beta-1}d\left(\frac{v}{\sigma}\right) = [\psi(\beta+1)-\psi(1)]$$

Let $\frac{v}{\sigma} = w$; then,

$$\int_0^\infty \beta w e^{-w}(1-e^{-w})^{\beta-1}dw = [\psi(\beta+1)-\psi(1)]$$

Thus,

$$E[w] = \int_0^\infty \sigma w e^{-w}(1-e^{-w})^{\sigma-1}dw = [\phi(\sigma+1)-\phi(1)]$$

We can calculate that

$$\begin{aligned} E(x^r) &= \int_0^\infty x^r f(x)dx = \int_0^\infty x^r 2\sigma\beta^2 x e^{-\beta^2x^2}\left(1-e^{-\beta^2x^2}\right)^{\sigma-1}dx \\ &= \int_0^\infty \sigma x^r e^{-\beta^2x^2}\left(1-e^{-\beta^2x^2}\right)^{\sigma-1}d\left(\beta^2x^2\right). \end{aligned}$$

Let $\beta^2x^2 = t$; then, $x = \frac{1}{\beta}t^{\frac{1}{2}}$. Thus,

$$\int_0^\infty \sigma x^r e^{-\beta^2x^2}\left(1-e^{-\beta^2x^2}\right)^{\sigma-1}d\left(\beta^2x^2\right) = \int_0^\infty \sigma\left(\frac{1}{\beta}t^{\frac{1}{2}}\right)^r e^{-t}\left(1-e^{-t}\right)^{\sigma-1}dt$$

$$= \sigma\beta^{-r}\int_0^\infty t^{\frac{1}{2}}e^{-t}\left(1-e^{-t}\right)^{\sigma-1}dt.$$

Let $e^{-t} = u$; then, $t = -\ln u$ and $dt = -\frac{1}{u}du$. Therefore,

$$\begin{aligned} \sigma\beta^{-r}\textstyle\int_0^\infty t^{\frac{1}{2}}e^{-t}(1-e^{-t})\sigma-1dt &= \sigma\beta^{-r}\textstyle\int_0^1(-\ln u)^{\frac{r}{2}}(1-u)\sigma-1du \\ &= \sigma\beta^{-r}\textstyle\int_0^\infty t^{\frac{1}{2}}e^{-t}(1-e^{-t})^{\sigma-1}dt. \end{aligned}$$

We can perform the polynomial decomposition of $(1-u)^{\sigma-1}$ to

$$(1-u)^{\sigma-1} = \sum_{i=0}^{\infty}\binom{\sigma-1}{i}(-1)^i u^i$$

Then,

$$\begin{aligned} \sigma\beta^{-r}\textstyle\int_0^1(-\ln u)^{\frac{r}{2}}(1-u)^{\sigma-1}du &= \sigma\beta^{-r}\textstyle\int_0^1(-\ln u)^{\frac{r}{2}}\sum\limits_{i=0}^{\infty}\binom{\sigma-1}{i}(-1)^i u^i du \\ &= \sigma\beta^{-r}\textstyle\sum\limits_{i=0}^{\infty}\binom{\sigma-1}{i}(-1)^i\int_0^1(-\ln u)^{\frac{r}{2}}u^i du \end{aligned}$$

According to the negative logarithmic gamma distribution:

$$\int_0^1(-\ln u)^{\frac{r}{2}}(1-u)^{\sigma-1}du = \frac{\Gamma\left(\frac{r}{2}+1\right)}{(1+i)^{\frac{r}{2}+1}}$$

Thus,

$$\sigma\beta^{-r}\sum_{i=0}^{\infty}\binom{\sigma-1}{i}(-1)^i\int_0^1(-\ln u)^{\frac{r}{2}}u^i du = \sigma\beta^{-r}\sum_{i=0}^{\infty}\binom{\sigma-1}{i}(-1)^i\frac{\Gamma\left(\frac{r}{2}+1\right)}{(1+i)^{\frac{r}{2}+1}}$$

$$E[x^r] = \sigma\beta^{-r}\sum_{i=0}^{\infty}\binom{\sigma-1}{i}(-1)^i\frac{\Gamma\left(\frac{r}{2}+1\right)}{(1+i)^{\frac{r}{2}+1}}$$

$$\frac{dE[x^r]}{dr} = E[x^r\ln x] = -\sigma\beta^{-r}\ln\beta\sum_{i=0}^{\infty}\binom{\sigma-1}{i}(-1)^i\frac{\Gamma\left(\frac{r}{2}+1\right)}{(1+i)^{\frac{r}{2}+1}}$$

$$+\sigma\beta^{-r}\sum_{i=0}^{\infty}\binom{\sigma-1}{i}(-1)^i\frac{\Gamma'\left(\frac{r}{2}+1\right)\frac{1}{2}(1+i)^{\frac{r}{2}+1}-\Gamma\left(\frac{r}{2}+1\right)(1+i)^{\frac{r}{2}+1}\frac{1}{2}\ln(1+i)}{(1+i)^{r+2}}$$

If $r = 0$, then

$$E[\ln x] = -\sigma\ln\beta\sum_{i=0}^{\infty}\binom{\sigma-1}{i}(-1)^i\frac{\Gamma(1)}{(1+i)} + \sigma\sum_{i=0}^{\infty}\binom{\sigma-1}{i}(-1)^i\frac{\frac{1}{2}\Gamma'(1)-\frac{1}{2}\Gamma(1)\ln(1+i)}{(1+i)}$$

$$= \sigma \sum_{i=0}^{\infty} \binom{\sigma-1}{i} (-1)^i \frac{\frac{1}{2}\Gamma'(1) - \Gamma(1) \ln\left[\beta(1+i)^{\frac{1}{2}}\right]}{1+i}$$

Thus,

$$\begin{aligned} H_s(f) &= -\ln 2\sigma\beta^2 - \sigma \sum_{i=0}^{\infty} \binom{\sigma-1}{i} (-1)^i \frac{\frac{1}{2}\Gamma'(1) - \Gamma(1) \ln\left[\beta(1+i)^{\frac{1}{2}}\right]}{1+i} \\ &\quad + [\phi(\sigma+1) - \phi(1)] - \tfrac{1}{\sigma} + 1 \end{aligned}$$

Hence, the proof of Theorem 1 is completed.

References

1. Kundu, D.; Raqab, M.Z. Generalized Rayleigh distribution: Different methods of estimations. *Comput. Stat. Data Anal.* **2005**, *49*, 187–200. [CrossRef]
2. Surles, J.G.; Padgett, W.J. Inference for reliability and stress-strength for a scaled Burr Type X distribution. *Lifetime Data Anal.* **2001**, *7*, 187–200. [CrossRef]
3. Yan, W.A.; Li, P.; Yu, Y.X. Statistical inference for the reliability of Burr-XII distribution under improved adaptive Type-II progressive censoring. *Appl. Math. Model.* **2021**, *95*, 38–52. [CrossRef]
4. Jamal, F.; Chesneau, C.; Nasir, M.A.; Saboor, A.; Altun, E.; Khan, M.A. On a modified Burr XII distribution having flexible hazard rate shapes. *Math. Slovaca* **2020**, *70*, 193–212. [CrossRef]
5. Chang, H.; Ding, J.J.; Wu, L.N. Performance of EBPSK demodulator in AWGN channel. *J. Southeast Univ. Natural Sci.* **2012**, *42*, 14–19.
6. Fu, R.L.; Ma, Y.X.; Dong, G.H. Researches on statistical properties of freak waves in uni-directional random waves in deep water. *Acta Oceanologica Sin.* **2021**, *43*, 81–89.
7. Feng, Y.; Song, S.; Xu, C.Q. Two-parameter generalized Rayleigh distribution. *Math. Appl.* **2022**, *35*, 128–136.
8. Çolak, A.B.; Sindhu, T.N.; Lone, S.A.; Shafiq, A.; Abushal, T.A. Reliability study of generalized Rayleigh distribution based on inverse power law using artificial neural network with Bayesian regularization. *Tribol. Int.* **2023**, *185*, 108544. [CrossRef]
9. Shen, B.L.; Chen, W.X.; Wang, S.; Chen, M. Fisher information for generalized Rayleigh distribution in ranked set sampling design with application to parameter estimation. *Appl. Math. Ser. B* **2022**, *37*, 615–630. [CrossRef]
10. Shen, Z.J.; Alrumayh, A.; Ahmad, Z.; Abu-Shanab, R.; Al-Mutairi, M.; Aldallal, R. A new generalized rayleigh distribution with analysis to big data of an online community. *Alex. Eng. J.* **2022**, *61*, 11523–11535. [CrossRef]
11. Rabie, A.; Hussam, E.; Muse, A.H.; Aldallal, R.A.; Alharthi, A.S.; Aljohani, H.M. Estimations in a constant-stress partially accelerated life test for generalized Rayleigh distribution under Type-II hybrid censoring scheme. *J. Math.-UK* **2022**, *2022*, 6307435. [CrossRef]
12. Raqab, M.Z.; Kundu, D. Burr Type X distribution: Revisited. *J. Probab. Stat. Sci.* **2006**, *4*, 179–193.
13. Xu, R.; Gui, W.H. Entropy estimation of inverse Weibull distribution under Adaptive Type-II progressive hybrid censoring schemes. *Symmetry* **2019**, *11*, 1463.
14. Chacko, M.; Asha, P.S. Estimation of entropy for generalized exponential distribution based on record values. *J. Indian Soc. Prob. St.* **2018**, *19*, 79–96. [CrossRef]
15. Shrahili, M.; El-Saeed, A.R.; Hassan, A.S.; Elbatal, I.; Elgarhy, M. Estimation of entropy for Log-Logistic distribution under progressive Type-II censoring. *J. Nanomater.* **2022**, *2022*, 2739606. [CrossRef]
16. Wang, X.J.; Gui, W.H. Bayesian estimation of entropy for Burr type XII distribution under progressive Type-II censored data. *Mathematics* **2021**, *9*, 313. [CrossRef]
17. Al-Babtain, A.A.; Elbatal, I.; Chesneau, C.; Elgarhy, M. Estimation of different types of entropies for the Kumaraswamy distribution. *PLoS ONE* **2021**, *16*, e0249027. [CrossRef] [PubMed]
18. Bantan, R.A.R.; Elgarhy, M.; Chesneau, C.; Jamal, F. Estimation of entropy for inverse Lomax distribution under multiple censored data. *Entropy* **2020**, *22*, 601. [CrossRef]
19. Shi, X.L.; Shi, Y.M.; Zhou, K. Estimation for entropy and parameters of generalized Bilal distribution under adaptive Type-II progressive hybrid censoring scheme. *Entropy* **2021**, *23*, 206. [CrossRef] [PubMed]
20. Liu, S.H.; Gui, W.H. Estimating the entropy for Lomax distribution based on generalized progressively hybrid censoring. *Symmetry* **2019**, *11*, 1219. [CrossRef]
21. Zhou, X.D.; Qi, X.R.; Yue, R.X. Statistical analysis for Type-II progressive interval censored data based on Weibull survival regression models. *J. Syst. Sci. Math. Sci. Chin. Ser.* **2023**, *43*, 1346–1361.
22. Almarashi, A.; Abd-Elmougod, G. Accelerated competing risks model from Gompertz lifetime distributions with Type-II censoring scheme. *Therm. Sci.* **2020**, *24*, 165–175. [CrossRef]

23. Ramadan, D.A. Assessing the lifetime performance index of weighted Lomax distribution based on progressive Type-II censoring scheme for bladder cancer. *Int. J. Biomath.* **2022**, *14*, 2150018. [CrossRef]
24. Hashem, A.F.; Alyami, S.A. Inference on a new lifetime distribution under progressive Type-II censoring for a parallel-series structure. *Complexity* **2021**, *2021*, 6684918. [CrossRef]
25. Luo, C.L.; Shen, L.J.; Xu, A.C. Modelling and estimation of system reliability under dynamic operating environments and lifetime ordering constraints. *Reliab. Eng. Syst. Saf.* **2022**, *218*, 108136. [CrossRef]
26. Ren, J.; Gui, W. Statistical analysis of adaptive Type-II progressively censored competing risks for Weibull models. *Appl. Math. Model.* **2021**, *98*, 323–342. [CrossRef]
27. Zhou, S.R.; Xu, A.C.; Tang, Y.C.; Shen, L.J. Fast Bayesian inference of reparameterized Gamma process with random effects. *IEEE Trans. Reliab.* **2023**, 1–14. [CrossRef]
28. Zhuang, L.L.; Xu, A.C.; Wang, X.L. A prognostic driven predictive maintenance framework based on Bayesian deep learning. *Reliab. Eng. Syst. Saf.* **2023**, *234*, 109181. [CrossRef]
29. Wang, X.; Song, L. Bayesian estimation of Pareto distribution parameter under entropy loss based on fixed time censoring data. *J. Liaoning Tech. Univ. (Nat. Sci. Edit.)* **2013**, *32*, 245–248.
30. Renjini, K.R.; Abdul-Sathar, E.I.; Rajesh, G. A study of the effect of loss functions on the Bayes estimates of dynamic cumulative residual entropy for Pareto distribution under upper record values. *J. Stat. Comput. Simul.* **2016**, *86*, 324–339. [CrossRef]
31. Han, M. E-Bayesian estimations of parameter and its evaluation standard: E-MSE (expected mean square error) under different loss functions. *Commun Stat.-Simul. Comput.* **2021**, *50*, 1971–1988. [CrossRef]
32. Rasheed, H.A.; Abd, M.N. Bayesian Estimation for two parameters of exponential distribution under different loss functions. *Ibn Al-Haitham J. Pure Appl. Sci.* **2023**, *36*, 289–300. [CrossRef]
33. Hassan, A.S.; Elsherpieny, E.A.; Mohamed, R.E. Classical and Bayesian estimation of entropy for Pareto distribution in presence of outliers with application. *Sankhya A* **2022**, *85*, 707–740. [CrossRef]
34. Kohansal, A. On estimation of reliability in a multicomponent stress-strength model for a Kumaraswamy distribution based on progressively censored sample. *Stat. Pap.* **2019**, *60*, 2185–2224. [CrossRef]
35. Luengo, D.; Martino, L.; Bugallo, M.; Elvira, V.; Särkkä, S. A survey of Monte Carlo methods for parameter estimation. *Eurasip J. Adv. Signal Process.* **2020**, *2020*, 101186. [CrossRef]
36. Das, A.; Debnath, N. Sampling-based techniques for finite element model updating in bayesian framework using commercial software. *Lect. Notes Civ. Eng.* **2021**, *81*, 363–379.
37. Karunarasan, D.; Sooriyarachchi, R.; Pinto, V. A comparison of Bayesian Markov chain Monte Carlo methods in a multilevel scenario. *Commun. Stat.-Simul. Comput.* **2021**, *81*, 1–17. [CrossRef]
38. Raqab Mohammad, Z. Discriminating between the generalized Rayleigh and Weibull distributions. *J. Appl. Stat.* **2013**, *40*, 1480–1493. [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.